American Society of Human Genetics 62nd Annual Meeting
November 6-10, 2012 San Francisco, California

POSTER ABSTRACTS

The program and abstract/poster board number next to each listing is followed by a **W** (Wednesday), or **T** (Thursday), or **F** (Friday) to indicate the day on which authors must be present at their poster boards. Posters will remain on the boards for all three days (Wednesday through Friday).

Session Title/Topic Area	*Start #*	*End #*
Genome Structure, Variation and Function	#412	#600
Pharmacogenetics	#601	#655
Metabolic Disorders	#656	#753
Therapy for Genetic Disorders	#754	#804
Development	#805	#854
Cytogenetics	#855	#980
Cancer Genetics	#981	#1279
Statistical Genetics and Genetic Epidemiology	#1280	#1583
Cardiovascular Genetics	#1584	#1715
Genetic Counseling and Clinical Testing	#1716	#1796
Ethical, Legal, Social and Policy Issues in Genetics	#1797	#1850
Genetics Education	#1851	#1863
Health Services Research	#1864	#1877
Complex Traits and Polygenic Disorders	#1878	#2369
Psychiatric Genetics, Neurogenetics & Neurodegeneration	#2370	#2661
Molecular Basis of Mendelian Disorders	#2662	#2966
Prenatal, Perinatal and Reproductive Genetics	#2967	#3034
Clinical Genetics and Dysmorphology	#3035	#3250
Evolutionary and Population Genetics	#3251	#3414
Epigenetics	#3415	#3525
Bioinformatics and Genomic Technology	#3526	#3735

POSTER AUTHOR SCHEDULE

Refer to the schedule below for presentation times and for the poster mounting;shremoval schedule. **Posters should remain up all three days.**

Wednesday, November 7

10:00 am– 10:30 am	Authors place posters on boards. Posters to remain up until Friday
10:00 am– 4:30 pm	Posters open for viewing
2:15 pm– 4:15 pm	**Poster Session I (W)**
	2:15 pm– 3:15 pm ***(odd poster board numbers; author must be present)***
	3:15 pm– 4:15 pm ***(even poster board numbers; author must be present)***

Thursday, November 8

7:00 am– 4:30 pm	Posters open for viewing
2:15 pm– 4:15 pm	**Poster Session II (T)**
	2:15 pm– 3:15 pm ***(odd poster board numbers; author must be present)***
	3:15 pm– 4:15 pm ***(even poster board numbers; author must be present)***

Friday, November 9

7:00 am– 4:30 pm	Posters open for viewing
2:15 pm– 4:15 pm	**Poster Session III (F)**
	2:15 pm– 3:15 pm ***(odd poster board numbers; author must be present)***
	3:15 pm– 4:15 pm ***(even poster board numbers; author must be present)***
4:15 pm– 4:30 pm	**Authors must remove posters**
4:30 pm	Exhibit Hall and Posters close

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

412W

Capture of chromosome conformation changes in breast cancer cell lines. *C. Liu, C. Woo, P. Yim, S. Dong, J. Stevens.* MBS R & D, Life Technologies, Foster City, CA.

Global changes in the epigenetic landscape are a hallmark of cancer. The initiation and progression of cancer is now realized to involve epigenetic abnormalities along with genetic alterations.In this study, we use breast cancer, the most common cancer in women, as a model to capture the chromosome conformation changes and corresponding gene expression level changes. Two breast cancer cell lines compared are the estrogen receptor alpha positive, weakly invasive, luminal epithelial-like MCF-7 and the ERalpha-negative, P53 mutated, highly invasive, fibroblast-like MDA-MB-231. 3C (Chromosome Conformation Capture) allows us to analyze the folding of chromatin at a resolution beyond that provided by current microscopy techniques. With the 3C technique, we are able to identify the physical interactions between distant DNA segments and of chromatin loops that are formed. In this study, we focused on PLD1 that has been shown to involve in many critical cellular pathways, such as EGF, PDGF, insulin, serum as well as Arf, Rho, PKC etc. through the hydrolysis of phophatidylcholine to generate the lipid second messenger phosphatidic acid (PA) and choline. The region in this study spans 245 kb, including both gene body and upper regulatory regions. 1st introns as well as 2nd promoter regions were used as anchors to capture the long distance loop interactions. 3C mappings were compared between MCF-7 and MDA-MB-231. The differences in the interaction patterns were correlated with the PLD1 expression levels. In addition to the two cell lines comparison, both cell lines were treated with epigenetic agents 5-Aza-2'-deoxycytidine, the inhibitor of methyltransferase which results in DNA demethylation (hypomethylation) and Trichostatin (TSA), the histone deactylase inhibitors which disrupts normal chromatin structure. Cells were treated with 5-Aza-2'-deoxycytidine and TSA for 7 days and 2 days respectively. Chromosome conformation changes upon drug treatments were captured and compared between these two cell lines. Epigenetic alterations along with the cancer metastasis, such as chromosome conformation changes (3D changes) allows us to get better understanding of epigenetic mechanisms which lead to the potential therapeutic possibilities of epigenetic drugs, such as 5-Aza-2'-deoxycytidine and TSA. In the future, chromosome 3D changes can be used as signatures for both diagnostic and therapeutic treatments.

413F

Myc-induced anchorage of the rDNA intergenic spacer region to nucleolar matrix modulates growth-stimulated changes in higher-order rDNA architecture. *C. Shiue, A. Nematollahi-Mahani, A. Wright.* Clinical Research Center (KFC), Department of Laboratory Medicine and Center for Biosciences, Karolinska Institute, SE-141 86 Huddinge Sweden.

Chromatin domain organization and the compartmentalized distribution of chromosomal regions are essential for packaging of DNA in the eukaryotic nucleus as well as regulated gene expression. Nucleoli are the most prominent morphological structures of cell nuclei and nucleolar organization is coupled to cell growth. It has been shown that nuclear scaffold/matrix attachment regions often define the base of looped chromosomal domains in vivo and that they are thereby critical for correct chromosome architecture and gene expression. Here, we show regulated organization of mammalian rRNA genes into distinct chromatin loops by tethering to nucleolar matrix via the non-transcribed intergenic spacer region of the rDNA. The rDNA gene loop structures are induced specifically upon growth stimulation and are dependent on the activity of the c-Myc protein. Matrix attached rDNA genes are hypomethylated at the promoter and are thus available for transcriptional activation. Epigenetically silenced rDNA genes are not recruited to the matrix. c-Myc, which has been shown to induce rDNA transcription directly, is physically associated with the matrix-attached rDNA gene looping structures via the intergenic spacer sequence. Such a role of Myc proteins in gene activation has not been reported previously.

414W

CENcode: linear assembly and epigenetic annotation of human centromeres. *K.E. Hayden, J. Kent.* Center for Biomolecular Science and Engineering, University of California, Santa Cruz, CA 95064, USA.

Centromere sequences remain vastly underrepresented in our reference assemblies and, consequently, remain isolated and unaccountable in ongoing studies in the genome sciences. This disparity results from the challenge of predicting the linear order of multi-megabase-sized regions that are composed almost entirely of near-identical satellite DNA. The incomplete nature of the human genome complicates interpretation of short-read epigenetic datasets, and as a result centromere genomics remain largely isolated from studies involving kinetochore regulation and structure. Here, we present a linear prediction of human centromeres, representing a sequence organization and content in a single, reference individual. Using the human Y centromere as a model of this approach, we intersect short-read functional datasets, assigning available ENCODE sequence libraries within the context of a centromeric reference. As a result, we present a high-resolution study of centromeric sequences organization, thus presenting an opportunity to identify genomic features underlying sites of kinetochore assembly.

415F

Impact of genetic variation on genome-wide epigenetic profiles and gene expression phenotypes. *A. Reymond[1], H. Kilpinen[2], S. Waszak[3], A. Gschwind[1,4], E. Migliavacca[1,4], R.M. Witwicki[1], S. Raghav[3], A. Orioli[1], L. Romano-Palumbo[2], M. Wiederkehr[1], S. Turnherr[3], D. Hacker[3], M. Gutierrez-Arcelus[2], L.J. Core[5], J.T. Lis[5], N. Hernandez[1], B. Deplancke[3], E.T. Dermitzakis[2].* 1) Center for Integrative Genomics, University of Lausanne, Lausanne, Switzerland; 2) Department of Genetic Medicine and Development, University of Geneva Medical School, Geneva, Switzerland; 3) Ecole Polytechnique Federale de Lausanne, Lausanne, Switzerland; 4) Swiss Institute of Bioinformatics, Lausanne, Switzerland; 5) Department of Molecular Biology and Genetics, Cornell University, Ithaca, New York.

Recent advances in genome-wide profiling of chromatin state and transcription factor (TF) binding have identified specific chromatin signatures related to various classes of functional elements in different cell types. However, their genetic basis and degree of variability across individuals remain largely unknown. We have generated genome-wide enrichment profiles of i) DNA methylation, ii) RNA Pol II, TFIIB, MYC, and PU.1 DNA occupancy, iii) histone modifications H3K4me1, H4K20me1, H3K27me3, H3K27ac, and H3K4me3, in addition to transcriptome-wide mRNA, small-RNA and Global Run-On RNA sequencing data, from lymphoblastoid cell lines of two trios sequenced in high depth as part of the 1000 Genomes project. We analyzed the allele-specific effects of sequence variation on TF binding, histone modifications, and gene expression and discovered allele-specific signals for each of the assays (8–57% of accessible sites), many of which showed the same allelic direction of effect at overlapping loci (for e.g. H3K4me3 sites in RNA Pol II, rho=0.73). We additionally used the allele-specific information to study the parentally transmitted (genetic) effects in each assay, and discovered strong patterns of transmission of allelic effects from parents to child (for e.g. PU.1, MYC, and H3K4me3), ranging from 20% to 70% depending on the assay. We have extended our analysis to lymphoblastoid cell lines of 54 unrelated individuals from the 1000 Genomes CEU population for the majority of the assays to identify genetic effects affecting chromatin properties on a population level in a context of known regulatory variation. We are also exploring the combinatorial patterns of all 12 assays across individuals to further dissect the components of the general transcriptional state of the cells. Together, this study will significantly improve our understanding of the biological landscape around regulatory and other functional elements of the genome, and provide better means to interpret the heritability and molecular basis of phenotypic diversity, such as disease susceptibility, in humans.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

416W

Capturing long-range chromatin interactions of chromatin regulatory elements. *N. Heidari, D.H. Phanstiel, W. Soon, M.P. Snyder.* Genetics, Stanford, Palo Alto, CA.

There is increasing evidence to suggest that long-range interactions between specific chromosomal regions play important roles in a diverse array of biological phenomena including immune response, gene regulation, and the development of human disease. Chromosomal folding brings together distal regulatory elements including enhancers, promoters, and insulators and these interactions have been shown to correlate with gene expression as well as translocation breakpoints in human cancer. However, methods to study these interactions have only recently been developed and are rapidly evolving. Chromatin conformation capture-based methods (3C, 4C, 5C, 6C, and Hi-C) allow for the detection of genome-wide chromosomal interactions while another variant, Chromatin Interaction Analysis by Paired-End Tag Sequencing (ChIA-PET), includes a chromatin immunoprecipitation step, and focuses on detection of interacting regulatory elements bound by a specific transcription factor or chromatin-associated protein. We have developed a robust ChIA-PET workflow as well as novel data analysis pipeline to characterize long-range chromatin interactions. A novel ChIA-PET analysis pipeline, chiapetter.py, was written to analyze the data. The pipeline was written in the programming language python and can be executed on either a desktop computer or a linux cluster for increased throughput. The pipeline consists of five major steps including: “linker parsing”, “mapping”, “(paired end tag) PET characterization”, “peak calling”, and “interaction calling”. It uses random shuffling of PETs to filter interactions to a fixed false discovery rate of 5%. This workflow has been applied to characterize protein-mediated long-range interactions of several regulatory factors as well as nuclear matrix proteins in chronic myelogenous leukemia (K562) as a workhorse cell line. Each experiment was performed in duplicates and was acquired at an average read depth of 99.5 million PETs per experiment. The vast majority of the peaks detected by ChIA-PET were also detected by a ChIP-Seq. Analysis of the RAD21 data set revealed 5,657 high confidence (FDR = 0.05) interactions comprising 5231 intrachromosomal and 336 interchromosomal interactions. With this methodology in place, we are prepared to explore biological questions including chromatin structural and functional variation in cancer, and other disease biology.

417F

PRDM9 directs genetic recombination away from functional genomic elements. *K. Brick[1], F. Smagulova[2], P. Khil[1], RD. Camerini-Otero[1], G. Petukhova[2].* 1) Genetics & Biochemistry Branch, NIDDK, National Institutes of Health, Bethesda, MD; 2) Uniformed Services University of Health Sciences, Department of Biochemistry and Molecular Biology, Bethesda, MD, USA.

Recombination initiates with the formation of programmed DNA double strand breaks (DSBs) at a small subset of genomic loci called hotspots. Elegant recent studies in mouse and human have determined that PRDM9, a meiosis-specific histone H3 methyl-transferase is involved in DSB hotspot site determination (Parvanov et al. Science 2010; Baudat et al., Science 2010; Myers et al., Science 2010), likely thorough DNA binding of its zinc-finger domain. We have recently generated the first genome-wide DSB hotspot map in a metazoan genome and have shown that the majority of mouse DSB hotspots are associated with testis-specific H3K4me3 chromatin marks, potentially formed by PRDM9 (Smagulova et al., Nature 2011). Curiously however, Prdm9 knockout mice remain proficient at initiating recombination. In this work, we describe several straightforward experiments that elucidate the nature and extent of the role of PRDM9 in determining DSB hotspots locations.

We used a novel ChIP-Seq variant developed by our group to detect ssDNA bound by the meiotic recombinase DMC1 (Khil et al., Genome Res., 2012). Using this method, we precisely mapped the genome wide distribution of DSB hotspots in seven mouse strains and in their F1 progeny. While hotspots in mice sharing a *Prdm9* allele mapped to almost identical loci, hotspots in other mice were dependent on the DNA binding specificity of the *Prdm9* allele. Importantly, in *Prdm9* knockout mice, hotspots were at completely different locations than in wild-type, definitively illustrating that PRDM9 determines practically all DSB hotspot locations. Intriguingly, DSBs in the pseudo autosomal region - the site of an obligate recombination event in every meiosis - were found to be *Prdm9*-independent and present in all strains. In *Prdm9* knockout mice, DSBs still accumulated in hotspots however, in the absence of PRDM9, most recombination initiated at H3K4me3 marks at promoters or enhancers. These sites are rarely targeted in wild-type mice illustrating an important, unexpected role for PRDM9 in sequestering the recombination machinery away from functional genomic elements where the efficient repair of DSBs may be problematic.

418W

ATRX interacts with the MRN complex to play a key role in the replication of non-canonical DNA structures associated with tandem repeats. *R. Gibbons, D. Clynes, C. Jelinska, H. Ayyub, D.R. Higgs.* Weatherall Institute of Molecular Medicine, University of Oxford, Oxford, United Kingdom.

ATRX is an X-linked gene, originally identified in patients with a form of syndromal mental retardation (ATR-X syndrome), which encodes a protein of the SWI/SNF family. ATRX preferentially binds tandem repeats (TRs) at rDNA, satellite repeats and telomeres. Within euchromatin ATRX also binds to TRs associated with a subset of genes, some of which are dysregulated when ATRX is mutated and the change in expression is determined by the size of the TR, producing skewed allelic expression. This reveals the characteristics of the affected genes, explains the variable phenotypes seen with identical ATRX mutations, and illustrates a new mechanism underlying variable penetrance in human genetic disease. Many of the TRs are G rich and predicted to form non-B DNA structures (including G-quadruplex) *in vivo*. Since ATRX and its partner DAXX are involved in the deposition of the variant histone H3.3 at telomeres and pericentromeric heterochromatin, it has been proposed that during nuclear processes such as transcription and replication ATRX influences the formation and/or resolution of non-canonical DNA structures by modifying their nucleosomal context. Consistent with this, loss of ATRX is associated with a marked decrease in H3.3 and an increased DNA damage response at telomeres during S-phase of the cell cycle when unresolved G4 DNA may impede DNA replication. Here we show that ATRX co-localises and interacts with the MRN complex (RAD50, MRE11 and NBS1) that is known to play a central role in sensing and resolving DNA damage during replication. In the absence of ATRX we find that DNA replication is perturbed with an increased frequency of stalled replication forks. This in turn is associated with an increased frequency of double strand breaks and a subsequent DNA damage response. It appears therefore that a key role of ATRX is to facilitate replication through TR sequences, via an interaction with the MRN complex. At telomeres, where the tandem TTAGGG repeats are replicated in a unidirectional manner, stalled forks can only be rescued by recombination, consistent with the observation that many cell lines and tumors that use recombination to maintain viable telomeres (alternative lengthening of telomere [ALT] phenotype) have mutations in the *ATRX* gene. In keeping with the proposed role of ATRX and MRN in replication and repair, here we also show that re-expression of ATRX in ALT cell lines reverses their characteristic phenotype.

419F

Genome-wide copy number variation analysis of a Branchio-Oto-Renal syndrome cohort identifies a recombination hotspot associated with an EYA1 deletion and novel candidate genomic regions. *F. Alasti[1], P. Brophy[2], B. Darbro[2], J. Dierdorff[3], C. Nishimura[1], B. Cobb[3], J. Clarke[2], M. Hakeman[3], A. Bassuk[2], R.J.H Smith[1], J.R. Manak[2,3,4].* 1) Department of Otolaryngology, University of Iowa, Iowa City, IA 52242, USA; 2) Department of Pediatrics, Carver College of Medicine, University of Iowa, IA 52242, USA; 3) Department of Biology, University of Iowa, IA 52242, USA; 4) Carver Center for Genomics, University of Iowa, IA 52242, USA.

Branchio-oto-renal (BOR) syndrome is an autosomal dominant disorder characterized by branchial cleft cysts, external ear abnormalities, preauricular pits or tags and hearing loss accompanied by malformations of the kidney. Both point mutations and large/small deletions in EYA1 are identified in about 40% of patients with the BOR phenotype. Mutations in SIX1 and SIX5 are also associated with BOR syndrome. Array-based comparative genomic hybridization (aCGH) has been used successfully to identify CNVs associated with a large number of human diseases, although this technique has only been used to a limited extent for BOR affecteds. We thus performed high-resolution aCGH on a cohort of 35 BOR probands negative for mutations in the three known BOR-causing genes. In this cohort, we identified >1,000 rare and novel CNVs, including an ~2.7 Mb heterozygous deletion, with nearly identical breakpoints in 4 unrelated BOR probands that removes all of EYA1 and multiple downstream genes. Both the distal and proximal deletion breakpoints are flanked by LTR-ERV_class I retrotransposon elements, which creates a recombination hotspot for non-allelic homologous recombination (P = 0.009). In a fifth BOR proband, we identified a novel 5.7 kb deletion that removes the last exon of EYA1 and the entire 3' UTR. In total, 5 probands (14%) had deletions of all or part of EYA1; further demonstrating that deletion in this gene is rather common in BOR. In addition to the deletions affecting EYA1, several novel deletions in other genomic regions were identified in this patient cohort, suggesting the involvement of other genes in the pathogenesis of BOR. Multiple methods were used to prioritize potential BOR-associated genes affected by the novel deletions, including mining protein-protein interaction data to create a hypothetical human EYA1 interactome network. These separate networks were then combined into one large network that was significantly enriched for genes involved in both morphogenesis of the metanephric bud and development of the inner ear, providing strong evidence for novel candidate disease-causing genes associated with BOR.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

420W

Delineation of the reliability of in silico copy number variation (CNV) calls from different Illumina SNP arrays. *S. Herms[1,2], L. Priebe[1,2], F. Degenhardt[1,2], M.M. Noethen[1,2,3], S. Cichon[1,2,4], P. Hoffmann[1,2]*. 1) Department of Genomics, Life & Brain Center, University of Bonn, Bonn, Germany; 2) Institute of Human Genetics, University of Bonn, Bonn, Germany; 3) German Center for Neurodegenerative Disorders (DZNE), Bonn, Germany; 4) Institute of Neuroscience and Medicine (INM-1), Structural and Functional Organization of the Brain, Genomic Imaging, Research Center Juelich, Juelich, Germany.

Illumina SNP arrays are widely used to identify common and rare susceptibility variants involved in the etiology of multifactorial diseases. Besides SNP-based genome wide association studies, SNP-arrays allow to in silico analyze Copy Number Variants (CNV). There has been a comprehensive discussion in the field to what extent in silico CNV calls are reliable. For individual associated CNVs, validation by quantitative PCR (qPCR) is usually performed which is labour-intense and requires large amounts of DNA. Moreover, it is almost impossible to qPCR-validate CNVs from in silico burden analyses which test the genome-wide frequency of CNVs between patients and controls. In order to get an objective overview of the reliability of in silico CNV calls and the influence of batch effects, we compared in silico CNV calls from a larger number of technical replicates which were genotyped as internal quality controls in different genotyping projects over the last three years. These data had been acquired on different Illumina arrays such as the Human660W, HumanOmniExpress, HumanOmni1S and HumanOmni1M-Quad. We compared these replicates on basis of genotypes and copy number variants to evaluate technical artifacts arising from batch effects. We will present data on deviances arising from quality parameters (e. g. gentrain score, logR ratio, B -allele frequency, intensity, chemistry batches) and the implications thereof.

421F

The Kaiser Permanente/UCSF Genetic Epidemiology Research Study on Adult Health and Aging: Detecting Copy Number Variation in a multi-ethnic cohort of 100,000. *C. Sabatti[1], Z. Zhang[2], Y. Banda[3], M. Kvale[3], T. Hoffman[3], S. Hesselson[3], H. Tang[4], P-Y. Kwok[2], C. Schaefer[5], N. Risch[3,5]*. 1) Department of Health Research and Policy, Stanford University, Stanford, CA; 2) Department of Statistics, UCLA, Los Angeles, CA; 3) Institute of Human Genetics, UCSF, San Francisco, CA; 4) Department of Genetics, Stanford University, Stanford, CA; 5) Division of Research, Kaiser Permanente, Oakland, CA.

The Genetic Epidemiology Research Study on Adult Health and Aging (GERA) cohort includes genotype information on over 100,000 individuals. Four ethnic-specific Axiom arrays were used to perform genome-wide genotyping on these subjects, including substantial numbers of individuals of European, Native American, Latino and East Asian background. Copy Number Variant (CNV) specific probes were not included in the array design, but information about DNA copy number can still be obtained from the raw signals of the genotyping reaction. We developed methods for detecting copy number variation from these data, and present the results of two focused analyses aiming to estimate the frequencies of CN polymorphisms in the four ethnic groups under study and the frequency of large sporadic or rare copy number variant events. For the first analysis, we worked with a set of CNP regions documented in the Database of Genomic Variants (http://dgvbeta.tcag.ca/dgv/app/home?ref=NCBI36/hg18) and dbVar (http://www.ncbi.nlm.nih.gov/dbvar/) that were covered by at least 10 SNPs on the Axiom platform. We inferred copy numbers for subjects from each of the ethnic groups under study and compared the estimated CNP frequencies across populations. For the second analysis, we identified subjects likely to harbor large CNVs by screening homozygosity patterns. We further analyzed raw signal from the genotyping experiments to confirm the presence of CNVs. Finally, we linked the information on the identified CNPs and CNVs to the extensive health-related phenotypic information available on these subjects in association analyses to detect copy number relationships with a variety of disease and risk factor outcomes.

422W

Genomic copy number variation affects brain structure volumes in schizophrenia. *T.H. Wassink[1], J. Manak[2], D. Rudd[1], E. Epping[1], F. Fleming[1], G. Zeien[1], S. Ziebell[1], B.C. Ho[1], N.C. Andreasen[1]*. 1) Dept Psychiatry, Univ Iowa Col Med, Iowa City, IA; 2) Dept Biology, Univ Iowa Col Med, Iowa City, IA.

BACKGROUND: Genomic copy number variation (CNV) is an important source of susceptibility variation for schizophrenia, accounting for ~10% of cases. Effects of CNVs on deeper aspects of the schizophrenia phenotype, however, have not been extensively investigated. We focused on brain structure, hypothesizing that patients with schizophrenia and disease-susceptibility CNVs would have more severe brain tissue deficits than patients with schizophrenia without such CNVs. METHODS: We screened for CNVs using the Nimblegen 2.1 million probe genome-wide microarray in 174 individuals with schizophrenia from whom we had previously obtained structural magnetic resonance (MR) brain imaging data. We used the Nexus suite of analysis tools to identify CNVs from the array data. CNVs were called “abnormal” if they were either a) known disease CNVs or b) novel, affected genes, and were > 500kb if deletions or > 1MB if duplications. CNVs were called “likely deleterious” if they were novel, affected genes, and were > 50 kb. The effects of CNVs were tested on cerebral lobe gray and white matter volumes and on cerebro-spinal fluid volume using ANCOVAs with appropriate covariates. RESULTS: 15 (8.6%) of patients had disease-susceptibility CNVs and 18 (10.3%) had suggestive CNVs. These included known CNVs such as 15q11.2 and 1q21.1 deletions, and numerous novel CNVs. ANCOVAs showed that patients with suggestive and abnormal CNVs showed similar patterns of brain volume differences in comparison to the non-CNV cases. Taken as a group, patients with CNVs had decreased frontal GM ($F=4.66$, $p=0.03$), frontal WM ($F=4.35$, $p=0.04$), temporal GM ($F=7.41$, $p=0.007$), and temporal WM ($F=9.27$, $p=0.003$) volumes in comparison to non-CNV patients. DISCUSSION: These data show first that CNVs defined as likely to be deleterious were associated with patterns of brain structure deficits that were similar to CNVs defined as abnormal. Second, the CNV cases taken as a group show a characteristic pattern of brain structure deficits that is distinct from non-CNV cases and that may suggest a helpful approach to subgrouping schizophrenia cases.

423F

The use of matepair sequencing in the detection of structural variation in patients with ID and/or congenital anomalies. *S. Vergult[1], E. Van Binsbergen[2], T. Sante[1], S. Nowak[1], O. Vanakker[1], S. Janssens[1], K. Claes[1], B. Poppe[1], N. Van der Aa[3], F. Roelens[4], A. De Paepe[1], F. Speleman[1], E. Cuppen[2], W. Kloosterman[2], B. Menten[1]*. 1) Center for Medical Genetics, Ghent University Hospital, Ghent, Belgium; 2) Department of Medical Genetics, University Medical Center Utrecht, Utrecht, The Netherla; 3) Department of Medical Genetics, UA, Antwerp, Belgium; 4) Heilig Hart Ziekenhuis, Roeselare, Belgium.

Background: Structural variations have been recognized as an important cause of intellectual disability and congenital malformations for many years. The past decade array CGH has emerged as the technique of choice for the detection of copy number variants in patients with ID and congenital malformations. Despite the many advantages of array CGH (much higher resolution and amenity for automation), it has two major drawbacks: it does not provide any information on balanced rerarrangements, nor does it provide information of the genomic architecture of duplications/insertions. With the advent of next generation sequencing, matepair sequencing capable of detecting both unbalanced as well as balanced aberrations has been developed.

Methods: Matepair sequencing was performed on 56 patients with a chromosomal phenotype. Of these, 36 patients were shown to have a structural rearrangement detected with conventional karyotyping and/or array CGH. Twenty patients with a normal array profile were included.

Results: In all patients for whom an unbalanced aberration was detected with array CGH, the aberration was confirmed by coverage analysis of the matepair sequencing data. For the detected duplications and complex rearrangements the genomic nature of the aberration was revealed by cluster analysis, except for those caused by LCR mediated NAHR. Balanced rearrangements were detected in six out of eight selected patients with (apparently) balanced aberrations. Due to the fact that one of the breakpoints resides in a highly repetitive region, the aberration was not identified in the remaining two patients. In the patients with a normal array profile, no balanced or unbalanced pathogenic aberrations were detected by matepair sequencing.

Conclusion: For the first time we show that matepair sequencing can be used in a diagnostic setting to detect structural variations. All unbalanced rearrangements identified by array CGH were confirmed by matepair sequencing and the main advantage over array CGH, the detection of (apparently) balanced aberrations, is also clearly underscored by this study. Nevertheless, in the patients with no detected aberrations by array CGH, no causal aberration was detected by matepair sequencing.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

424W

At the crossroads of structure and function: SNP genotypes indexed by copy number yield novel eQTLs. *L.K. Davis, E.R. Gamazon, N.J. Cox.* Section Genetic Medicine, The University of Chicago, Chicago, IL.

Current analysis methodologies used in genome-wide association studies (GWAS) (at any given autosomal locus) collapse the homozygous (i.e., A/A), hemizygous (i.e., A/0) and duplicative (i.e., A/A/A) genotype states. This approach assumes that these classes of genotype are equivalent and treats the genotype variable itself as a whole and single variable that does not contain components. However, the knowledge of large-scale copy number variation (CNV) throughout the genome clearly suggests that component complexity exists within the genotype construct. We developed a simple local covariate approach that incorporates CNV status of SNP genotype into a standard regression model with the resulting analysis yielding effect sizes and p-values for 1) SNP genotype, 2) CNV genotype and 3) copy number indexed SNP genotype. We applied this method to an expression quantitative trait loci (eQTL) analysis of HapMap lymphoblastoid cell lines (LCLs) and found that by properly accounting for allelic dosage at deletion sites we dramatically increased power to detect associations in genome wide analysis of eQTLs. Results remained robust after removing regions of the genome known to harbor significant copy number and eQTL complexity (e.g., HLA region). In total, the results yielded improved p-values for 8,172 eQTL SNPs. Moreover, 101 completely novel additional trans eQTLs, that did not reach statistical significance in a previous standard GWAS, were discovered. This finding carries with it implications for the analysis of all GWAS and suggests that at least some portion of missing heritability may be recovered by indexing SNPs according to copy number status.

425F

Extensive genetic variation in somatic human tissues. *M. O'Huallachain[1,2], K.J. Karczewski[1,3], S. Weissman[4], A.E. Urban[1,5], M. Snyder[1].* 1) Genetics Department, Stanford University, Stanford, CA; 2) Department of Molecular, Cellular and Developmental Biology, Yale University,; 3) Biomedical Informatics Training Program, Stanford University School of Medicine, Stanford, CA; 4) Department of Genetics, Yale University, New Haven, CT; 5) Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, CA.

Genetic variation between individuals has been extensively investigated, but differences between tissues within individuals are far less understood. It is commonly assumed that all healthy cells that arise from the same zygote possess the same genomic content with a few known exceptions in the immune system and germ line. However, a growing body of evidence shows that genomic variation exists between differentiated tissues. We investigated the scope of somatic genomic variation between tissues within humans. Analysis of copy number changes by high-resolution aCGH in diverse tissues from six subjects reveals a significant number of intra-individual genomic changes between tissues. Many of these events affect genes. Our results have important consequences for understanding normal genetic and phenotypic variation within individuals, and they have significant implications for both the etiology of genetic diseases such as cancer, and for immortalized cell lines that might be used in research and therapeutics.

426W

Copy Number Variations in a cohort of Brazilian Sickle Cell Anemia Patients with and without Cerebrovascular Accident. *G. Ananina[1], F. Menaa[1,2], M.A. Bezerra[3], A.S. Araujo[3], P.R.S. Cruz[1], G.P. Gil[1], F.F. Costa[2], M.B. Melo[1,2].* 1) Mol Biol and Genet Eng (CBMEG), Universidade Estadual de Campinas (UNICAMP), Campinas, São Paulo State, Brazil; 2) Hematology and Hemotherapy Center, University of Campinas/Hemocentro-Unicamp (Campinas, São Paulo State, Brazil); 3) Hematology and Hemotherapy Center of Pernambuco/HEMOPE (Recife, Pernambuco State, Brazil).

Backgound - Sickle cell anemia (SCA) is a blood disorder caused by the homozygous state of a single point mutation in the beta globin gene resulting in an abnormal form of hemoglobin (HbSS). SCA is characterized primarily by chronic anemia, however, severity varies considerably among individuals. The present study aims to investigate the association of genetic factors, particularly copy number variations (CNVs), with cerebrovascular accident (CVA), one of the most severe complications in SCA patients. **Materials and Methods** - SCA patients with and without CVA were recruited at two Brazilian centers of hematology: Hemocentro (Campinas, SP) and HEMOPE (Recife, PE). Recruitment methods and blood collection were approved by the Ethics Committee of FCM-UNICAMP (Campinas/SP, Brazil). Genotyping protocol was carried out using Affymetrix® Genome-Wide Human SNP 6.0 Array (Affymetrix Inc., CA, USA) following manufacturer's recommendations. CNVs calling was performed with Affymetrix Genotyping Console software. Only samples that passed the recommended control values: contrast QC > 0.4; median absolute intensity pairwise differences MAPD < 0.3, were used in the analysis. **Results** - An exploratory analysis of distribution of copy number (CN) events was performed to compare CVA-affected (n=18) and non-affected (n=16) subjects. The analysis was applied to all autosomal CN events (gains and losses) larger than 1 kb. We observed the following distribution of CN events per individual in the affected and non-affect cohorts, respectively: *max* number - 1927/576; *min* number - 107/88; mean - 583/256. Hence, the total number of autosomal CNVs per individual were higher in CVA-affected individuals than in the non-affected cohort. The augment of CNVs observed in the CVA-affected cohort occurred primarily due to the increase in the number of relatively small CN events (of the size less than 100 kb). The group of affected individuals also demonstrated higher variance in the number of CN events. **Conclusions** - Pathophysiology of CNVs is still poorly understood and copy number genomic disbalance potentially can affect expression of dosage sensible genes thus influencing complex phenotypic traits. As the next step we plan to expand our sample size and refine the analysis. **Acknowledgments**: CNPq, Brazil (573780/2008-0, 151463/2009-3, 158367/2011-1); FAPESP, São Paulo, Brazil (2008/57441-0); Laboratório Nacional de Luz Síncrotron (LNLS), Campinas-São Paulo, Brazil.

427F

Prevalence and effects of the obesity associated deletion on chromosome 16p11.2 in the Swedish Obese Subjects (SOS) study. *J.C. Andersson-Assarsson, S. Romeo, L. Sjöström, L.M.S. Carlsson.* Department of Molecular and Clinical Medicine and Sahlgrenska Center for Cardiovascular and Metabolic Research, Sahlgrenska Academy at Gothenburg University, Vita Stråket 15, 41345 Gothenburg, Sweden.

Common obesity results from a combination of environmental and genetic factors. Studies have shown a very high heritability, up to 70%, for many obesity-related traits, including BMI and body fat. Much effort has been made to identify candidate genes and gene variants in common obesity, with some notable successes such as the discovery of the *FTO* gene. However, the majority of the heritability remains to be explained. Copy number variants (CNV) explain part of this “missing heritability”. They are an important source of genetic variability between individuals and have been associated with normal phenotypic diversity as well as with common diseases, including obesity. We and others have previously identified a rare CNV (a deletion) on chromosome 16 that give rise to a highly penetrant form of obesity. The aim of this study was to determine the frequency of the deletion and to evaluate its effects on baseline characteristics in the Swedish Obese Subjects (SOS) study. The SOS study started in 1987 and is a prospective, controlled, intervention study involving 4047 obese individuals; 2010 individuals have undergone bariatric surgery and 2037 conventional treatment (matched control group). DNA was available for 3678 individuals (1835 surgery cases (548 males, 1289 females) and 1843 controls (548 males, 1295 females)). Using genotyping and MLPA, we identified 14 individuals carrying the deletion (0.4%). This frequency is somewhat lower than in our previous study. However, the deletion is also associated with autism, schizophrenia and developmental delay and since psychiatric disabilities were an exclusion criteria in SOS we believe this can account for the lower frequency compared to previous studies. Of the 14 carriers of the deletion (6 males, 8 females), 9 underwent bariatric surgery (4 males, 5 females). Compared to individuals not carrying the deletion, baseline characteristics of the carriers of the deletion differed in terms of anthropometry, metabolism and social factors.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

428W

Molecular study of the inverted repeats responsible for triplication of the PLP1 locus. *C.R. Beck[1], C.M.B. Carvalho[1], J.R. Lupski[1,2,3].* 1) Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX; 2) Department of Pediatrics, Baylor College of Medicine, Houston, TX; 3) Texas Children's Hospital, Houston, TX.

Pelizaeus-Merzbacher disease (PMD, MIM #312080) is most commonly caused by an increased dosage of the gene PLP1. Recently, the presence of duplication-inverted triplication-duplication (DUP-TRP/INV-DUP) structures for this region has been described, and triplication of PLP1 that is flanked by duplications has also been observed. The common feature of these complex genomic rearrangements is the presence of a pair of inverted low copy repeats (LCRs) presumably mediating the rearrangement, with one of the breakpoints located within the inverted LCRs. Although the hypothesized mechanism for formation of this complex structure is DNA replication fork collapse followed by microhomology mediated template switching (FoSTeS/MMBIR) and break-induced replication (BIR) at a region of ectopic homology, to date no formal experimental demonstration at the PLP1 locus has been performed. Therefore, we sought to further characterize the inverted repeats responsible for the formation of the structure of DUP-TRP/INV-DUP, as well as to obtain further insights into the molecular mechanism underlying such alterations. Our observations indicate that the LCRs that mediate the DUP-TRP/INV-DUP rearrangements have predisposed the region for inversion, which is highly prevalent in the human population. Indeed, the human genome structural variation track of the UCSC genome browser indicates that 8 of 9 individuals contain fosmids on which paired-end sequencing indicates that they carry inversions at this location. Therefore, the hypothesized breakpoint for the BIR event within the LCRs could occur at two different locations depending on the original structure of the region. We have implemented further experimental analyses elucidating the structure and sequence of the rearrangements within the LCRs for four patients with PMD who carry DUP-TRP/INV-DUP rearrangements.

429F

Novel sequence-based CNV detection framework allows fine-mapping of the CFH region in 150 Chinese Singaporeans with Age-related Macular Degeneration. *E. Bellos[1], L. Coin[1], T.Y. Wong[2], C.Y. Cheng[2], G. Cheung[2], M. Hibberd[3], V. Kumar[3], S. Davila[3].* 1) School of Public Health, Imperial College London, London, United Kingdom; 2) Singapore Eye Research Institute, Singapore; 3) Genome Institute of Singapore, Singapore.

Copy Number Variation (CNV) is pervasive in the human genome and has been shown to play a causal role in genetic diseases. Recent advances in sequencing technologies have provided the means for identifying genomic variation at an unprecedented resolution. A single next-generation sequencing (NGS) experiment offers a multitude of features that can be used to extract CNV signatures. These features comprise read depth, read pair distance and split reads, and exhibit varying degrees of sensitivity for detecting different types of variants. Achieving a high sensitivity across the CNV spectrum requires taking into account the strengths and weaknesses of each feature and incorporating them into a unified model. Here we present cnvHiTSeq, an integrative method for CNV discovery and genotyping that jointly analyzes multiple features of sequence data at the population level. cnvHiTSeq utilizes a population-haplotype framework to incorporate diverse data sources into a single probabilistic model. An hidden Markov model is used to capture the spatial properties of CNV across a single chromosome copy. To account for sample-independent variation in the data, cnvHiTSeq uses the population distribution of measurements to update the parameters of the model. By organically combining evidence from multiple features, cnvHiTSeq achieves sensitive and precise discovery of all CNV classes even from low-coverage data. Using data from the trio and low-coverage phases of the 1000 Genomes Project, our method identified considerably more variants than competing methods, while maintaining a low false discovery rate (FDR). cnvHiTSeq detects ~80% of all CNVs greater than 100bp with a FDR of 4.3% (9.8%) for deletions (duplications). Our unified approach also benefits from population-level modelling to achieve a high CNV genotyping accuracy of 98.2%. We applied cnvHiTSeq to a cohort of 150 Chinese Singaporeans with Age-related Macular Degeneration (AMD). The cohort comprises high-coverage sequence data of the extended 200kb CFH region. Though previously associated with AMD, this region remains hard to characterize as it contains complex patterns of CNV. The high-resolution obtained by cnvHiTSeq allowed us to fine-map CNV in the region and elucidate its complexity. Thus, we have demonstrated that cnvHiTSeq offers a complete solution to sequence-based CNV detection and genotyping, aiming to further our understanding of CNV impact on disease and evolution.

430W

Incidental copy-number variants identified by routine genome testing in a clinical population. *P.M. Boone, Z.T. Soens, I.M. Campbell, P. Stankiewicz, S.W. Cheung, A. Patel, A.L. Beaudet, S.E. Plon, C.A. Shaw, A.L. McGuire, J.R. Lupski.* Dept. of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA.

Purpose: Mutational load of susceptibility variants has not been studied on a genomic scale in a clinical population, nor has the potential to identify these mutations as incidental findings, i.e. not related to the patient's current diagnosis and contributing to the “incidentalome”, during clinical testing been systematically ascertained. Methods: We analyzed data from chromosome microarray analysis (CMA), a method for genome-wide detection of DNA copy-number variations (CNVs) including deletion and duplication mutations, performed on DNA from 9,005 patients, fetuses, and their family members. CNVs encompassing or disrupting single genes were identified computationally and analyzed for their potential to confer predisposition to dominant, adult-onset disease. Multi-gene CNVs affecting dominant, adult-onset cancer predisposition syndrome genes were also assessed. Results: Of 1,812 single-gene CNVs present in our cohort, 83 occurred in 40 unique genes associated with dominant, adult-onset disorders and unrelated to the patients' referring diagnoses (i.e. incidental). Fourteen (17%) of these mutations are likely disease-predisposing, 25 (30%) are likely benign, and 44 (53%) are of unknown clinical consequence. When incidental CNVs spanning up to 20 genes were considered, 27 CNVs affected 17 unique genes associated with dominant, adult-onset cancer predisposition syndromes. Conclusion: CNV mutations potentially conferring susceptibility to adult-onset disease can be identified as incidental findings during routine genome-wide testing. Some mutations in known disease genes may be medically actionable, enabling disease surveillance or prevention; however, most incidentally observed single-gene CNVs are currently of unclear significance to the patient.

431F

Parental age and burden of genomic copy number variation. *J. Buizer-Voskamp[1], H.M. Blauw[2], M.P.M. Boks[1], KR. Van Eijk[1,3], J.H. Veldink[2], E.A.M. Hennekam[3], J.A.S. Vorstman[1], F. Mulder[3], H. Tiemeier[4,5], A.G. Uitterlinden[4,6], L.A. Kiemeney[7], L.H. Van den Berg[2], R.S. Kahn[1], C. Sabatti[8], R.A. Ophoff[1,9].* 1) Rudolf Magnus Institute of Neuroscience, Department of Psychiatry, University Medical Center Utrecht, The Netherlands; 2) Rudolf Magnus Institute of Neuroscience, Department of Neurology, University Medical Center Utrecht, The Netherlands; 3) Department of Medical Genetics, University Medical Center Utrecht, The Netherlands; 4) Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands; 5) Department of Child and Adolescent Psychiatry, Erasmus Medical Center and Sophia Children's Hospital, Rotterdam, The Netherlands; 6) Department of Internal Medicine, Erasmus Medical Center, Rotterdam, The Netherlands; 7) Department of Urology, Radboud University Nijmegen Medical Center, Nijmegen, The Netherlands;; 8) 11Department of Health Research and Policy, Stanford University School of Medicine, Stanford, California, USA; 9) Center for Neurobehavioral Genetics, University of California, Los Angeles, California, USA.

Genomic copy number variations (CNVs) and increased parental age are both associated with the risk to develop neuropsychiatric disorders. As advanced parental age is associated with increase of germline *de novo* mutations, we studied the relationship between parental age and various types of CNV burden in a large population sample (n=6 773) from The Netherlands. Analysis of global measures as well as proxies for *de novo* CNV events in this unique sample, however, did not support a correlation between parental age and CNV burden in the offspring. While it remains possible that local genomic effects may exist for specific phenotypes, our results strongly suggest that global CNV burden and advanced parental age are independent risk factors for psychiatric disorders.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

432W

Genome-wide analyses of recurrent genomic rearrangements - identification of three novel genomic disorders. *P. Dittwald[1,2,3], T. Gambin[4], P. Szafranski[1], J. Li[1], R.S. Amato[5], M. Divon[6], K. Maliszewski[6], E.E. Dolgado Bohorquez[7], L. Elton[8], A.C-H. Tsai[9], D. Scott[1], S.-H.L. Kang[1], A.M. Breman[1], S.R. Lalani[1], C. Bacino[1], W. Bi[1], A. Milosavljevic[1], J.R. Lupski[1], A.L. Beaudet[1], C.A. Shaw[1], A. Patel[1], A. Gambin[2,10], S.W. Cheung[1], P. Stankiewicz[1]*. 1) Dept of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX;; 2) Institute of Informatics, University of Warsaw, Warsaw, Poland; 3) College of Inter-Faculty Individual Studies in Mathematics and Natural Sciences, University of Warsaw, Warsaw, Poland; 4) Institute of Computer Science, Warsaw University of Technology, Warsaw, Poland; 5) Department of Medical Genetics, Eastern Maine Medical Center, Bangor, Maine; 6) Lenox Hill Hospital, New York, NY; 7) Instituto de Referencia Andino, Bogota, Columbia; 8) Clinical Genetics, Specially for Children, Austin, TX; 9) Molecular and Medical Genetics, OHSU, Portland OR; 10) Mossakowski Medical Research Centre, Polish Academy of Sciences, Warsaw, Poland.

Using bioinformatic approaches, Sharp et al. (2005) predicted the genome-wide occurrence of 130 genomic rearrangements between directly oriented paralogous low-copy repeats (LCRs) > 10 kb in size, of > 95% DNA sequence identity, and located between 0.05–10 Mb apart from each other. To date, approximately 40 nonallelic homologous recombination (NAHR)-mediated rearrangements have been described as microdeletion and microduplication syndromes. Several other regions are thought to be pathogenic but have yet to be fully characterized. We queried our chromosomal microarray analysis database of 24,021 patients, in which all LCR-flanked regions are interrogated. The most common recurrent CNVs in our database are *NPHP1* duplications (214, benign variant), *CHRNA7* duplications (175, likely pathogenic with incomplete penetrance), and 22q11.2 deletions (161, pathogenic). After subtracting ~ 75% of cases, in which parental studies were not available, we found that out of 178 de novo recurrent CNVs, the most frequent are deletions in the DiGeorge syndrome (45), 16p11.2 autism (27), and Williams-Beuren syndrome (11) regions. Further, we analyzed various DNA features to correlate with recurrent CNV frequencies and to identify the NAHR recombination sites. We also bioinformatically analyzed the remaining unknown genomic regions flanked by directly oriented LCRs. Here, we report three novel recurrent and likely pathogenic genomic rearrangements. An ~ 1.7 Mb deletion in 2q12.2q12.3 was found in patients with developmental delay (DD) and dysmorphic features. It harbors *ST6GAL2*, previously associated with schizophrenia, and was mediated by an LCR 18.4 kb in size with 98.98% DNA sequence identity. An ~ 0.6 Mb deletion of 2q12.3 in patients with moderate DD and dysmorphic features, which encompasses the *SLC5A7* gene (choline transporter) was mediated by a 25.1 kb LCR with 97.62% sequence identity. An ~ 1.5 Mb deletion in 2q12.3q13 containing the *EDAR* and *RANBP2* genes was detected in a patient with skeletal anomalies. Heterozygous mutations in *EDAR* have been reported in patients with ectodermal dysplasia (OMIM 129490) and *RANBP2* mutations are seen in patients with necrotizing encephalopathy (OMIM 608033). This deletion was mediated by a 29.2 kb LCR with 97.53% sequence identity. These studies further elucidate the importance of NAHR-mediated genome instability in human disease and have resulted in the identification and characterization of three novel recurrent genomic disorders.

433F

DUF1220 Domain Copy Number Loss Implicated in 1q21-associated Microcephaly. *L. Dumas[1], M. O'Bleness[1], J. Davis[1,2], C.M. Dickens[1], N. Anderson[1], J.G. Keeney[1], J. Jackson[1], M. Sikela[1], A. Raznahan[4], J. Giedd[4], J. Rapoport[4], S.S.C. Nagamani[3], A. Erez[3], N. Brunetti-Pierri[5], R. Sugalski[6], J.R. Lupski[3], T. Fingerlin[2], S.W. Cheung[3], J.M. Sikela[1]*. 1) Department of Biochemistry and Molecular Genetics, Human Medical Genetics and Neuroscience Programs, University of Colorado School of Medicine, Aurora, CO; 2) Department of Epidemiology, University of Colorado School of Public Health, Aurora, CO; 3) Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX; 4) National Institute of Mental Health, National Institutes of Health, Bethesda, MD; 5) Telethon Institute of Genetics and Medicine, Naples, Italy and Department of Pediatrics, Federico II University of Naples, Italy; 6) Brooke Army Medical Center, San Antonio, TX.

DUF1220 protein domains show the largest human lineage-specific increase in copy number of any protein coding region in the human genome and map primarily to 1q21, where deletions and reciprocal duplications have been associated with microcephaly and macrocephaly, respectively. Given these findings and the high correlation between DUF1220 copy number and brain size across primate lineages (R2=0.98; p=1.8 × 10-6), DUF1220 sequences represent plausible candidates to underlie 1q21-associated brain size pathologies. To directly investigate this possibility, we used specialized bioinformatics tools developed for scoring highly duplicated DUF1220 sequences to implement targeted 1q21 array CGH on individuals (N=42) with 1q21-associated microcephaly and macrocephaly. We show that of all 1q21 genes examined (N=53), DUF1220 copy number shows the strongest association with brain size among individuals with 1q21-associated microcephaly, particularly with respect to the three evolutionarily conserved DUF1220 clades, CON1, CON2, and CON3 (CON1, p=0.0079; CON2, p= 0.0134;CON3, p=0.0116034). Notably, all DUF1220-encoding genes belonging to the NBPF family show significant correlations with FOC Z-score in the deletion group. In a similar survey of 1q21 copy number differences in a non-disease population we show that DUF1220 copy number exhibits the strongest correlation with gray matter volume (CON1, p=0.0246; CON2, p= 0.0334019). When comparing the arrayCGH results for disease and non-disease populations, only the DUF1220 sequences are consistently significant in both groups. Taken together these data strongly implicate DUF1220 copy number loss in the etiology of 1q21-associated microcephaly and are fully compatible with the view that DUF1220 domains function as key general effectors of evolutionary, pathological and normal variation in brain size.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

434W

Age-related somatic structural changes in the nuclear genome of human blood cells. *L.A. Forsberg[1], C. Rasi[1], H.R. Razzaghian[1], G. Pakalapati[1], L. Waite[2], K.S. Thilbeault[2], A. Ronowicz[3], N.E. Wineinger[4], H.K. Tiwari[4], D. Boomsma[5], M.P. Westerman[6], J.R. Harris[7], R. Lyle[8], M. Essand[1], F. Eriksson[1], T.L. Assimes[9], C. Iribarren[10], E. Strachan[11], T.P. O'Hanlon[12], L.G. Rider[12], F.W. Miller[12], V. Giedraitis[13], L. Lannfelt[13], M. Ingelsson[13], A. Piotrowski[3], N.L. Pedersen[14], D. Absher[2], J.P. Dumanski[1]*. 1) Dept. of Immunology, Genetics and Pathology, Rudbeck laboratory, Uppsala University, 75185 Uppsala, Sweden; 2) HudsonAlpha Institute for Biotechnology, 601 Genome Way, Huntsville, AL 35806, USA; 3) Dept. of Biology and Pharmaceutical Botany, Medical University of Gdansk, Hallera 107, 80–416 Gdansk, Poland; 4) Section on Statistical Genetics, Department of Biostatistics, University of Alabama at Birmingham, Ryals Public Health Building, Suite 327, Birmingham, Alabama 35294-0022, USA; 5) Department of Biological Psychology, VU University, Van der Boechorststraat 1, 1081 BT Amsterdam, The Netherlands; 6) Hematology Research, Mount Sinai Hospital Medical Center, 1500 S California Ave, Chicago, Illinois, 60608, USA; 7) Department of Genes and Environment, Division of Epidemiology, The Norwegian Institute of Public Health, Post Box 4404 Nydalen, N-0403 Oslo, Norway; 8) Department of Medical Genetics, Oslo University Hospital, Kirkeveien 166, 0407 Oslo, Norway; 9) Department of Medicine, Stanford University School of Medicine, Stanford, California 94305, USA; 10) Kaiser Foundation Research Institute, Oakland, California 94612, USA; 11) Dept. of Psychiatry and Behavioral Sciences and University of Washington Twin Registry, University of Washington, Box 359780 Seattle, WA 98104, USA; 12) Environmental Autoimmunity Group, National Institute of Environmental Health Sciences, National Institutes of Health Clinical Research Center, NIH 10, Room 4-2352, 10 Center Drive, MSC 1301, Bethesda, MD 20892-1301, USA; 13) Department of Public Health and Caring Sciences, Division of Molecular Geriatrics, Rudbeck laboratory, Uppsala University, 751 85 Uppsala, Sweden; 14) Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, SE-171 77 Stockholm, Sweden.

The adult human body is composed of cells that by mutations of different kind have become more or less genetically divergent. These during life acquired changes are referred to as somatic genetic variation and are to be distinguished from the well studied inherited genetic variation. Somatic variation can manifest, for instance, as various forms of cancer but is still often ignored in studies, despite it being predicted ubiquitous. We have recently shown that structural somatic genetic variation accumulates in blood with age and is common also in phenotypically normal cells and people (Forsberg et al. 2012 PMID: 22305530). In brief, we studied age-stratified cohorts of 318 monozygotic (MZ) twins and 296 single-born subjects using SNP-array, and described age-related accumulation of copy number variation (CNV) in the nuclear genomes in vivo. Large structural aberrations were found in 3.4% of subjects older than 60 years but not in younger controls. Recurrent region- and gene-specific mutations were observed. Longitudinal analyses suggested variation in the rate of accumulation of clones carrying structural changes and a natural self-removal of aberrant cell clones from blood. We also showed that different sub-compartments of blood could be affected differently by structural changes. Important disease-related aspects emerge from our results. For instance, we observed in phenotypically normal people, somatic aberrations typical for myelodysplastic syndrome, suggesting that these aberrations represent a sub-clinical form of this disease. Our results may also help explain the age-related reduction of cell clone complexity in blood: a well described hallmark of immuno-senescence and aging. The mutations we describe likely provide the affected cells with a mild proliferative advantage, without transforming them into immortalized cancer clones. However, the proliferative advantage of a few cell clones could affect the overall complexity of clones present in the blood. To further describe how somatic variation contributes to human aging and development of late-onset disease we are now analyzing a substantial number of elderly people using a higher-density SNP-array platform. These samples are collected longitudinally and preliminary results support our published data. Thus, we confirm that cell clones in blood of normally aging people carry large somatic aberrations in high frequency. The progress of our project and new data will be presented at the meeting.

435F

Genome-wide scan of Common Copy Number Variant Regions in Post-Traumatic Stress Disorder. *G. Guffanti[1], A. Aiello[2], M. Uddin[3], D. Wildman[4], S. Galea[5], K. Koenen[5]*. 1) Department of Psychiatry, Columbia University, New York, NY; 2) School of Public Health, University of Michigan, Ann Arbor, MI; 3) Department of Psychiatry & Behavioral Neurosciences, Wayne State University, Detroit, MI; 4) Center for Molecular Medicine and Genetics and Department of Obstetrics and Gynecology, Wayne State University, Detroit, MI; 5) Department of Epidemiology, Columbia University Mailman School of Public Health, New York, NY.

Post-Traumatic Stress Disorder (PTSD) is a common heritable neuropsychiatric disorder of unknown etiology. CNVs represent a specific form of genetic variation that is crucially relevant to integrate the information made available by SNP microarray technologies in studies of complex traits. Since CNVs can cover whole gene loci, CNV dose-related effects may have a strong impact on important biological functions, e.g. alterations of protein expression levels and disruption of coding regions. We investigated the genome-wide distribution of CNVs in subjects exposed to traumatic events who developed PTSD and subjects exposed to the same traumatic events who did not develop PTSD using the HumanOmniExpress BeadChip. Participants were recruited from the Detroit Neighborhood Health Study (DNHS), a longitudinal cohort of predominately African American adults (18+) living in Detroit, Michigan and screened for lifetime trauma exposure. PTSD was assessed using a modified version of the PTSD checklist, a 17-item self-report measure of DSM-IV symptoms. After data quality control, 748 subjects were analyzed: 136 cases of PTSD and 612 controls who were exposed to traumatic events who did not develop PTSD. We detected a total of 26,335 autosomal CNVs using PennCNV algorithm. After filtering for CNVs with segment size shorter than 20 Kb, we included in subsequent analysis 9,210 CNVs with an average of 10.5 segments per subjects in cases and 12.1 in controls (average size of 86 Kb). We found higher number of CNVs of relatively short length (< 20 Kb) in DNHS sample, which is consistent with findings that report smaller haplotype blocks in African Americans. The burden analysis revealed no significant differences in CNVs frequencies or size overall and by CNV type between cases and controls. Out of the 9,210 CNV, 7,757 clustered in 884 CNV-Regions across all individuals. Comparison with data downloaded from the Database of Genomic Variants (DGV) revealed substantial overlap between our CNVRs and entries of single CNVs in the DGV. Within the CNVRs, 2,336 CNVs identified in our sample are not overlapping with DGV-CNVs. Of the 884 CNVRs, 130 are regions of duplication events, 468 of deletions and 287 of both duplications and deletions. Approximately half (49%) of our CNVRs are found to affect genes. To our knowledge, this is the first study that provides a comprehensive map of common CNVRs in PTSD on genome-wide scale. Further analyses are warranted to validate current findings.

436W

Genetic variation among somatic cells revealed by single nucleus sequencing. *I.M. Hall[1,2], M.L. Lindberg[1], J. Piper[3,4], K.J. Brennand[4], F.H. Gage[4], M.J. McConnell[3,4]*. 1) Department of Biochemistry and Molecular Genetics, University of Virginia School of Medicine, Charlottesville, VA; 2) Center for Public Health Genomics, University of Virginia, Charlottesville, VA; 3) Crick-Jacobs Center for Theoretical and Computational Biology, Salk Institute for Biological Studies, La Jolla, CA; 4) Laboratory of Genetics, Salk Institute for Biological Studies, La Jolla, CA.

The extent to which somatic genome variation contributes to diseases other than cancer is an important and unresolved question in human genetics. This question has been extraordinarily difficult to study at the genome-wide level due to the relatively large input DNA requirements of second generation sequencing instruments and the difficulty of isolating clonally defined cellular lineages from human somatic tissues. Single nucleus sequencing (SNS) overcomes these limitations. Building upon protocols developed for tumor samples, we have implemented a high-throughput and cost-effective SNS method that employs flow sorting, cell barcoding and minimal DNA amplification, and that can identify DNA copy number variants (CNVs) at megabase resolution from individual nuclei. We have applied this method to various post-mortem and cultured somatic cells from healthy human individuals. We find a surprisingly high incidence of aneuploidy and large CNVs. But, we have also identified insidious artifacts related to DNA amplification. Finally, we present preliminary results from ongoing efforts to perform deep whole-genome sequencing and comprehensive variant detection from single somatic cells.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

437F

Structural haplotypes and recent evolution of the human 17q21.31 locus. *R.E. Handsaker[1,2], L.M. Boettger[1,2], M.C. Zody[1,3], S.A. McCarroll[1,2].* 1) Harvard Medical School, Department of Gentics, Boston, MA; 2) Program in Medical and Population Genetics, Broad Institute, Cambridge, MA; 3) Genome Sequencing and Analysis Platform, Broad Institute, Cambridge, MA.

Structurally complex regions of the human genome are not yet understood. Human 17q21.31 is a particularly interesting such locus: it contains a megabase-long inversion polymorphism and many complex, previously uncharacterized copy-number variations (CNVs); it associates with fertility, female meiotic recombination rate, and neurological disease; and it appears to be under selection in Europeans. To elucidate how complex genome structures vary in populations, we developed a population-genetic approach based on (i) accurate typing of structural features in large populations and (ii) phasing of these structural features using populations and families to determine structural haplotypes.

We identified nine distinct structural forms of 17q21.31 segregating in the human population, each distinguished by the gain, loss, or rearrangement of a large genomic segment (>100 kb) relative to each of the other structural forms. These nine structural forms suggest a phylogeny that offers a structural history of this rapidly evolving locus.

We find that five of these structural forms, including forms of both the H1 and H2 inversion types) involve partial duplications of the *KANSL1* gene (formerly *KIAA1267*). These duplications have arisen independently on the H1 and H2 forms of the 17q21.31 inversion polymorphism and both have subsequently risen to high frequency (26% and 19%) in West Eurasian populations, although nearly absent in other parts of the world. The haplotypes around these duplications coalesce to recent ancestors (< 40 kya), suggesting that the parallel expansions in frequency of these duplications were recent. Both duplications produce novel, truncated transcripts of the *KANSL1* gene, whose *Drosophila* ortholog regulates the timing of the differentiation of female germ cells.

Relating complex structural alleles to phenotype is important. We show that these complex, multi-allelic structural forms - often assumed to be poorly captured by SNPs - can in fact be imputed from SNP genotypes with high accuracy using either existing methods that support bi-allelic markers or emerging methods that can directly impute multi-allelic variants. These approaches allow complex genome structures to be efficiently related to human phenotypes using existing genome variation data.

438W

Population genetics of haptoglobin and haptoglobin-related protein: 50 years of human copy number variation. *E. Hollox[1], A. Menard[1], F. Yang[2], B. Fu[2], M. Sironi[3], R. Hardwick[1,2].* 1) Department of Genetics, University of Leicester, Leicester, United Kingdom; 2) Wellcome Trust Sanger Institute, Cambridge, United Kingdom; 3) Scientific Institute IRCCS E. Medea, Bosisio Parini, Italy.

Haptoglobin, coded by the HP gene, is a serum protein that acts as a scavenger for free heme, and haptoglobin-related protein (coded by the HPR gene) forms part of the trypanolytic factor targeting ApoL1 to the surface of trypanosome pathogens. Recent genome-wide association studies have shown association of a SNP in the intron of HPR with total cholesterol levels, supporting previous studies suggesting a link between haptoglobin and the risk of myocardial infarction. In 1962, Oliver Smithies and others showed that a common polymorphism of haptoglobin (with alleles Hp1 and Hp2) was the result of an intragenic duplication, later determined to be 1.7kb in size. Some studies have associated the Hp2 allele with susceptibility to malaria but this remains controversial. We reasoned that analysis of the global distribution of alleles at these two CNVs, and analysing the genomic region for signatures of natural selection, might shed light on any putative pathogen selection pressure. Here were determine global allele frequencies of this polymorphism using the Human Genome Diversity Project panel, and place it in the context of flanking SNP variation. We also use the paralog ratio test, fiber-FISH and sequence-read depth analysis to characterise the rare duplication of the HPR gene, and show that is at a frequency of 15% in certain West African populations. We also examine the case for natural selection in response to pathogen pressure at both copy number variable loci.

439F

Microduplication in 5p15.33 including *CLPMT1L* and *TERT* genes identified in a patient with cleft lip and palate, cardiopathy and urogenital abnormalities. *G. Izzo[1], E.L. Freitas[1], A.C.V. Krepischi[2], D.R. Bertola[1], M.R.S. Passos-Bueno[1], C. Rosenberg[1].* 1) Geneticts and Evolutionary Biology Department, Instituto de Biociencias USP, Sao Paulo, Brazil; 2) CIPE, Fundação Antonio Prudente, A.C. Camargo Hospital, Sao Paulo, Brazil.

Cleft of the lip with or without cleft palate (CL/P) are common birth defects of complex etiology and may lead to deleterious effects on language and communication skills, as well as on the appearance, leading to psychological issues for patients and their families. A number of genomic regions have been associated with CL/P, and the discovery of new candidate risk factors is essential for prognosis, management, and genetic counseling. In addition, several cases of CL/P often appear associated with genetic syndromes that involve additional birth defects, including developmental delay and/or mental retardation. Here we describe a patient with syndromic CLP. The major phenotypical abnormalities include dysmorphic facial features, neuro-psychomotor and developmental delay, hyperactivity and attention deficit disorder, atrio-ventricular septal defect, cryptorchid testes and micropenis. Array-CGH analysis revealed a rare interstitial microduplication on chromosome 5p15.33 (chr5:1,062,209-1,364,929; GRCh37/hg19) of approximately 300 kb encompassing five Refseq genes, including *TERT*, *CLPMT1L* and three *SLC* family members. This chromosomal abnormality was confirmed by real-time quantitative PCR. The CLPMT1L protein has a major domain homology with a protein previousy associated with CLP due to triggering an apoptotic pathway. Furthermore, *CLPMT1L* overexpression appears to lead to increased aptosis in vitro. In addition, *TERT* gene mutations have also been previoulsy associated with cardiopathies. Thus, we suggest that this microduplication has a causative relation with the patient phenotype. We are the first group to report an association of the *CLPMT1L* gene with the CL/P phenotype.

440W

Integration of microarray and NGS generated data for copy number analysis. *R. Keshavan, S. Verma, S. Shams.* BioDiscovery Inc, Hawthorne, CA.

High-throughput genomic data generation technologies produce massive amounts of data at ever increasing rates. One of the major scientific obstacles now is efficient analysis of this data to establish a better understanding of the underlying biology. Although the growing volume and higher density microarrays were a challenge to efficient analysis, Next-Generation Sequencing (NGS) technologies have complicated the problem by orders of magnitude. In this paper, we will explore a single dimensionality of the analysis issue that is related to efficient aggregation, integration and visualization of copy number data generated by microarrays as well as NGS instrumentation. We will discuss methods for estimating copy number data from exome and whole genome sequencing data and combining this data with previously generated array data. We will described the areas of commonality and differences of each platform and show how such integrated approaches might lead into better analysis results.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

441F

Comparative performance of three algorithms in detecting CNV associated with obesity and non-alcoholic fatty liver disease. *W. Li[1], J. Littrell[1], S. Gawrieh[2], M. Olivier[1]*. 1) Biotechnology and Bioengineering, Medical College of Wisconsin, Milwaukee, WI; 2) Division of Gastroenterology and Hepatology, Department of Medicine, Medical College of Wisconsin, Milwaukee, WI.

Our research explores the potential role of copy number variation (CNV) in obesity and non-alcoholic fatty liver disease (NAFLD). CNV, which may take the form of addition or removal of paralogous segments from a genome, has been shown previously to be associated with diverse disease phenotypes, but its role in obesity and its associated co-morbidities is unclear. To assess this effect of CNV on obesity, we screened a set of Affymetrix SNP6 chips generated from a sample of 753 individuals from 354 families who had BMIs ranging from 13.8 to 84.71. In addition, we included 147 samples from unrelated morbidly obese individuals (BMI>40) with or without NAFLD. We analyzed our Affymetrix SNP6 data with three algorithms to identify CNV: (1) Affymetrix Genotyping Console (AGC); (2) PennCNV; and (3) cn.FARMS. The output from each algorithm was subsequently compared to identify a set of loci corroborated by each analysis. Comparisons of the three genotyping-algorithms reveal potential biases in each method. Results to date indicate that AGC tends to report the greatest number of loci. However, approximately 70% of these were not corroborated by competing methods. AGC loci also have a greater average length (170 ± 130 Kb) than those detected by other methods, suggesting a performance bias with respect to especially lengthy paralogous segments. In contrast, PennCNV detected loci with a smaller average length (12 ± 31 Kb) and, overall, identified fewer loci that were not corroborated by other algorithms. Pathway analysis of the set of corroborated loci suggests that several detected loci are metabolically linked to obesity or potentially interact with such loci or their encoded products. Forthcoming analyses will characterize proteins encoded by previously unstudied loci to determine their metabolic roles and potential causal role in obesity and obesity associated phenotypes.

442W

Increased Rate of Rare Genic Copy Number Variants in Parkinson's Disease Among Ashkenazi Jews. *X. Liu[1], C. Cheng[2], S. Kisselev[3], H. Mejia-Santana[2], E. Louis[1,2,4,5], L. Cote[2], H. Andrews[6], C. Waters[4], B. Ford[4], S. Frucht[1], S. Fahn[1], K. Marder[1,2,4,7], J. Lee[1,2,5,8], L. Clark[1,3,8]*. 1) Taub Institute for Research on Alzheimer's Disease and the Aging Brain, College of Physicians and Surgeons, Columbia University, New York, NY, USA; 2) Gertrude H. Sergievsky Center, College of Physicians and Surgeons, Columbia University, New York, NY, USA; 3) Department of Pathology and Cell Biology, College of Physicians and Surgeons, Columbia University, New York, NY, USA; 4) Department of Neurology, College of Physicians and Surgeons, Columbia University, New York, NY, USA; 5) Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY, USA; 6) New York State Psychiatric Institute, Data Coordinating Center, New York, NY, USA; 7) Department of Psychiatry, Columbia University Medical Center, NYC, NY, USA; 8) Center for Human Genetics, College of Physicians and Surgeons, Columbia University, New York, NY, USA.

Background: Copy number variation (CNV) has been associated with neuropsychiatric and neurodegenerative disease. In Parkinson's disease (PD) patients, rare large genomic duplications or deletions have been observed in *SNCA* (*PARK1*), *Parkin* (*PARK2*) and DJ1 (*PARK7*) and Tyrosine Hydroxylase (*TH*). To date, only one genome-wide case-control study has assessed the contribution of CNVs to familial PD in a North American dataset. Method: We conducted a genome-wide scan for CNVs in a case-control dataset of Ashkenazi Jewish (AJ) origin (268 PD cases and 178 controls). Illumina 610/660 chip raw intensity array data was normalized using GenomeStudio. CNV calls and quality control were generated by PennCNV. Global genome wide burden analysis of rare CNVs (≤1% in the dataset,≥100 Kb) and rare CNV region were analyzed by PLINK. Ingenuity Pathway Analysis software (IPA) was used to search for biological relationships among the genic CNVs and known Mendelian PD genes. Result: A total of 986 rare CNVs were observed in our dataset of 432 subjects. CNVs in the PD genes like *SNCA*, *PARK2*, *DJ-1*, *PINK1*, *LRRK2* and *TH* were not identified. Overall global burden analyses did not reveal differences between cases and controls in CNV rate, distribution of CNV or number of genes affected by CNVs. Overall deletions (total CNV size) (P=0.019) and large CNVs (P=0.046) with length≥500kb were found 1.4 fold and 1.24 fold respectively more often in cases than in controls. For rare CNV regions, global burden was elevated and a significant association between PD and ovostatin 2 (*OVOS2*) was observed (Chr11p11.21)(P=0.028). 7 male sporadic PD cases with overlapping heterozygous duplications at *OVOS2* were found(CNV length range: 118Kb-148Kb), but not in 171 controls. A total of 81 PD cases carried a rare genic CNV (single occurrence or <5 PD cases as a cut-off) that was absent in controls. The IPA network identified *ATXN3*, *FBXW7* in the same disease pathway as *PARK2* (*Parkin*), *CHCHD3* and *SDF4* as *SNCA*, *HSF1* and *KLC1* as *MAPT* and *MBD3* as *PARK7* (*DJ-1*). Conclusions: Our study suggests that deletions (total CNV size) and large deletions (≥500Kb) are more frequent in cases than in controls and are associated with increased risk in PD. *OVOS2* was found as a significant risk factor for PD in the AJ population. We observed an increased rate of sporadic (de novo) and recurrent rare genic CNVs in PD cases and IPA analysis mapped those genic CNVs to known Mendelian PD networks.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

443F

A Map of Copy Number Variation in Admixed Populations in Xinjiang. *H. Lou[1], S. Li[2], R. Fu[1], X. Pan[2], L. Jin[1,2], S. Xu[1].* 1) CAS-MPG PICB, Shanghai, China; 2) MOE IBS, Fudan, Shanghai, China.

Xinjiang, the largest Chinese administrative division, is geographically located in Central Asia and played an important role historically in connecting Eastern Eurasian (EEA) and Western Eurasian (WEA). However, human population genomic studies in this region have been largely underrepresented in similar efforts both worldwide and within China. Especially, copy number variations (CNVs), which cover a large portion of human genome and are expected to harbour considerable genetic diversity that could contribute to neurological diseases and complex traits, have never been explored in any of the ethnic groups residing in Xinjiang. Here we investigated CNVs in the three major ethnic groups including the Kazakh (KZK), Kirgiz (KGZ) and Uyghur (UIG) with Affymetrix Genome-Wide Human SNP Array 6.0. A total of 1,461 CNV regions have been identified in 131 samples representing the three groups, and more than 70% of the CNVs could be also found in either of Eastern Eurasian or Western Eurasian popualtions. Interestingly, all the three groups have shown increased heterozygous deletions and duplications at both population level and individual level, as an expected consequence of population admixture. The estimated ancestral proportions inferred based on biallelic CNVs for KGZ, KZK and UIG were 63:37, 65:35 and 45:55 (EEA:WEA), respectively. These results were consistent with genetic relationship revealed by principle component analysis and phylogenetic analysis, KZK was closer to EEA, UIG was closer to WEA and KGZ was in the middle. The deletion and duplication allele frequency in the three groups also followed an admixture pattern, that is, a mixture of ancestral allele frequencies weighted by their contribution. Finally we identified ancestry biased and population specific CNVs which are likely to be involved in adaptation to local environment. Our study provided a first genome-wide perspective of CNVs three major ethnic groups in Xinjiang which will be valuable for both evolutionary and medical studies.

444W

Human CNVs and ultraconserved elements (UCEs): a rapid progression to mutual exclusivity. *R.B. McCole, C.Y. Fonseka, C.-T. Wu.* Department of Genetics, Harvard Medical School, Boston, MA.

Ultraconserved elements (UCEs) are genomic regions showing exceptionally strong and unexplained levels of DNA sequence conservation between related species. We have hypothesized that UCEs function as `copy counters,' helping genomes to maintain exactly the correct number of chromosomes and the right amount of genetic information. This model predicts that perturbations in the copy number of UCEs will be highly deleterious to the cell and, ultimately, to the individual. Early tests of the model examined the co-occurrence (overlap) of UCEs with copy number variants (CNVs) and segmental duplications (SDs) in the human genome. These studies demonstrated that the profiles of human CNVs and SDs are highly depleted for UCEs, thereby lending support to our model. We have now extended these studies to address the following questions: How quickly do UCEs become depleted from human CNV profiles? When does this depletion occur? Does it require one or more rounds of human reproduction? Using new sets of UCEs and the most recent datasets of CNVs, we find that depletion is rapid and does not necessarily involve passage through the germline. These findings show that UCE depletion from human CNV profiles is an ongoing phenomenon in current human populations, possibly occurring during development, thus propelling the study of ultraconservation into the realm of human disease, including cancer, where genome instability is common. Funding from R01 GM085169-01A1, to C.-T. Wu.

445F

Prediction of 3551 human haploinsufficient genes. *S. Meader, F. Honti, C. Webber.* MRC Functional Genomics Unit, University of Oxford, Oxford, United Kingdom.

Haploinsufficient (HIS) genes are those for which the absence of one fully functioning copy in a diploid organism causes an abnormality. Many genes with causative roles in human disease, such as neurodevelopmental disorders and cancer, are haploinsufficient. Thus, the identification of causative disease genes from among a patient's genetic variants may be greatly aided if we are able to reliably and sensitively identify first which genes in the human genome are haploinsufficient. We sought to score haploinsufficiency likelihoods for as many human genes as possible. As training data, we employed a previously published set of 301 known human haploinsufficient disease genes, and collated a second set of 1266 genes whose mouse orthologues, when knocked out in a hemizygous state, produce an abnormal phenotype. Haplosufficient (HS) genes were taken from common human copy number variable regions. We evaluated a range of functional and evolutionary properties for these genes. From an evolutionary perspective, we found that haploinsufficient genes evolve at a slow rate over short evolutionary distances ($p = 1.5 \times 10\text{-}4$), and are more likely to be retained as 1:1 orthologues over longer evolutionary distances ($p = 1.0 \times 10\text{-}10$). We observed that HIS genes are significantly more likely to be highly expressed in fetal tissues ($p = 0.017$), and tend to be expressed in a tissue specific manner. HIS genes were also significantly enriched for involvement in complex formation ($p = 1.0 \times 10\text{-}8$), and were enriched for a range of Gene Ontology categories. Using both publically available (STRING) and in-house functional linkage networks, we demonstrated that HIS genes cluster very strongly with other known HIS genes in these networks ($p < 1.0 \times 10\text{-}5$). We trained a Support Vector Machine on these characteristics to classify ~17,000 genes as either HIS or HS and to provide an associated likelihood. With our training set derived from orthologues of HIS mouse genes, we demonstrate a high degree of accuracy in our predictions (AUC = 0.948, previously published best AUC = 0.812), with the strongest single predictor being our in-house functional linkage network. Overall, employing our most conservative data set, we predict 3551 genes within the human genome to be HIS.

446W

Prioritization of copy number variation loci associated with autism from AutDB: An integrative multi-study genetic database. *I. Menashe, E. C. Larsen, S. Banerjee-Basu.* MindSpec, McLean, VA.

Copy number variants (CNVs) are thought to play an important role in the predisposition to autism spectrum disorder (ASD). However, their relatively low frequency and widespread genomic distribution complicates their accurate characterization and utilization for clinical genetics purposes. Here we present a comprehensive analysis of multi-study, genome-wide CNV data from AutDB (http://mindspec.org/autdb.html), a genetic database that accommodates detailed annotations of published scientific reports of CNVs identified in ASD individuals. Overall, this study included 4,926 CNVs in 2,373 ASD subjects from 48 scientific reports, encompassing ~2.12×10-9 bp of genomic data. Remarkable variation was seen in CNV size, with copy number gains (CNGs) being significantly larger than copy number losses (CNLs), ($P = 3\times10\text{-}105$; Wilcoxon rank sum test). Examination of the CNV burden across the genome revealed 15 loci with a significant excess of CNVs among ASD subjects ($P < 10\text{-}5$). Altogether, these loci covered 15,846 kb of the genome and contained 170 genes. Specifically, eight loci included ≥2 genes, six loci encompassed only one gene, and one locus (9p23) did not include any known genes. Interestingly, our analysis indicates that 15q11.2–13.3, a genomic region prone to chromosomal rearrangements of various sizes, contains three distinct ASD susceptibility CNV loci that vary in their genomic boundaries, CNV types, and inheritance patterns. We anticipate that the results of our analysis will promote the clinical utility of these 15 loci and facilitate future investigation into the underlying mechanism of CNVs in ASD susceptibility.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

447F

Copy Number Variant analysis in a deeply phenotyped cohort of individuals with Intellectual Disability. *E. Mercier[1,2], Y. Qiao[3], J. Gillis[2], S. Lewis[4], E. Separovic[3], P. Pavlidis[2]*. 1) Graduate Program in Bioinformatics, University of British Columbia, Vancouver, British Columbia, Canada; 2) Department of Psychiatry and Centre for High-Throughput Biology, University of British Columbia, Vancouver, British Columbia, Canada; 3) Department of Pathology, University of British Columbia, Vancouver, British Columbia, Canada; 4) Department of Medical Genetics, University of British Columbia, Vancouver, British Columbia, Canada.

The goal of our project is to gain insight into phenotypic variability as it relates to features of copy number variants (CNVs) such as their number, type (unique or common), size, gene content and position. We analyzed a cohort of 64 individuals with idiopathic intellectual disability (ID). Thirty three subjects had unique CNVs (de novo or rare familial) while 31 had only common CNVs. In agreement with previous work, de novo CNVs were larger ($p= 1.5\times10^{-8}$) and encompassed more genes ($p= 8\times10^{-9}$) than common CNVs. In our cohort, unique CNVs were preferentially found within 5Mb of chromosome ends ($p<0.001$). An analysis of Gene Ontology categories for the de novo or non-commons CNVs did not provide substantial insight into functional commonalities linking the CNVs. In our preliminary analysis we used 9 coarse phenotype categories: abnormalities of prenatal and early postnatal development (systemic congenital abnormalities, birth and neonatal), facial dysmorphism, abnormalities of brain, extremities, skin, trunk, reproductive system as well as neurological abnormalities. To take an unsupervised multivariate approach to discovering patterns in the phenotype data, we applied K-means clustering. The optimal number of clusters was two (Calinsky criterion), which were designated Group 1 (31 individuals) and Group 2 (33). Features which distinguished these groups ($p<0.01$) included prenatal and systemic congenital abnormalities, birth and neonatal abnormalities, all which were enriched in group 1. However, there was no significant association of CNV features (e.g., de novo vs. common) with either group. Further analysis using all 97 fine phenotype entries as well as gene content analysis of the CNVs and their association with specific patterns of phenotype and/or gene pathways is under way.

448W

A novel susceptibility gene, UGT2B17, is associated with familial Ankylosing Spondylitis. *D. O'Rielly[1], M. Uddin[1], R. Inman[2], D. Gladman[2], W. Maksymowych[3], R. Yazdani[1], P. Rahman[1]*. 1) Faculty of Medicine, Memorial University, St. John's, Newfoundland, Canada; 2) University of Toronto, Toronto, Ontario, Canada; 3) Department of Medicine, University of Alberta, Edmonton, Alberta, Canada.

Ankylosing spondylitis (AS) exhibits a strong genetic predisposition that is only partially accounted for with SNP-based associations. To date, no structural variations such as copy number variations (CNVs) have been reported to be associated with AS. Our primary objective was to identify highly penetrant, novel CNVs associated with familial AS by employing a custom designed genome-wide microarray. We applied our custom microarray comprised with 2 × 1 million probes targeting rearrangement hotspots with a mean spacing of 280 bp on a large multiplex three generational Caucasian AS family of Northern European Ancestry. This family consisted of five (5) individuals affected with AS, two members with systemic lupus erythematosus (SLE), and three unaffected family members. An ethnically-matched control was used for hybridization. At least five (5) probes were required to call an aberration with an average intensity >0.25 for a duplication and <-0.25 for a deletion. For CNV's identified using the microarray, independent validation was carried out using quantitative PCR (qPCR) following Taqman assay protocols. Microarray analysis revealed complex rearrangements within AS patients in multiple regions that were absent in unaffected family members. The most compelling evidence for structural variation was for the UGT2B17 gene. Duplication of the UGT2B17 gene was present in all five (5) patients with AS (two males and three females) and in the family member diagnosed with SLE. This genomic gain was absent in unaffected family members. Validation using qPCR demonstrated complete concordance with the microarray results and revealed that all five (5) AS individual contains ≥2 copies of UGT2B17, whereas unaffected family members had only one copy. The population distribution among Caucasians of the UGT2B17 CNV is 0.15 for a deletion, 0.45 for one copy and 0.25 for two or more copies. The UGT2B17 gene encodes a key enzyme responsible for glucuronidation of androgens which can stimulate bone formation. Interestingly, deletion of UGT2B17 has been associated with osteoporosis, but this has not been replicated in an independent study. Given the proposed function of this gene in stimulating bone formation, it represents a high priority susceptibility gene in familial AS. Replication studies and further functional studies are warranted.

449F

Evolution of copy number variation in the rhesus macaque β-defensin region. *B. Ottolini[1], T. Schwarzacher[2], C.L. Bevins[3], J.V. Solnick[3], G. Doxiadis[4], E.J. Hollox[1]*. 1) Department of Genetics, University of Leicester, Leicester, United Kingdom; 2) Department of Biology, University of Leicester, Leicester United Kingdom; 3) Department of Microbiology and Immunology, School of Medicine, University of California, Davis CA; 4) Department of Comparative Genomics & Refinement, Biomedical Primate Research Center, Rijswijk, Netherlands.

Beta defensins are multifunctional secreted short peptides: they present antibacterial and antiviral action in many species, possess immune cell signal activity in humans, recruiting immature dendritic cells to infection sites, and control coat color in dogs and presumably in cattle. In humans the β-defensin region is known to be copy number variable (CNV) and contains six genes repeated as a block, with a diploid copy number between 1 and 12 and an approximate repeat length of 2.5 Mb. Although genome-wide comparative genomic hybridisation arrays (aCGH) provided evidence of this region also being CNV in chimpanzee and macaque, the extent and nature of CNV in different mammals remains unknown. The rhesus macaque (*Macaca mulatta*) is the most widespread non-human primate and represents a good model for immunity studies. Its genome has been sequenced, although there is poor assembly quality in repeated segments such as the β-defensin region. For all these reasons, we studied the genomic architecture of the rhesus macaque -defensin region using a variety of methods. We performed high-resolution aCGH on 16 non-related macaque individuals, PCR-based methods (Paralogue Ratio Test and microsatellite assay) on a cohort of 70 samples, and metaphase spread FISH and fibre-FISH on lymphoblastoid cell lines from 5 related individuals. Combining these approaches, we aimed to overcome the limitations of the assembly and to define an absolute copy number for this region in rhesus macaque. We present here preliminary results showing that only the region containing the human DEFB4 (DEFB2L) orthologue is CNV, while the rest of the genes show no variation. The repeated block encompasses only 8.5kb, considerably smaller than that found in humans, and this will be discussed in an evolutionary context. Also, we will present how the combined use of PCR-based and cytogenetic approaches allows absolute copy number calling and may be applied to other mammalian species to give insight into the genomic evolution of the β-defensin region.

450W

Characterization of copy number variation involving the salivary agglutinin gene *DMBT1*. *S. Polley[1], D. Hains[2], E. Hollox[1]*. 1) Department of Genetics, University of Leicester, Leicester, Leicestershire, United Kingdom; 2) Nationwide Children's Hospital, Columbus, Ohio, USA.

Salivary agglutinin encoded by the gene *DMBT1* is a multifunctional protein with roles in innate immunity, regulation of inflammation and cell differentiation. The central region of *DMBT1* contains locus specific tandem repeated scavenger receptor cysteine-rich (SRCR) domains which bind to a broad spectrum of Gram-positive and Gram-negative bacteria as well as viruses. The main objective of the study was to characterize the copy number variation involving the DMBT1 gene. We have developed a cost effective, high throughput and easily reproducible PCR-based copy number typing system using the paralog ratio test (PRT) for estimating the exact copy number of SRCR domains in the *DMBT1* gene of different people. We validate that this method is accurate and precise when compared to array CGH and long-range PCR. The SRCR copy number estimation in different populations provides the pattern of genetic variation in SRCR domain of *DMBT1*. A worldwide analysis of SRCR copy number shows a high frequency of high SRCR copies in Africa and America and high frequency of low SRCR copies in East Asia, South Asia and Oceania. We are currently comparing copy number between healthy individuals and patients will help us to explain the genetic contribution of *DMBT1* (mainly SRCR copy number) to common inflammatory and infectious diseases.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

451F

Single exon deletion in *PCCA* gene in patient with Propionic Acidemia highlights challenges to clinical whole exome CNV analysis. *G. Scharer[1,2], G. Creadon-Swindell[2], E. Spector[2], C. Coughlin[1], T. Shaikh[1,3]*. 1) Dept. of Pediatrics-Division of Clinical Genetics & Metabolism, University of Colorado Denver SOM, Aurora, CO; 2) DNA Diagnostic Laboratory, Dept. of Pediatrics- Division of Clinical Genetics & Metabolism, University of Colorado SOM, Aurora, CO; 3) Colorado Intellectual and Developmental Disabilities Research Center, University of Colorado SOM, Aurora, CO.

The application of oligonucleotide arrays (aCGH) for the detection of genomic copy number variation (CNV) has become accepted clinical practice. This can include comprehensive whole-genome arrays that also cover exons of known disease genes, or those with dense, targeted coverage of selected genes. The latter type of arrays is especially useful to identify exonic deletions in autosomal recessive disorders, where standard sequencing has found only 1 disease causing mutation. Identification of a small CNV is typically based on a defined gain or loss in signal (log2) ratios of several consecutive probes by software analysis. Reporting threshold is generally set arbitrarily at 5 probes to decrease number of false positives, while accepting a decrease in sensitivity. Here we report the case of a patient identified by biochemical testing with propionic academia (PA), who was found with one point mutation only by Sanger sequencing in the *PCCA* gene. Subsequent targeted exonic deletion testing by aCGH utilizing a custom 135k oligonucleotide array (a minimum of 3 probes are placed in each exon) did not identify an exonic deletion after data analysis utilizing two commercial software packages at standard settings. Only after manual review a possible heterozygous 3 probe deletion was noted in exon-4 of the *PCCA* gene. Due to the small size of exon-4 (69 bp) and a relatively large intron-4 (43 kb) only 3 probes covered this coding region. Of note, the 3 probes in the adjacent exon-3 (48 bp) did not show loss of signal. Follow-up testing utilizing qPCR (Dual Realtime-PCR) did verify the heterozygous deletion of exon-4 in the patient, and hence confirmed the diagnosis of PA. This case highlights common difficulties in array design, but also the analytical challenges in identifying very small copy number variants in coding sequences. Simply lowering the reporting threshold to 3 or less probes would create too many false positives hits and make verification impractical. While manual review is possible in targeted CNV analysis (gene panels or single genes with a single point mutation), it is virtually impossible for arrays designed for whole exome/genome analysis. In addition to considering denser probe coverage in coding regions, coverage needs to be extended into intronic sequences, thus requiring high-density arrays carrying 1–2 million probes, i.e. for whole exome CNV analysis.

452W

Identifying Autism Loci using Homozygosity Analysis and Copy Number Variation Analysis. *K.E. Schmitz-Abe[1,2,5], T.W. Yu[1–6], T.K. Kim[9], R.S. Hill[1–3,5], E.M. Morrow[4,7,8], M.E. Greenberg[5], K. Markianos[1,2,5], C.A. Walsh[1–6]*. 1) Division of Genetics, Department of Medicine, Children's Hospital Boston, Boston, Massachusetts, USA, 02115; 2) Manton Center for Orphan Disease Research, Children's Hospital Boston, Boston, Massachusetts, USA, 02115; 3) Howard Hughes Medical Institute, Children's Hospital Boston, Boston, Massachusetts, USA, 02115; 4) The Autism Consortium, Boston, Massachusetts, USA, 02115; 5) Harvard Medical School, Boston, Massachusetts, USA, 02115; 6) Department of Neurology, Massachusetts General Hospital, Boston, Massachusetts, USA, 02114; 7) Department of Molecular Biology, Cell Biology and Biochemistry, Brown University, Providence, Rhode Island, 02912; 8) Department of Psychiatry and Human Behavior, Brown University, Providence, Rhode Island, 02912; 9) UT Southwestern Medical Center 5323 Harry Hines Boulevard Dallas, Texas 75390.

Analysis of large Autism data sets has provided statistical and functional evidence for the role of rare deleterious variants, point mutations as well as copy number variants (CNVs). Although such lesions explain only a few percent of Autism Spectrum Disorders (ASD) they provide valuable insights in the diverse developmental errors that lead to ASD. Here we present strong statistical evidence for a mechanism rarely implicated before in ASD: copy number 0 CNVs (homozygous deletions). We performed CNV analyses in 213 consanguineous families from the Middle East (258 affected individuals, 552 unaffected individuals) using Affymetrix 6.0 and 500K SNP arrays, using five algorithms (Birdsuite, PennCNV, Affymetrix Genotyping Console, Nexus and Golden Helix) and requiring concordance of at least two calls. To distinguish between common and rare copy number events, we compiled a CNV catalog of 1,258 samples from the International HapMap project using the same analysis pipeline. Applying this pipeline to our dataset, we observe an excess of rare, homozygous deletions (p=0.008) in cases versus controls. This enrichment remains significant after controlling for overall homozygosity in cases vs. controls (p=0.02). The 27 homozygous deletions in 24 cases delete exons (13/27), just introns (3/27) or are apparently non-coding (11/27). Currently we are testing whether CNV0 events are present and play a similar role in the Autism Genetic Research Exchange (AGRE) Collection (1,258 affected, 1,746 unaffected) where homozygosity is much reduced but the sample size is much higher. Functional analysis of a large, previously reported CNV0 deletion near NHE9, (a gene that shows excess rare mutations in epilepsy and autism cases compared to controls) suggests that the deletion removes key regulatory sequences for NHE9 (Morrow et al, 2008). ChIPseq analyses reveal that mouse NPAS4 and CBP, transcriptional factors implicated in neuronal activity and contextual learning, are recruited by neuronal activity to bind sequences upstream of NHE9; human sequences orthologous to these binding sites are deleted by the previously reported CNV0 in this region, providing a biological mechanism by which this deletion likely disrupts the regulation of NHE9 expression by neuronal depolarization. Supported by the NIMH, NINDS, the Simons Foundation, and the HHMI.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

453F

Rare Exonic Deletions Contribute to Autoimmune Diabetes Susceptibility. *C. Shtir[1], D. Smith[1], H. Guo[1], J. Barrett[2], M. Hurles[2], N. Walker[1], V. Plagnol[3], J. Cooper[1], J. Howson[1], O. Burren[1], S. Rich[4], J. Todd[1]*. 1) JDRF/WT Diabetes & Inflammation Lab, Cambridge Inst Medical Research, Cambridge, United Kingdom; 2) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, UK; 3) University College London, London, UK; 4) University of Virginia, Charlottesville, VA, USA.

In this study we investigate rare deletions (DELs) in individuals diagnosed with type 1 diabetes (T1D) by analyzing 6,808 T1D cases, 9,954 controls, and 2,206 T1DGC families, genotyped on the high density ImmunoChip (iChip) array. Investigation of LRR distributions allowed correction of batch effects through principal components analysis (PCA). CNV calls were estimated based on PCA-corrected LRRs with PennCNV. We examined the burden of rare DELs at the iChip level, for all rare, genic, and exonic DELs, respectively. Based on candidate gene and GWAS studies we also examined 53 genomic regions that most convincingly implicate loci that alter T1D risk. Case-control Study: The median size of rare DELs across the iChip was 11kb for either cases or controls, and there was no correlation between gene-SNP density and number of rare DELs per gene (R2 = 0.05). We detected an excess of 4.2% case DELs at the 200kb threshold, with 12.5% case DELs vs. 8.3% control DELs [OR(CI)=1.77 (1.32,2.39);P=9.84E-05]. Rare DELs produced case-control ORs ranging from 1.24 to 1.35 (P=1.11E-04 to 2.65E-04) across the iChip, and from 1.51 to 2.33 in T1D regions (4.47E-07 to P=8.29-05), depending on whether all, genic, or exonic, rare DELs were studied. Non-T1D regions gave on average an OR of 1.1 (P>0.1). T1D regions remained significantly enriched, after 100K permutations from regression analysis with adjustment for gene size, DEL frequency and length. Family Cohort: T1D families reflect a different ascertainment from the case: control study, as affected offspring innately bear a stronger HLA heritability component. Affected offspring revealed a similar rare DEL load as the unaffected parents, at the iChip level and within T1D regions. However, 7.48% vs. 4.86% affected vs. unaffected offspring carried rare de novos [OR(CI)=1.58(1.07,2.41);P=1.91E-02], while transmitted DELs were detected in 19.37% vs. 17.29% affected vs. unaffected offspring [OR=1.14(0.91,1.45);P=2.4E-01]. These findings suggest that when de novo DELs arise in T1D offspring, they occur within an already compromised genetic background, surpassing a liability threshold beyond which an additional deleterious de novo mutation may predispose to T1D. Our results indicate a major new source of variation: rare, short, gene-disrupting CNVs for T1D and probably other common diseases, and highlight the necessity to analyse other diseases for which large numbers of samples have been typed using dense marker or probe arrays.

454W

An atypical microduplication overlapping the C2orf27A gene at 2q21.1 co-segregates with Tourette Syndrome in a three generation family. *M. Uddin[1], D.D. O'Reilly[1], P. Rahman[1], T.R. Purchase[2], H. White[2], S. Luscombe[3], S.J. Moore[3], T.L. Young[1], K.A. Hodgkinson[1]*. 1) Faculty of Medicine, Memorial University, St. John's, Canada; 2) Pediatric Psychiatry, Health Sciences Centre, St. John's, NL, Canada; 3) Pediatric Development and Rehabilitation, Health Sciences, St. John's, NL, Canada.

Tourette syndrome (TS) is a developmental neuropsychiatric disorder characterized by the presence of both motor and verbal tics. TS may be accompanied by features associated with obsessive compulsive disorder (OCD), attention deficit hyperactivity disorder (ADHD) and poor impulse control. The general population prevalence of TS is between 0.3–1% and affect males more than females. It has a major genetic component with concordance rates for monozygotic and dizygotic twins estimated at 50–70% and 10–23% respectively. Despite numerous linkage and association studies no strong common candidate genes have been identified. Recent analysis of copy number variations (CNV) in TS show an association with genes previously implicated in autism although none have been reported unique to TS. We have begun an epidemiological study of families with TS in Newfoundland, and have ascertained one multiplex family. Five family members have been clinically diagnosed and the 6th is potentially affected (by collateral family history). The family includes four young siblings, their mother and maternal grandfather. Three of the siblings are significantly affected with concomitant OCD, ADHD and anxiety. The fourth has TS with fewer co morbidities. We designed a custom array CGH targeting genomic hotspots prone to rearrangements. The high density array included 2 × 1 million probes targeting regions with a mean spacing of 270bp. We found a common ~30kb microduplication overlapping the C2orf27A gene at 2q21.1 that co-segregates with the six affected family members. The atypical duplication breakpoint is 80kb larger in the three siblings with significant comorbidities. The fourth sibling, mother and grandfather have the smaller microduplication. The function of the C2orf27A gene is currently unknown. In summary, we have identified a CNV that co-segregates with Tourette syndrome in a multiplex family. The importance of this and other disease-associated CNVs to Tourette (and associated syndromes) pathogenesis is being explored by ascertaining other multiplex families from the Newfoundland population.

455F

Detection of Copy Number Variation from Exomes in the DDD and UK10K Projects. *P. Vijayarangakannan, T. Fitzgerald, C. Joyce, S. McCarthy, M. Hurles, The DDD and UK10K projects.* Wellcome Trust Sanger Institute, Cambridge, United Kingdom.

Large-scale targeted-resequencing projects have become routine in studies that involve large sample sizes not only due to their cost and technological efficiency, but also to extend the analysis easily to a wider range of genetic abnormalities. The DDD (Deciphering Developmental Disorders) project and the UK10K project are two studies in the UK that aim to sequence thousands of individuals using whole-exome sequencing approaches. The DDD project aims to advance clinical genetic practice for children with developmental disorders through systematic development and application of latest microarray and next-gen sequencing methods. Starting from April 2012, the project aims to recruit 12,000 families in three years through the National Health Service Regional Genetics Services in the UK. The UK10K project aims to study diseases in 6000 patients caused by rare genetic changes in the human genome. Copy Number Variation (CNV) in the human genome has been implicated in a range of developmental disorders. Here, we report CoNVex, a novel algorithm for the detection of CNV from targeted-resequencing data. CoNVex utilises the read depth information in probe regions, compares it to a reference (median depth across samples) and detects copy number variable segments within the log2 ratio (of depth over median depth) using an error-weighted score and the Smith-Waterman algorithm. We have applied this to exome data from 230 DDD patient-parent trios. The DDD exome plus design includes probes that are used in the Agilent SureSelect 50Mb library in addition to custom probes. We evaluate the algorithm using the same samples for which exon-resolution aCGH and SNP array data are available for validation. We also compare our method to other exome-CNV calling algorithms. Our results show that CNV calling from exome data can have comparable resolution to single-chip aCGH for calling genic CNVs with high sensitivity and specificity.

456W

Alterations in *RBFOX1*: Real finding or red herring? *L. Walters[1], A. Bailes[1], L. Erdman[1], A. McKinney[1], S. Ramsey[1], C. Weber[1], S. Hashimoto[1], L. Lamb Thrush[1,2], J. Atkin[2,3], G. Herman[2,3], S. Hickey[2,3], K. Manickam[2,3], C. Crowe[4], S. Thomas[5], C. Astbury[1,6], J. Gastier-Foster[1,2,6], S. Reshmi[1,6], R.E. Pyatt[1,6]*. 1) Department of Pathology and Laboratory Medicine, Nationwide Children's Hospital, Columbus, OH; 2) Department of Pediatrics, The Ohio State University, Columbus, OH; 3) Division of Molecular and Human Genetics, Nationwide Children's Hospital, Columbus, OH; 4) Division of Clinical Genetics, MetroHealth Medical Center, Cleveland, OH; 5) Center for Child Development, Wood County Hospital, Bowling Green, OH; 6) Department of Pathology, The Ohio State University, Columbus, OH.

RBFOX1/A2BP1 (OMIM 605104) encodes an RNA-binding protein with six isoforms that regulate splicing in a tissue-dependent fashion. It was first identified as a binding protein of ATXN2, suggesting a role in neurologic function. RBFOX1 is highly expressed in the brain, skeletal muscle, and the heart. Studies in mice have demonstrated a dynamic role for RBFOX1 in regulating neuronal excitation and calcium homeostasis, and knockouts are prone to spontaneous seizures. In humans, translocation-based alterations, which disrupt regulatory regions of the locus, produce a variable phenotype including developmental delay (DD), seizures, and mild dysmorphisms [Bhalla et al. 2004; Martin et al. 2007]. Smaller alterations, including both gains and losses, have been found in association with a variety of neuropsychiatric disorders, including ADHD, schizophrenia, autism, and DD [Elia et al. 2010; Xu et al. 2008; Sebat et al. 2007; Mikhail et al. 2011]. We present a series of 11 cases ascertained between 2009–2012 that showed copy number changes in the *RBFOX1* locus by array CGH. The alterations, both gains and losses, ranged in size from 50 kb to 970 kb and clustered around the third intron of the largest isoform. Most patients displayed neurologic dysfunction, including DD, behavioral problems, and ADHD, along with a highly variable degree of physical dysmorphism. In addition, several cases presented with congenital heart defects. The most severe of these had dextrocardia, single ventricle with double inlet left ventricle, and situs invertus abdominalis. This case is similar to a published report of an infant with dextrocardia, single ventricle, and situs solitus, along with other malformations, with an intragenic deletion of *RBFOX1* [Lale et al. 2011]. Previous studies have demonstrated that disrupting the first exon and intron of the long isoform causes decreased expression of all isoforms [Martin et al. 2007]. However, the impact of disrupting more distal introns has not yet been established. Unfortunately, evidence from the International Standards for Cytogenomic Arrays supporting dosage sensitivity for *RBFOX1* is lacking. Given the clustering of our array findings and the persistent, albeit variable, neurological phenotype, which is consistent with previously published reports, we conclude that alterations in *RBFOX1* may be a susceptibility factor for neuropsychiatric disease, although its role in congenital heart defects has yet to be established.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

457F

Analyses of copy number variation reveal putative susceptibility loci in MTX-induced mouse neural tube defects. *JH. Wang[1], XW. Wang[1], T. Guan[2], Q. Xiang[3], MS. Wang[3], Z. Zhang[3], Z. Guan[1], GL. Wang[1], ZQ. Zhu[1], Q. xie[1], GN. Li[1], J. Guo[1], F. Wang[1], ZG. Wang[3], B. Niu[1,2], T. Zhang[1].* 1) Department of biotechnology, Capital Institute of Pediatrics, Beijing, China; 2) Department of Biochemistry and Molecular Biology, Shanxi Medical University, Taiyuan, China; 3) Chinese Academy of Medical Sciences, Peking Union Medical College, Beijing, China.

Copy number variations (CNVs) are thought to act as an important genetic mechanism underlying phenotypic heterogeneity. Impaired folate metabolism is considered to be a risk factor for neural tube defects (NTDs). However, the precise nature of this link between low folate status and NTDs remains unclear. Using an array-comparative genomic hybridization (aCGH) assay, we investigated whether CNVs can be detected in the NTD embryonic neural tube tissues of of methotrexate (MTX)-induced folate dysmetabolism pregnant C57BL/6 mice, and validated the aCGH findings by RT-PCR. Prior to performing the experiments, this study was reviewed and approved by the ethic evaluation committee of Capital Institute of Pediatrics, Beijing, China. The CNVs were then comprehensively investigated by informatics methods for prioritize candidate genes. Dihydrofolate reductase (DHFR) activity of maternal serum was detected, and the concentrations of folate and relevant metabolites were analyzed using LC/MS/MS. Three high confidence CNVs on XqA1.1, XqA1.1-qA2 and XqE3 were found In the NTD embryonic neural tissues. DHFR activities, 5-methyltetrahydrofolate (5-MeTHF), 5-formyltetrahydrofolate (5-FoTHF) and S-adenosylmethionine (SAM) concentrations of maternal serum decreased significantly after MTX injection. Twelve putative genes and three microRNAs were selected as the candidates for susceptibility to MTX-induced NTDs and might play roles in the NTD pathogenesis. These findings indicated that CNVs were involved in MTX-induced NTDs by causing the disorder of folate metabolism and would shed new light on the pathogenesis of NTDs. Jian-Hua Wang, Xiu-Wei Wang, contributed equally to this manuscript. Corresponding author: Bo Niu, Ting Zhang.

458W

Large range of copy number variation (CNV) for human alpha-defensin genes *DEFA1* and *DEFA3*: lower copy number is a risk factor for autoimmune diseases. *H. Wang[1], B. Zhou[1], YL. Wu[1], S. Bowden[1], R. Hoffman[1], J. Ahearn[3], D. Birmingham[2], B. Rovin[2], L. Hebert[2], A. Schwaderer[1], D. Hains[1], CY. Yu[1].* 1) The Research Institute at Nationwide Children's Hospital, Columbus, OH; 2) Wexner college of Medicine, The Ohio State University, Columbus OH; 3) Allegheny Singer Research Institute, Temple University School of Medicine, Pittsburgh, PA.

Human α-defensins (DEFAs) are antimicrobial peptides important in innate immunity. The current version (v37.3) of Human Reference Genome sequence reveals a total of six functional genes and seven pseudogenes in the *DEFA* locus spanning 150 kb in size at chromosome 8p23.1. The functional genes include *DEFA3*, *DEFA1* and *DEFA1B*. Because these genes are highly homologous in sequence identities and have extensive inter-individual gene CNV, they are collectively referred to as *DEFA1A3*. The objective of this study is to copy-type the *DEFA1A3* in different racial groups and autoimmune diseases. Dot-plot analyses of genomic DNA sequences and long-range genomic mapping by PacI and PmeI digested DNA resolved by pulsed field gel electrophoresis revealed segmental duplications of 19.1 kb for the *DEFA1A3* genes. Using a 1.2 kb genomic DNA probe specific for 5' region of human *DEFA1A3* for Southern blot analysis, we detected seven polymorphic variants characterized by 8.7, 8.2, 7.0, 6.0, 4.5, 3.7 and 3.2 kb TaqI restriction fragments in human populations. The CNVs and associated polymorphisms of *DEFA1A3* have been characterized by TaqI genomic RFLPs in 657 human subjects, among them are 192 healthy Caucasians, 164 healthy African Americans and 302 autoimmune patients (183 type 1 diabetes (T1D) and 119 systemic lupus erythematosus (SLE)). The diploid gene copy number (GCN) for *DEFA1A3* in our study population varies from 4 to 17. While >85% of human subjects have a GCN between 5 and 9 copies, the distribution patterns of GCN are significantly different among groups ($p<0.0001$). Healthy African Americans have highest GCN with a mean of 7.41 (95% CI=7.15–7.67), compared to healthy Caucasians with a mean GCN of 7.03 (6.79–7.27). Both Caucasian SLE and T1D patients have a mean GCN of 6.7 and these autoimmune patients have significantly lower copy number of *DEFA1A3* than their race-matched controls ($p=0.03$). Further analyses revealed that the TaqI-8.7 kb *DEFA1A3* variant is specific for Caucasians, while the TaqI-3.2 kb variant is specific for African Americans. Moreover, the frequency of the TaqI-6.0 fragment is significantly lower in Caucasian SLE than in race-matched controls ($p=0.03$). In summary, the GCN for *DEFA1A3* is highly variable among human subjects and racial groups. Low copy number of *DEFA1A3* appears to be a common risk factor for autoimmune diseases.

459F

Analysis of structural variation in the Genome of the Netherlands (GoNL) project. *K. Ye[1], V. Guryev[2], W. Kloosterman[3], L. Francioli[3], J.Y. Hehir-Kwa[4], E. Lameijer[1], A. Abdellaoui[5], J. de Ligt[4], V. Koval[6], N. Amin[6], F. van Dijk[7], L.C. Karssen[6], H. Mei[8], Genome of the Netherlands consortium.* 1) Leiden University Medical Center, Leiden, Netherlands; 2) Hubrecht Institute and UMC, Utrecht, Netherlands; 3) University Medical Center, Utrecht, Netherlands; 4) Radboud University Medical Centre, Nijmegen, Netherlands; 5) VU University Medical Center, Amsterdam, Netherlands; 6) Erasmus Medical Center, Rotterdam, Netherlands; 7) University Medical Center, Groningen, Netherlands; 8) Netherlands Bioinformatic Centre, Nijmegen, Netherlands.

The Genome of the Netherlands (GoNL) is a national collaboration that aims at characterizing genetic variations in the Dutch population of 250 families. Here we report on the pilot results of the structural variation analysis for 18 families. Our analysis employs several methodologies for detection of different types and size ranges of variants. Using GATK Unified Genotyper, we identified 1,459,968 small indels of which 23% are novel compared to 1000 Genomes phase 1 data, and 86% overlap with Pindel's short indel calls. In silico functional analysis indicates that 819 are causing premature stop codons and frameshifts in 749 genes. A combination of 4 approaches, read depth (CNVnator, DWAC-Seq), read pair (123SV, BreakDancer and GenomeSTRiP), split-read (Pindel), and de novo assembly (SOAPdenovo, CLC) ensures detection of structural variants of different types and size ranges. For example, we identified a 1.8 kb insertion absent in genome reference, but common in Dutch population (allele frequency=42%), while rare (5%) in 1000 Genomes project. Homozygous DNA segments were identified using PLINK and VCFtools. We performed hundreds of PCR/Sequencing assays to determine false-positive rate for each tool and to establish de novo mutation rate of indels and structural variants. The set of wide size range, multi-type, high-quality SVs calls, together with GoNL SNP set (reported separately) describes common genomic variants in the Dutch genomes. This variation catalogue is essential for understanding population history, interpretation of GWAS and analysis of other studies involving West-European samples.

460W

Copy number variation of the Y chromosome in autism spectrum disorder. *R.K. Yuen[1,2], E. Kolomietz[3], M. Vlasschaert[3], A. Merkoulovitch[1,2], J.R. MacDonald[1,2], C.R. Marshall[1,2], S.W. Scherer[1,2].* 1) McLaughlin Centre, The Hospital for Sick Children, Toronto, ON, Canada; 2) The Centre for Applied Genomics, The Hospital for Sick Children, Toronto, ON, Canada; 3) Division of Molecular Diagnostic Genetics, Department of Laboratory Medicine and Pathology, Mount Sinai Hospital, Toronto, ON, Canada.

Autism spectrum disorder (ASD) presents a complex and heterogeneous etiology with a strong genetic basis. While it affects 1 in 88 people, ASD has a strong gender bias and males are four times more likely to be on the spectrum than girls. Genetic mutations and copy number variations (CNVs) on the X chromosome have been associated with the ASD phenotype. However, the Y chromosome has not been well studied with respect to autism due to the fact that its complex and repetitive DNA sequence structure complicates typical analyses. Moreover, the existing microarrays do not provide a good coverage of probes along the Y chromosome. To investigate the CNV architecture of the Y chromosome in ASD, 32 male probands and their fathers along with 46 male controls were run on a custom Agilent 8×15K CGH array that we have developed. This array contains over 10K commercial and custom probes (size 44~60nt) with an average probe spacing of ~1.5kb across the euchromatic region of the Y chromosome. In our preliminary data, we found one *de novo* XYY male in ASD cases and a potential mosaic loss of chromosome Y in one male control. In addition to known CNVs in loci such as TSPY genes and partial AZFc region, we identified a novel CNV of ~10kb in the exonic regions of DAZ repeats in almost half of the male population (ASDs and controls). Although most of the CNVs in the probands are inherited and there is no significant difference between ASDs and controls for any CNVs detected, we found a de novo gain of a ~200kb region encompassing *TTTY20*, *FAM197Y2*, *TSPY1*, *TSPY3* and *TSPY4* in one proband from a simplex family. The copy number gain of this region was not observed in any of the male controls examined. The current findings suggest that large common CNVs on the Y chromosome are unlikely to be a major contributory cause of ASD, but some of the *de novo* CNVs detected may warrant further investigation. A larger sample cohort will be used for future studies of potential involvement of rare CNVs on the Y chromosome in ASD.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

461F

Novel method of CNV analysis in FcγR locus and its application to immune-related diseases. *A. Zhernakova[1, 2, 3], L. Franke[1], A. Ioan[2], G. Trynka[1], K. Fransen[4], R. Weersma[4], S. Rantapää-Dahlqvist[5], P. Gregersen[6], C. Wijmenga[1], T. Huizinga[2], R. Toes[2].* 1) department of department of Genetics, UMCG, Groningen, The NetherlandsGenetics, UMCG, Groningen, Netherlands; 2) department of Rheumatology, LUMC, Leiden, The Netherlands; 3) Complex Genetics Group, UMCU, Utrecht, The Netherlands; 4) department of Gastroenterology, UMCG, Groningen, The Netherlands; 5) Department of Public Health and Clinical Medicine/Rheumatology, Umea University, Sweeden; 6) Feinstein Institute for Medical Research, Manhasset, New York, USA.

Genetic variants near the FC-gamma receptor (FcγR) locus are associated with several immune-related diseases. However, most FcγR genes are located in complex regions of segmental duplications (SD) and they are therefore not well covered by the genotyping platforms. To be able to identify copy-number variants (CNVs) in this locus, we developed a method to analyse CNVs using principal component analysis of the raw intensity values of single nucleotide polymorphisms (SNPs) genotyped on the Immunochip platform. This platform includes 1,159 SNPs in the SD block of FcγR genes; of these only 140 (12%) passed our quality control for SNP analysis but their intensity values are informative for the CNVs estimation. We identified several CNV loci in the FcγR block. We confirmed our results via an independent method - arrayCGH genotyping - and observed a perfect correlation in CNV estimation between both methods. Next, we applied our method to case-control cohorts of rheumatoid arthritis (RA), celiac disease and inflammatory bowel disease (in total, 15231 individuals). We found no associations between CNVs with these diseases ($p > 0.05$), however individuals with RA more often had a rare complete deletion of FCGR3B. Additionally, by performing functional studies we observed a correlation between the number of FCGR3A gene copies and FCGR3 (CD16) expression on T-cells. Conclusion: We have developed a method to accurately estimate CNVs based on SNP intensity data that can be extended to other phenotypes and to other SD loci in the human genome. We have also established the functional effect of the number of copies of FCGR3A gene on the CD16 expression on T-cells.

462W

An improved diagnostic workflow using SNP arrays reveals novel insights into CNVs in mental retardation. *X. Zhou[2,1], JM. Wang[1], B. Wang[3], J. Wang[3], T. Ji[1], X. Zhang[2], X. Ma[3], Y. Jiang[1].* 1) Department of Pediatrics, Peking University First Hospital, Beijing; 2) Bioinformatics Division, Tsinghua National Laboratory of Information Science and Technology, Tsinghua University, Beijing; 3) National Research Institute For Family Planning, Beijing.

Chromsomal microarray analysis has been widely adopted among clinicians for the genetic testing of children with developmental delay or mental retardation (DD/MR). While copy number variants (CNVs) can be readily detected by array CGH or SNP chips, distinguishing pathogenic CNVs from benign CNVs is challenging and requires considerable skills. SNP arrays can also generate genotype information, but may suffer from low signal-to-noise ratio; so it is not clear how to make best use of this platform. We developed an improved diagnostic workflow for this task by introducing the following three innovations: (1) Combining CNV discovery and targeted genotyping to improve the sensitivity of detecting small CNVs; (2) Using an in-house database CNVs ascertained by similar platforms and computational protocols from over 5,000 samples; (3) Judging the pathogenicity of CNVs by the number of affected genes or cross-reference genes or chromosomal regions to a curated phenotype ontology. We applied our diagnostic workflow on 288 DD/MR patients genotyped with Illumina CytoSNP-12 bead arrays. We identified pathogenic CNVs in 38 patients (25 males, 13 females). In addition to very large CNVs or CNVs corresponding to known chromosomal syndromes, we were able to detect small CNVs (<250kb) in five patients (1.7%) disrupting a single gene, three of which were known to cause monogenic forms of DD/MR. Examining the patient's clinical record supported their pathogenicities. Most published studies used samples of European ancestry; conclusions drawn from previous studies remained to be replicated in other populations. In our cohort of Chinese patients, we replicated the previous finding that DD/MR patients, especially for those with multiple congenital anomalies, have higher burden of rare and large CNVs as compared with natural populations. Most large and rare CNVs were de novo, and preferentially originated from paternal germline mutations. But not all de novo CNVs can be pathogenic: the rate of de novo CNVs identified as non-pathogenic by our workflow was not significantly different from the natural populations. We also discovered in four cases, pathogenic CNVs were inherited from mothers; and one pathogenic CNV in the mother but not transmitted to patients. It can be explained if females are more tolerable to pathogenic CNVs than males.

463F

Copy number variation in autism spectrum disorders. *B. Sheppard[1], C. Ladd-Acosta[1], B.K. Lee[2], J. Bonner[3], G. Windham[4], L. Schieve[5], L. Croen[6], A. Reynolds[7], D. Schendel[5], C. Newschaffer[2], M.D. Fallin[1].* 1) Department of Epidemiology, Johns Hopkins School of Public Health, Baltimore, MD; 2) Drexel University School of Public Health, Philadelphia, PA; 3) Michigan State University Biomedical Research Informatics Core, East Lansing, MI; 4) California Department of Public Health; 5) Centers for Disease Control and Prevention, Atlanta, GA; 6) Kaiser Permanente Northern California Division of Research; 7) University of Colorado Denver School of Medicine, Denver, CO.

There is considerable emerging evidence for the role of copy number variation (CNV) in autism spectrum disorder (ASD) susceptibility. In order to further characterize variation in ASD, we assessed genome-wide CNV burden, genic CNV burden, and regional CNV associations at previously implicated loci in ASD cases compared to developmentally healthy controls from the Study to Explore Early Development (SEED). Approximately 900 subjects from the SEED cohort were genotyped at nearly 1 million single nucleotide polymorphic markers (SNPs) and nearly 130,000 monomorphic copy number probes on the Illumina HumanOmni1 array. A hidden Markov model (HMM) approach was then used to detect CNVs using the SNP array data. Burden was assessed by case-control differences in CNV rate and size, with respect to frequency, length, and type. Regional associations were preformed at previously implicated ASD loci, including 7q11.23, 16p11.2, 15q11.2-q13.1, 15q13.2-q13.3, 22q13, NRXN1, and CNTN4. Overall, cases exhibited larger CNVs when compared to controls, particularly large duplications. Six large (> 400 kb) genic CNVs were identified; two were found in CNTN4 and one in PARK2; both previously implicated ASD genes involved in neuronal-cell adhesion and ubiquitin pathways, respectively. However, no significant associations were detected in the seven previously implicated CNV and gene regions. These results support many previous findings, especially with respect to evidence for larger CNVs in cases compared to controls and large rare events affecting neuronal-cell adhesion and ubiquitin gene networks. Future research is needed to confirm these findings as well as to implicate additional genes and pathways that could elucidate the biological mechanisms involved in ASD. We hope to supplement this analysis with additional SEED samples in the near future.

464W

Understanding revisions to the human reference genome and assembly model. *D. Church[1], P. Flicek[2], T. Graves[3], T. Hubbard[4], V. Schneider[1], The Genome Reference Consortium.* 1) Natl Ctr Biotech Infor/NIH, Bldg 45 rm 5AS43, Bethesda, MD; 2) EBI, Hinxton, Cambridge, UK; 3) The Genome Institute at Washington University, St. Louis, MO; 4) The Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK.

The publication, more than a decade ago, of an assembly for the human genome was a milestone event in biology. This resource has transformed basic and clinical research. One of the key insights into human biology made possible by the reference assembly was the discovery of an unrecognized degree of genetic variation among individuals. In today's era of whole genome sequencing, alignment of next generation sequencing reads against the high quality reference assembly remains a critical step in the interpretation of variation data. However, these analyses have made it increasingly clear that the linear chromosome models used in the human reference assembly do not always adequately represent the most variant and complex regions of the human genome. To address this issue, the Genome Reference Consortium (GRC; http://genomereference.org), the group overseeing the human reference assembly, developed a new assembly model. GRCh37, the current reference assembly, was the first to use this model, which has now also been implemented in the recently released mouse reference assembly, GRCm38. This model retains the intuitive linear chromosomes of previous assemblies, but now provides alternate assembly representations for more diverse and complicated genomic regions that are placed in a chromosome context via alignment. Assembly patches comprise the second key feature of the assembly model. Representing novel sequences and assembly corrections, the patches enable the GRC to provide timely updates to the reference assembly without disruptive coordinate changes. Like the alternate assemblies, the patches are stand-alone scaffold sequences placed in chromosome context via alignment. Released quarterly, there have been 10 patch releases associated with GRCh37, including more than 71 novel sequence representations and greater than 69 assembly corrections. We will show how the adoption of this assembly model has enabled GRCh37 to capture substantial amounts of human sequence not represented in previous assembly versions. We also will present data showing how the use of the alternate loci and patches improves the ability of the reference to act as an alignment substrate and discuss the need for new analytic resources to take advantage of these assembly features. Lastly, we will discuss ongoing GRC efforts to address assembly issues, including the use of single haplotype resources to resolve complex regions and 1000 genomes data to update rare or erroneous bases.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

465F

Evolutionary history and genome organization of DUF1220 protein domains. *M. O'Bleness[1], C.M. Dickens[1], D. Albracht[2], H. Kotkiewicz[2], T. Graves[2], R. Wilson[2], J. Sikela[1]*. 1) Biochemistry, Univ Colorado Denver AMC, Aurora, CO; 2) The Genome Institute at Washington University School of Medicine, St. Louis, MO, 63108 USA.

DUF1220 protein domains exhibit the most extreme human lineage-specific (HLS) copy number increase of any protein coding sequence in the human genome. From >270 copies in human, DUF1220 copy number markedly decreases as a function of a species evolutionary distance with human (125 in chimp, 35 in macaque, and 1 in mouse/rat) and shows a strong evolutionary correlation with brain size (R2=0.98; p=1.8 × 10-6). The dramatic HLS increase in DUF1220 copy number was driven by domain hyperamplification involving a 4.7 kb, tandemly repeated three DUF1220 domain unit called the HLS DUF1220 triplet. All copies of the HLS DUF1220 triplet lie in a dynamic genomic environment characterized by a pericentric inversion (1p12-1q21.2) and a polymorphic heterochromatin expansion (1q12) that are unique to the human genome. The great majority of DUF1220 sequences map to 1q21 and DUF1220 copy number loss has recently been implicated in 1q21-associated microcephaly, further strengthening the association of DUF1220 copy number and brain size. An accurate characterization of DUF1220 domains and their genome organization within the 1q21 region is critical to understanding the mechanism behind the HLS DUF1220 triplet expansion and its evolutionary importance. However a comprehensive analysis of DUF1220 domains and the gene family that encodes them (NBPF) has been hindered by the incomplete nature of the reference assembly of the 1q21.1-q21.2 region. As part of an effort to create accurate sequence assemblies of complex genomic regions, a new 1q21 assembly has been generated by targeted 1q21 sequencing of a haploid hyditiform mole BAC library. This new resource has allowed the 1q21.1-q21.2 region to be closed to just a few remaining gaps. Analysis of this new assembly, combined with the growing number of available primate genome assemblies, has allowed the most comprehensive evolutionary history of DUF1220 domains and their genomic organization to be generated.

466W

Retrotransposition of gene transcripts leads to structural variation in mammalian genomes. *A.D. Ewing[1], T. Ballinger[1], D. Earl[1], D. Haussler[1,2].* 1) Center for Biomolecular Science and Engineering, University of California, Santa Cruz, CA; 2) Howard Hughes Medical Institute, University of California, Santa Cruz, CA.

Processed pseudogenes are a source of material for new gene formation on evolutionary time scales. Most prior work on pseudogene discovery involves comparing copies that are present in reference genome assemblies to their parent genes. Here, we explore pseudogene insertion variants that are present as segregating insertions in the germlines of individual humans and mice. Through analysis of whole-genome sequence data, we find evidence for 39 retroposed transcripts in the genomes of one or more of over 900 humans sequenced at low coverage as part of the 1000 Genomes Project, but not present in human reference assembly GRCh37. Similarly, we find evidence for 755 pseudogene insertions at distinct locations present in one or more of 17 inbred mouse strains but not present in the C57BL/6J reference assembly NCBI m37, and 23 pseudogene insertions across a cohort of 10 chimpanzee genomes not present in the panTro2 reference genome assembly. Many of these insertions are new members of existing pseudogene families whose progenitors are highly and widely expressed, and the majority of these insertions have detectable hallmarks of processed pseudogene formation. Using these data, we estimate the rate of novel germline pseudogene formation in humans at roughly 1 new pseudogene for every 6000 individuals.

467F

Whole genome sequencing of iPS cells reveals somatic mutation arising at latest stages of reprogramming. *M. Cortes, J. S. Seo*. Seoul National University College of Medicine, GMI, Seoul, South Korea.

Purpose: Fibroblasts Induced to pluripotency with the four factors are going trough different steps of development, with an uncertain fate at the beginning of Area 51 (Day 8 to 15) as described by Nagy A. & al. (Nature 2011), whereas their fate seems determinate after Day 16. We therefore performed a whole genome sequencing (WGS) at high throughput of stages before, between and after Area 51 to understand mechanisms involved in the reprogramming process. Methods: We generated a WGS with the Next Generation Sequencer (NGS) Illumina HighSeq 2000 at a high coverage of 28.95X of 6 samples, primary iPSC (1iPS), three intermediate stages of reprogramming from 1iPSC to the secondary iPSC (2iPSC) at Day 0 (D0), Day 11 (D11) and Day 18 (D18) and finally of the secondary iPSC. Databases were generated with MySQL to store and handle the analyses of large data files from the NGS; CNVs were called with CNVnator, Breakdancer and SNSD (publication ongoing). We also performed SNPs genotyping with our in-house SNPs caller. Results: Unlike recent published data, we identified more than 6 millions mutations per sample and we found 379 de novo non synonymous SNPs (nsSNPs) arising at a later stage of reprogramming, for which Gene Ontology (GO) revealed that 3 genes Nkap, Pcm1 and Fgfr2 responsible for stem cell development differentiation and division. Meanwhile early stage of reprogramming didn't show more than 32 de novo nsSNPs affecting cell membrane proteins. With SNSD we found 6,028 germline CNV losses present in all stages of reprogramming, where 28 Genes identified with DAVID are involved in Cell morphogenesis in the differentiation process (p-value 4.5E-37). We found 11 intragenic somatic CNVs arising between 1iPSC and 2iPSC, one gene affected Prox1 is directly involved in stem cell maintenance, thus our analysis with Breakdancer shows an increase of Copy Number when cells are entering Area 51 followed by a dramatic decrease in Copy Number at D18, the end of Area 51. Conclusion: Unlike previously reported, our results show that later stages of reprogramming contain more genetic aberrations than early stages, thus we discovered SV in Nkap, Pcm1, Fgfr2 and Prox1 that could influence stem cell fate at the end of Area 51. The way these findings change the understanding of the iPSCs genome alter approaches to induced pluripotent stem cells and provide a new look at mechanisms responsible for reprogramming to pluripotency.

468W

Possible genomic impact of SVA retrotransposon in humans similar to that in gibbons. *T. Hara, Y. Hirai, H. Hirai, A. Koga.* Primate Research Institute, Kyoto University, Inuyama City, Japan.

The superfamily Hominoidea (hominoids) comprises the following two families: Hominidae (hominids) and Hylobatidae (gibbons, also called small apes). SVA transposon is a composite retrotransposon that shares a common lineage with that of hominoids and is considered to have been generated by the fusion of three genetic elements: SINE-R, a variable number of tandem repeat (VNTR) sequence, and *Alu*. This retrotransposon has been reported to cause various human diseases, including Fukuyama muscular dystrophy and X-linked dystonia-parkinsonism. We have recently discovered two SVA-related genetic occurrences that probably affect the genome structure of gibbons.

One occurrence was the large-scale expansion of tandem repeat sequences in the centromere regions in the hoolock gibbon (*Hoolock hoolock*). The repeat sequences are 40 kb or more in length and are carried by 28 of the 46 chromosomes of the somatic cells. The sequences consist of 35- to 50-bp repeat units and exhibit a sequence similarity with the VNTR region of SVA. Thus, it has been inferred that the large-scale tandem repeats in the centromere regions were derived from the SVA transposon, or that the repeat sequences served as a source for SVA. If the former is true, similar events may also occur in the human genome due to the presence of numerous SVA copies scattered throughout.

The second occurrence observed was the proliferation of a new composite retrotransposon in the genome of the white-cheeked gibbon (*Nomascus leucogenys*). The basic structure of this transposon is the same as that of SVA with one prominent difference being the presence of part of a single-copy nuclear gene in place of the SINE-R region. Thus it has been inferred that the new retrotransposon was formed in the gibbon genome independent of SVA but by similar mechanisms. The copy number of this new retrotransposon was estimated to be of the order of 10^2 and probably exceeds the copy number of SVA in the same genome; thus, it may have a higher transposition frequency than SVA. It is suggested that formation of new composite retrotransposons could also occur in the human genome due to the presence of respective solo elements.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

469F

Functional studies of tandem repeat variation in the human genome using multiplexed capture and high-throughput sequencing. *A. Guilmatre[1], G. Highnam[2], D. Mittelman[2, 3], A.J. Sharp[1].* 1) Genetics and Genomic Sciences, Mount Sinai School of Medicine, New York, NY; 2) Virginia Bioinformatics Institute, Virginia Tech, Blacksburg, VA; 3) Department of Biological Sciences, Virginia Tech, Blacksburg, VA.

Tandem repeats (TRs) are stretches of DNA sequence comprised of two or more contiguous copies of a sequence of nucleotides arranged in a head-to-tail pattern, and comprise ~2% of the human genome. They are characterized by a high mutation rate due to replication slippage, and represent an important source of genetic variation in most genomes. A recent study in yeast showed that genes containing TRs in their promoters display high transcriptional divergence and that variations in repeat length were associated with changes in expression level, suggesting a functional role for TRs in regulating gene expression. However, TRs have long been considered as mere junk DNA, and due to a number of technical issues in genotyping them remain poorly studied. Furthermore, GWAS provide little information on repeat variation because most TRs are poorly tagged by SNPs. To date, whole-genome sequencing is the only high-throughput technique enabling large-scale TR genotyping but it remains expensive. Therefore, to facilitate studies of the role of TR variations in human, we have developed a method to perform targeted enrichment and high-throughput sequencing of TRs, and applied this to study TRs located in gene promoter regions in the human genome. By adapting the exome capture method to enrich TR regions and using a specific bioinformatic pipeline designed for genotyping TRs from sequencing reads (Fondon et al 2012, PLoS One v7:e33036), we attained a mean of 319 reads per target. One third of these completely span the TR and were therefore informative, allowing >4,000 promoter-associated TRs to be genotyped, of which 79% were polymorphic. Our method allows multiplexing of ≥8 individuals per lane of HiSeq, enabling the cost-effective study of thousands of repeats on a population level. We are also estimating potential PCR biases by sequencing hydatidiform moles that are homozygous genome-wide, studying inheritance patterns using trios, and utilizing PCR-free library preparation to ensure genotype accuracy. Based on the success of our preliminary experiments we are currently studying the effect of TR variations on local gene expression by performing sequencing of TRs in HapMap individuals. These studies will give major insights into the functional effects of TR variations in the human genome.

470W

Meta-analysis with dbGaP Data on Genome-wide Association Studies for Alzheimer's Disease. *J. Kim.* Ilsan hospital, Koyang, South Korea.

The genome-wide association studies (GWAS) usually find risk genes with low odd ratios because they target common variants. The GWAS of Alzheimer's disease (AD) found several risk genes during recent five years. However, the odd ratios of these genes for AD are so low that we cannot explain the heritability of AD, yet. The imputation with 1000 genomes can be an alternative solution. The 1000 genomes project provides detail human genome and rare variants. We will try to find rarer SNPs related with AD by using imputation of 1000 genomes. The most strong association was found in ApoE. The other assocciated genes for Alzheimer's disease were CR1 and CD2AP which were founded in other GWAS studies.

471F

A survey of the mutational load in genes associated with rare Mendelian disorders and autism. *A.R. Rao, S.F. Nelson.* Department of Human Genetics, UCLA Geffen School of Medicine, Los Angeles, CA.

In this study, we surveyed the number of individuals in human populations with a severely protein-damaging rare variant in genes associated with a rare Mendelian disorder or autism. Genetic variants that severely disrupt protein function are of clinical interest, and it is important to accurately predict whether a variant is a genuine loss of function (LoF) variant. Since we expect these variants to be rare, the abundance of such variants in a given gene across a human population provides a basis for validating whether they are in fact severely protein-damaging. We combined data from the 1000 Genomes Project, which included exomes of 1092 individuals, and the NIEHS Environmental Genome Project, which consisted of the exomes of 95 samples. For each gene with a known inheritance model, we calculated the number of individuals that are heterozygous or homozygous for the minor allele of a variant predicted to be protein-damaging. Under the assumption that rare deleterious missense mutations—predicted to be damaging by various protein function prediction methods—, stop codon gain/loss, inframe codon gain/loss, splice site disrupting and frameshift variants are protein-inactivating, the number of individuals with supposed disruptions in protein function far exceed our expectation, given the known prevalence of various Mendelian disorders. If we exclude the missense mutations and inframe codon loss/gain variants, the number of individuals with a LoF variant in a given gene is closer to the expected value. Additionally, the remaining variants in genes with surprisingly large counts may be filtered further, e.g. stop gain variants may be vetted based on their proximity to the end of a gene. Our results suggests that most missense mutations that are predicted to be deleterious are not severely damaging and may simply increase risk towards common, complex disease, at most. Furthermore, our method of analysis illustrates how one may filter LoF variants beyond the common method of comparing minor allele frequencies, in future studies that aim to discover the gene for a rare Mendelian disorder of unknown cause.

472W

Inferring mechanisms of indel mutation from whole genome sequence. *I.HC. Turner[1], Z. Iqbal[2], G. McVean[1,2].* 1) Department of Statistics, University of Oxford, Oxford, United Kingdom; 2) The Wellcome Trust Centre for Human Genetics, Roosevelt Drive, Oxford, OX3 7BN, United Kingdom.

Mutations in DNA arise through the combination of damage, errors during replication (e.g. slippage) and transcription, and the cell pathways that attempt to repair DNA damage. Although the mutational signatures of DNA repair mechanisms are understood in specific cases in model organisms, their relative contributions to the germline are unknown. Inferring contributions from sequence data has previously been impossible as mutation discovery was biased towards shorter events and by the use of a reference genome. We use a new variant discovery method based on de novo assembly to provide unbiased variant discovery from whole genome sequence, thereby allowing us to reason about underlying cellular processes (focusing on those causing insertions and deletions).

We made variant calls in humans (85 Luhya from the 1000 Genomes Project [1]), chimpanzees (10 from PanMap [2]) and in birds. These calls consistently show that twice as many deletions as insertions have occurred, although their size distributions are similar. Almost all insertions (90%) are tandem and most (93%) non-tandem events are deletions. The high incidence of micro-homology in deletions, between flank ends and deletion edges, indicates a 2bp homology requirement for error-prone DNA repair. Indels of size 4bp occur at an increased rate in some species, possibly due to up-regulation and preference of the Ku pathway. Increases in the number of variants at the 3' end of genes may be evidence of Transcription Associated Mutagenesis in the germline.

From a variant's sequence and context we infer the most likely pathway responsible. Using this we quantify the relative effect of cellular processes on the genome. We describe a statistical model to explain double-stranded break repair by the Non-Homologous End Joining (NHEJ) pathway, which we parameterize using discovered variants. The model can explain the mutation patterns between species, and is validated by comparison with experimental data.

[1] A map of human genome variation from population-scale sequencing, Nature (2010) [2] A Fine-Scale Chimpanzee Genetic Map from Population Sequencing, Science (2012).

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

473F

A complex rearrangement of the OCA2 gene in two unrelated albino patients. *B. Arveiler[1,2], E. Lasseaux[1], D. Cailley[1], M.L. Vuillaume[1,2], C. Rooryck[1,2], A. Rouault[1], J. Toutain[1], D. Lacombe[1,2], F. Morice-Picard[1,2].* 1) Service de Génétique Médicale, CHU de Bordeaux, Bordeaux, France; 2) Univ Bordeaux Segalen, Maladies Rares: Génétique et Métabolisme (MRGM), EA4576, Bordeaux, France.

Oculocutaneous albinism (OCA) is an autosomal recessive disease characterized by hypopigmentation of the skin, hair and eyes. Molecular analysis allows to classify patients, depending on the gene that is mutated, into OCA1 (TYR gene), OCA2 (formerly called P gene), OCA3 (TYRP1) and OCA4 (SLC45A2). We routinely perform an exhaustive analysis of the 4 genes, including a search for point mutations and for intragenic microrearrangements. Deletions and duplications are investigated using ultra-high resolution array-CGH, in which the 4 genes are covered at a density of 1 probe/102bp on average (1). We report two patients with a complex rearrangement. In patient 1 mutation p.Asn489Asp was identified in the heterozygous state. Array-CGH analysis indicated the presence of two heterozygous deletions restricted to introns 2 and 20 respectively, the intervening segment being present in two copies. Both deletions were inherited from her mother, thus indicating that both deletions were on the same allele. The patient's father was heterozygous for the p.Asn489Asp mutation. Patient 2 was heterozygous for the same deletions of introns 2 and 20 as patient 1. No other mutation was found. The patient's mother did not harbor these deletions. His father was not available for study. Remarkably, the 5' breakpoint of the intron 2 deletion and the 3' breakpoint of the intron 20 deletion seemed to be at the same position as the breakpoints of an intron 2–20 deletion that we identified previously in three patients of polish origin (1). A haplotype analysis showed that the deletions in patients 1 and 2 were on the same haplotype as that associated with the intron 2–20 deletion in the polish patients. This suggested that instead of having a bipartite deletion the patients may in fact have the “polish” intron 2–20 deletion, with the intervening segment being inserted somewhere else in the genome. A PCR-sequence analysis performed with primers located on each side of the intron 2–20 deletion confirmed that patients 1 and 2 indeed harbored the “polish” deletion. FISH analysis using a probe covering the exon 3–20 segment identified no exogenous signal, but two signals at the OCA2 locus, therefore suggesting that the intervening segment was inserted inside of, or immediately next to, the OCA2 gene. High throughput sequencing of the entire 15q11–13 region is underway to determine the exact nature of the rearrangement(s). (1)Rooryck et al., Hum Genet (2011) 129:199–208.

474W

A four base pair insertion in WNT16a identified through exome sequencing is associated with type 2 diabetes. *L. Been[1], M. Fortner[1], E. Howard[2], M. Lerner[3], D. Brackett[4], D. Sanghera[1].* 1) Pediatrics/Gen, Univ Oklahoma HSC, Oklahoma, OK; 2) Cell Biology, Univ Oklahoma HSC, Oklahoma, OK; 3) Surgery, Univ Oklahoma HSC, Oklahoma, OK; 4) Veteran Affairs, VA Medical Center, Oklahoma, OK.

Several *in vitro* and *in vivo* studies have shown that many components of the WNT pathway are involved in β cell proliferation, insulin secretion, and production of glucogon-like peptide-1 (GLP-1). We hypothesize that any alternation in WNT pathway will have profound consequences in insulin secretion and generation of new β cells, as this pathway needs to be tightly regulated. A four base pair frame-shift insertion in the *WNT16a* gene was detected in the conserved region through a genome-wide exome sequencing carried out on two participants (one type 2 diabetic (T2D) case and one control). A genetic screening of this insertion in 1,512 consisting of 792 type 2 diabetic cases and 720 normoglycemic control individuals suggested that the number of carriers of this insertion did not differ significantly among cases (18%) than in controls (15%) (p=0.128). However, the insertion carriers showed significantly higher levels of serum C-peptide (β=0.087, p= 0.001) when compared to the wild type carrier. In gene expression studies on 18 human pancreatic tissues, our data revealed a 2.86 fold increase in the expression of WNT16a among insertion carriers compared to non-carriers. Further *in vitro* transfection assays reformed by cloning ~1.2-kb 5' region of the putative *WNT16a* promoter into luciferase reporter vector showed 70 fold increase in promoter activity associated with the insertion carriers as compared to the wild-type carriers. The WNT signaling regulates and stabilizes β catenin binding with *TCF7L2*,which is critical for expression of *GLP-1* and many other genes involved in β cell development, our results suggest that *WNT16a* insertion may play a significant role in T2D pathogenesis by altering the transcription of WNT pathway genes involved in glucose homeostasis.

475F

Obstetric complication-associated ANXA5 promoter polymorphisms affect gene expression via DNA secondary structures. *H. Inagaki[1], S. Ota[1], H. Nishizawa[2], H. Miyamura[1,2], K. Nakahira[3], M. Suzuki[1], S. Nishiyama[2], Y. Udagawa[2], I. Yanagihara[3], H. Kurahashi[1].* 1) ICMS, Fujita Health University, Toyoake, Aichi, Japan; 2) Dept Obstetrics and Gynecology, Fujita Health University, Toyoake, Aichi, Japan; 3) Dept Developmental Medicine, Osaka Medical Center and Research Institute for Maternal and Child Health, Izumi, Osaka, Japan.

Recent findings have highlighted the possibility that polymorphisms within the annexin A5 gene (ANXA5) promoter contribute to the etiology of various obstetric complications. However, the underlying mechanisms are unknown. In our present study, the transcriptional activity of the ANXA5 gene promoter was examined in the context of G-quadruplex formation. The M2 haplotype of the promoter polymorphisms that confers a high risk of onset for these disorders shows lower activity and less expression of ANXA5 mRNA. This gene promoter region has a motif that potentially forms a G-quadruplex structure, and the M2 allele possesses less potential than the major allele. The promoter activity determined by luciferase reporter analysis correlated with the in vitro G-quadruplex propensity estimated by circular dichroism. Treatment with a G-quadruplex ligand inhibited this promoter activity. The ANXA5 promoter region was found to be hypomethylated even on the M2 allele. Taken together, our data indicate that the M2 allele instigates G-quadruplex-mediated transcriptional regulation and consequently downregulates the placental anticoagulation factor annexin A5 levels, leading to obstetric disorders. Our data lend support to the developing paradigm that genomic variation affects gene expression levels via DNA secondary structures leading to the disease susceptibility.

476W

The Influence of Genomic Context on Mutation Patterns Inferred from Extremely Rare Variants. *V.M. Schaibley[1], M. Zawistowski[2], D. Wegmann[3], D. Kessner[4], M.G. Ehm[5], M.R. Nelson[5], P.L. St. Jean[5], G. Abecasis[2], J. Novembre[4], S. Zöllner[2], J.Z. Li[1].* 1) University of Michigan, Department of Human Genetics, Ann Arbor, MI, 48109; 2) University of Michigan, Department of Biostatistics, Ann Arbor, MI, 48109; 3) School of Life Sciences, École Polytechnique Fédérale de Lausanne, Switzerland; 4) University of California Los Angeles, Department of Ecology and Evolutionary Biology, Los Angeles, CA, 90095; 5) GlaxoSmithKline, Quantitative Sciences, Research Triangle Park, NC, 27709.

Spontaneous mutations drive human genome evolution and disease patterns, but our understanding of the mutation process is limited. Previous studies have shown correlations between both nucleotide diversity in humans and inter-species substitution rates with local GC content and recombination rate, suggesting that genomic context impacts mutation patterns. However, these data are influenced by natural selection, biased gene conversion, and demographic history. A new approach to studying spontaneous mutation patterns is to examine extremely rare variants, which have arisen recently and therefore are relatively less affected by the potentially confounding effects of demography, selection, and biased gene conversion. We utilized rare variant data that was previously obtained by sequencing the exons of 202 genes in genomic DNA samples of >14,000 individuals to identify genetic associations with disease. This dataset contains >19,000 variants with an allele frequency $\leq 10^{-4}$ in individuals of European ancestry. We quantified the effect of GC content and recombination rate on the total mutation rate and each mutation subtype (GC>AT, GC>TA, etc.). Our results show that mutations, as inferred from rare variants, toward a G or C base pair (AT>GC or AT>CG) are less likely in regions of high GC content. This trend is not observed among common variants or substitutions. Rather, the probability of observing a common variant or substitution toward an A or T base (GC>AT or GC>TA) is decreased in GC-rich regions, which is not seen in rare variants. This suggests that GC content influences the initial mutation and the fixation probability of an existing variant differently. Our results also show that recombination rate has little effect on rare variants, but a significant impact on common variants and substitutions, especially the AT>GC and AT>CG subtypes, consistent with biased gene conversion primarily driving correlations between inter-and intra-species diversity with recombination. Understanding the spontaneous mutation process and the genomic context effects will aid in our understanding of observed patterns of DNA variation and in the subsequent interpretation of the downstream genotype-phenotype analyses. Our study demonstrates that rare variants emerging from deep sequencing studies substantially expand the data available for studying spontaneous mutation patterns.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

477F

Sequencing the Unsequenceable: Expanded CGG Repeats in the Human *FMR1* Gene. *P. Hagerman[1,2], J. Eid[3], P. Peluso[3], D. Rank[3], J. Yin[1], L. Hickey[1], E. Loomis[1]*. 1) Department of Biochemistry and Molecular Medicine, University of California, Davis, School of Medicine, Davis, CA; 2) MIND Institute, University of California, Davis, Health System, Sacramento, CA; 3) Pacific Biosciences, Menlo Park, CA.

Alleles of the *FMR1* gene with more than 200 CGG repeats generally undergo methylation-coupled gene silencing, resulting in fragile X syndrome, the leading heritable form of cognitive impairment. Smaller expansions (55–200 CGG repeats) result in elevated levels of *FMR1* mRNA, which is directly responsible for the late-onset neurodegenerative disorder, fragile X-associated tremor/ataxia syndrome (FXTAS). For mechanistic studies and genetic counseling, it is important to know with precision the number of CGG repeats; however, no existing DNA sequencing method is capable of sequencing through more than ~100 CGG repeats, thus limiting the ability to precisely characterize the disease-causing alleles. The recent development of single molecule, real-time sequencing represents a novel approach to DNA sequencing that couples the intrinsic processivity of DNA polymerase with the ability to read polymerase activity on a single-molecule basis. Further, the accuracy of the method is improved through the use of circular templates, such that each molecule can be read multiple times to produce a circular consensus sequence (CCS). We have succeeded in generating CCS reads representing multiple passes through both strands of repeat tracts exceeding 700 CGGs (>2 kb of 100 percent CG) flanked by native *FMR1* sequence, with single-molecule readlengths exceeding 12 kb. This sequencing approach thus enables us to fully characterize the previously intractable CGG-repeat sequence, leading to a better understanding of the distinct associated molecular pathologies. Real-time kinetic data also provides insight into the activity of DNA polymerase inside this unique sequence. The methodology should be widely applicable for studies of the molecular pathogenesis of an increasing number of repeat expansion-associated neurodegenerative and neurodevelopmental disorders, and for the efficient identification of such disorders in the clinical setting. New data will be presented at the Annual Meeting.

478W

Three novel NR5A1 gene mutations in patients with 46,XY karyotype and disorders of sex development. *H.C. Fabbri[1], J.G. Andrade[2], A.T. Maciel-Guerra[2,3], G. Guerra-Junior[2,4], M.P. de Mello[1]*. 1) Centro de Biologia Molecular e Engenharia Genética (CBMEG), UNICAMP, Campinas, Brasil; 2) Grupo Interdisciplinar de Estudos da Determinação e Diferenciação do Sexo (GIEDDS), UNICAMP, Campinas, Brasil; 3) Departamento de Genética Médica, UNICAMP, Campinas, Brasil; 4) Departamento de Pediatria, UNICAMP, Campinas, Brasil.

Disorders of sex development (DSD) is the term used for congenital conditions in which development of chromosomal, gonadal, or anatomic sex is atypical. The DSD with 46,XY as karyotype are characterized by ambiguous or female external genitalia, dysgenetic testes, and the presence or absence of Mullerian structures. Among them, we can mention the 46,XY DSD (previous male pseudohermaphrodite, XY male undervirilization, or XY male undermasculinization) and the XY gonadal dysgenesis. The gonadal dysgenesis refers to a set of abnormalities in individuals who have dysgenetic gonads caused by defects in genes involved in the embryonic gonadal development. Several genes participate in the cascade of sex determination and differentiation, and defects in some of them may lead to dysgenesis. SRY is the first candidate gene to be investigated is such cases, however only 15% present mutations in SRY gene. The NR5A1 gene, which encodes the SF-1 transcriptional regulator, has been described as resposible for DSD phenotypes. Variations in NR5A1 had been identified in association with DSD, hypospadias, anorchia, male infertility, female primary ovarian insufficiency and, in some cases, with adrenal tumors and endometriosis. In this study we report the molecular evaluation of NR5A1 gene in 67 Brazilian patients with XY gonadal dysgenesis. Direct sequencing of the 7 exons, including the promoter region and intron/exon boundaries and 3'UTR was performed. The study reveled three novel heterozygous mutations. Two of them had been identified in patients with 46,XY partial gonadal disgenesis: the c.94G>A nucleotide change within exon 2 resulted in p.S32N mutation and the C.1138+1G>T which is located in the splice donor region of intron 6. The p.C65Y resulted from the c.195G>A within exon 3 and was identified in three siblings with 46,XY DSD. Both p.S32N and p.C65Y occur within SF-1 DNA-binding domain and might interfere with its transcriptional activity. The c.1138+1G>T suppresses the consensus splice donor site and probably leads to exon 7 skipping. The findings described here show the variability in phenotype presentations for different NR5A1 mutations. Therefore, more research must be done to elucidate the important role of NR5A1 in DSD patients.

479F

A deletion and a novel nonsense SRY mutation in 46,XY Brazilian patients with disorder of sex development. *CSC. Piveta[1, 5], BPB. Araújo[2], PD. Pereira[2], FB. Coeli-Lacchini[1], AT. Maciel-Guerra[3,4], G. Guerra-Junior[3,5], IL. Monlleó[2], MP. de Mello[1]*. 1) CBMEG, UNICAMP, Campinas, São Paulo, Brasil; 2) Universidade Estadual de Ciências da Saúde de Alagoas, Brasil; 3) Grupo Interdisciplinar de Estudos da Determinação e Diferenciação do Sexo (GIEDDS). Universidade Estadual de Campinas, São Paulo, Brasil; 4) Departamento de Genética Médica.Faculdade de Ciências Médicas da Universidade Estadual de Campinas, São Paulo, Brasil; 5) Departamento de Pediatria. Faculdade de Ciências Médicas da Universidade Estadual de Campinas,São Paulo, Brasil.

Keywords : SRY, disorders of sex development, deletion, nonsense mutation. DSD (Disorders of sex development) are among the most common genetic diseases in humans and are characterized by incomplete or disordered gonadal or genital development, leading to divergences between genetic, gonadal and phenotypic sex, especially in individuals with 46,XY karyotype. In humans, male or female sexual development primarily depends on the presence or absence of the SRY gene (sex-determining region on Y) located on the short arm of the Y chromosome. The SRY gene contains a single exon and encodes a 204-amino-acid protein whose central 79 amino acids encode a HMG box (High Mobility Group) DNA binding domain. This transcription factor binds to target DNA sequences and bends it. The main regulatory activity of the SRY gene results from its DNA interaction and DNA flexing ability. Beside its regulatory role, SRY contains important sequences for SRY protein nuclear signaling. The majority of SRY mutations leading to XY gonadal dysgenesis occur within the HMG domain, predominantly causing defective binding to target DNA sequences or impairing the ability of the SRY protein to bend DNA. The objective of the present study was to investigate mutations in SRY in two Brazilian patients with XY gonadal dysgenesis. Genomic DNA was obtained from peripheral blood leukocytes and polymerase chain reaction (PCR) of SRY was performed. The first case presented total deletion of SRY gene sequences although other Y-chromosome sequences were preserved. A nonsense mutation within the HMG-box of the SRY gene was identified in the second patient: a cytosine was replaced by a thymine at position 340 (c.340C>T), resulting in the replacement of the glutamine (CAG) codon at residue 114 of by a stop codon (TAG) (p.Q114X). This mutation occurred de novo because the father is a normal male who carries a normal SRY copy. This point mutation is been described here for the first time. It can be assumed that this point mutation will result in a truncated nonfunctional SRY protein or that the protein will not translated due to the activation of the nonsense-mediated mRNA decay mechanism. Both hypotheses can explain the female phenotype and the 46,XY gonadal dysgenesis.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

480W

Whole genome sequencing in autism identifies hotspots for germline mutation. *J. Michaelson[1,2], H. Zheng[4], M. Gujral[1], Y. Shi[4], D. Malhotra[1,2], X. Jin[4], D. Greer[1], A. Bhandari[1], R. Corominas[2], A. Peoples[1,2,5], A. Koren[6], S. McCarroll[6], L. Iakoucheva[2], Y. Li[4], J. Wang[4], J. Sebat[1,2,3].* 1) Beyster Center for Genomics of Psychiatric Diseases; 2) Department of Psychiatry; 3) Department of Cellular Molecular Medicine, University of California, San Diego, La Jolla, CA 92093, USA; 4) BGI-Shenzhen, Shenzhen, China; 5) Trinity College Dublin, Ireland; 6) Department of Genetics, Harvard Medical School, Cambridge, MA.

De novo mutation plays an important role in Autism Spectrum Disorders (ASDs). Notably, most disease-associated copy number variants (CNVs) are characterized by high mutation rates and moderate-to-high penetrance. Extrapolating from these observations, we hypothesize that hypermutability is a general property of the major risk loci in ASD, and may also include nucleotide substitution hotspots. Clear evidence for the latter is lacking however because the magnitude of variability in base substitution rates is not known. We investigated global patterns of germline mutation by whole genome sequencing of 10 monozygotic (MZ) twins concordant for ASD and their parents, and we examined the genetic association of genes and nucleotide substitution hotspots with disease in independent datasets. A total of 581 germline *de novo* mutations (DNMs) were detected in 10 MZ twin pairs, and DNMs displayed a remarkably non-random positioning in the genome ($P < 10^{-10}$). We applied a simple statistical approach to clustering mutations and defined > 20 regions where DNMs occur at significantly high density. Notably, some hotspots included genes that have been previously implicated in ASD. Five of the 26 genes found in our study were impacted by mutations in an independent dataset from recent exome sequencing studies of ASD (7 DNMs in cases and 0 in controls), a higher degree of gene overlap than would be expected by chance ($P=10^{-7}$). In addition, ASD cases showed a 3-fold enrichment of DNMs in hotspots compared to controls. The observed distribution of DNMs could be explained in part by intrinsic properties of the genome. On the whole, DNMs were biased toward regions with reduced GC-content, low nucleosome occupancy, and we observed a weak bias toward late-replicating, heterochromatic regions. However, DNMs were also enriched within certain high-GC features, including recombination hotspots and transcribed sequences. Our findings suggest that rates of nucleotide substitution vary widely throughout the genome as a function of DNA sequence elements and their chromatin states. In addition, regional hypermutability is a significant factor shaping patterns of genetic variation and disease risk in humans.

481F

Molecular Basis of Male Infertility: An Indian Perspective. *H. Sharma[1], R. Prasad[1], R. Mohan[2], S.K. Singh[2].* 1) Dept. of Biochemistry, PGIMER, Chandigarh, India; 2) Dept.of Urology, PGIMER, Chandigarh, India.

The majorities of males affected with Cystic Fibrosis (CF) are infertile due to high incidence of mutated Cystic Fibrosis Transmembrane Conductance Regulator (CFTR) gene, leading to congenital absence of Vas-deference (CAVD) and is considered as genital form of Cystic Fibrosis. However clinically affected CF patients present spectrum of genital phenotypes ranging from impaired spermatogenesis to CAVD and obstruction in reproductive tract other then vas-deferens or epididymis. Recently research has been focused to investigate the role of CFTR other then vas formation. The Present study was undertaken to establish the spectrum and frequency of CFTR mutation in Indian infertile male population and to understand the possible involvement of CFTR gene in the etiology of male infertility other than CAVD. Blood samples from infertile males due to CBAVD (n=25), obstructive Azospermia (n=25) and oligospermia or non-obstructive azoosermia(n=100) were used for genomic DNA isolation and screening mutations in seven exons (2,3,4,7,8,10,12) hot spots of CFTR gene. Delta F-508, N1303K, R553X, G551D, G542X, 621+1G>T and W1282X were the other known most common mutation screened in Indian infertile males through allele specific ARMS PCR analysis. IVS8-Poly T allele polymorphism was also determined in all infertile male patients. Out of 25 CBAVD males 20 (80%) were found to have mutation in one allele of CFTR gene in remaining 5 (20%) mutation could not be identified. Among 25 cases of obstructive azospermia 17(68%) were found to have mutated CFTR allele where as among 100 oligospermic or Non obstructive azospermia males only 17% were found to have mutation in CFTR gene. Delta F-508 (9.5%) and IVS8 (5T) allele (22%) were the most common mutations identified in Indian infertile male population. Moreover the increased frequency of CF mutations in males with reduce sperm quality and quantity and in males with azospermia without CBAVD suggest CFTR protein may be involved in the process of spermatogenesis or sperm maturation apart from playing critical role in the development of epididymal gland and vas-deferens.

482W

Phase-defined complete HLA genotyping by amplicon sequencing and read-backed phasing. *K. Hosomichi, T. Hayano, I. Inoue.* Division of Human Genetics, National Institute of Genetics, Mishima, Shizuoka, Japan.

The human leukocyte antigen (HLA) region, the 3.8-Mb segment of the human genome at 6p21, has been associated with more than 100 different diseases, mostly autoimmune diseases. Recently, the HLA region attracts attentions because strong drug adverse effects are associated with particular HLA alleles. Due to the complex nature of the HLA genes, classical direct sequencing method could not comprehensively elucidate genomic makeup of the HLA genes. We applied long ranged PCR to amplify six HLA genes (*HLA-A, -B, -C, DPB1, -DQB1* and *-DRB1*) followed by Nextera DNA Sample Preparation Kit and multiplex sequencing with MiSeq sequencer (illumina). 150 bp paired-end reads were aligned to the 6 HLA gene segments of hg19 on a condition that maximum 20 mismatches per read are allowed. In each heterozygous SNV position called from the first alignment, we classified the alleles using read-backed phasing : two alleles were divided into two phased haplotypes using information of the variants set on each paired-reads. According to these two phased haplotypes, we created two new self-reference sequences from an individual, which were applied for second stage alignment. After the second stage alignment, two HLA gene sequences based on phased SNVs and Indels were identified. Taking advantages of sequencing distantly separated fragments of a chromosome by paired-end sequencing at high depth read, and highly polymorphic nature of HLA genes, unambigous phasing becomes possible by a series of methods consisting of two alignment steps and read-backed phasing. Notably our phasing method does not require the IMGT/HLA database for creating HLA gene sequence unlike thus far HLA typing methods using Next-gen sequencers. In conclusion, we are able to determine the phase-defined complete HLA gene sequences including intron, upstream and downstream regions, even though alleles are rare or novel.

483F

Characterization of sequence variants from the first direct-to-consumer exome pilot project. *A. Shmygelska, E. Harrington, C. McLean, A. Chowdry, B. Naughton.* 23andMe Inc, Mountain View, CA.

We present the analysis of exomes of forty-five individuals. This data set represents a subset of 23andMe's exome pilot project data sequenced at high coverage (80x) using the Agilent SureSelect 50Mb targeted exome capture platform. We summarize the number and pattern of genetic variants emerging from the study, which is to our knowledge the first direct-to-consumer exome project. We focused particularly on loss of function (LoF) variants, i.e., variants predicted to disrupt the function of protein-coding transcripts. In order to assess functional potential of the variants, we developed a filtering and annotation pipeline that ensures that variants called meet quality control measures (including coverage depth, genotype quality, and sequence uniqueness). An average exome has ~30,000 called variants in the exonic regions, ~700 of which are LoFs. This category includes frame shifts, stop gains and losses, start losses, and splice site modifications. Non-synonymous substitutions represent ~11,000, and synonymous substitutions account for ~12,000 variants. We identified and characterized a set of high-quality LoF variants (~400 per exome) including their occurrence in a set of genes involved in Mendelian disorders, disruption of protein domains, frequencies, evolutionary conservation, and functional classes of LoF-containing genes. We found that rare mutations (< 5% allelic frequency) are located more often than expected in domains present in membrane-associated proteins. As did previous studies, we found that in a healthy individual the number of variants predicted to substantially impact proteins is on the order of hundreds and most are carried in the heterozygous state. We also found that LoF-containing genes are enriched to encode drug-metabolizing enzymes, particularly cytochrome P450 (CYP) genes (18 out of 57 genes coding for various CYPs contain LoFs, adjusted p=4.4e-5), and transport proteins (ATP-powered pumps, transporters, and ion channels; adjusted p= 8.8e-4). Additionally, we replicated previous findings of LoF enrichment in olfactory receptor and immune response genes. CYPs are a major enzyme class involved in the oxidative metabolism of a diverse set of molecules, including drugs, dietary chemicals and endogenous compounds. Together with ATP-binding cassette pumps, solute carrier transporters, and ion channels, CYPs play a key role in adverse drug interactions. We discuss further analysis of the LoFs found in drug response genes.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

484W

Targeted Sequencing Applications and Project Design. *R. Fulton.* The Genome Institute, Washington University School of Medicine, 4444 Forest Park Boulevard, St. Louis, MO.

With significant advances in DNA sequencing technology and targeted sequencing techniques, there has been an explosion in genomic applications exploiting these platforms. DNA sequencing not only has become significantly higher throughput and less expensive, but it also produces digital outputs that can be used for interpretation of sample characteristics, thus enabling many new applications. This presentation outlines current capabilities, including timelines, general cost ranges, example applications, lessons learned, and project highlights involving targeted sequencing projects, in an effort to spark collaborations and interactions aimed at further exploitation of these technologies for genomic discovery. Target designs range from gene lists and/or chromosomal regions identified through various research efforts including genome wide association studies, literature searches, or related studies. These regions can range from tens of kilobases to tens of megabases of targets, with study designs also varying from family studies, quantitative trait studies, case control studies, or sample panels as examples. Human genetics projects such as age-related macular degeneration, allelic spectrum of metabolic disorders, cleft lip, breast and prostate susceptibility, and retinitis pigmentosa highlight successes in discovery. Validation and extension successes are demonstrated by cancer resequencing efforts involving Acute Myeloid Leukemia, ovarian cancer, and breast cancer; while pediatric cancer, AML, and breast cancer project applications show the power of deep digital read counts, and demonstrates the ability to show clonality, copy number, and clonal progression.

485F

Inverted low-copy repeats and genome instability: A genome-wide analysis. *T. Gambin[1], P. Dittwald[2,3], C. Gonzaga-Jauregui[4], C. M. B. Carvalho[4], J. R. Lupski[4,5,6], A. Gambin[2,7], P. Stankiewicz[4,8].* 1) Institute of Computer Science, Warsaw University of Technology, Warsaw, Poland; 2) Institute of Informatics, University of Warsaw, Warsaw, Poland; 3) College of Inter-Faculty Individual Studies in Mathematics and Natural Sciences, University of Warsaw, Warsaw, Poland; 4) Dept of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX; 5) Dept of Pediatrics, Baylor College of Medicine, Houston, TX; 6) Texas Children's Hospital, Houston, TX; 7) Mossakowski Medical Research Centre, Polish Academy of Sciences, Warsaw, Poland; 8) Dept of Medical Genetics, Institute of Mother and Child, Warsaw, Poland.

To date, about 40 unstable genomic loci flanked by directly oriented LCRs have been associated with recurrent deletions and reciprocal duplications causing genomic disorders. Inverse Paralogous LCRs (IP-LCRs) can also cause genome instability both by acting as nonallelic homologous recombination (NAHR) substrates, in this case mediating balanced inversions of the intervening genomic intervals, and by facilitating complex DUP-TRP/INV-DUP rearrangements. When disrupting an overlapping dosage sensitive gene(s), balanced inversions can lead to abnormal phenotypes. We delineated the genome-wide distribution of IP-LCRs greater than 1-kb in size and 95% sequence identity and mapped the intersecting genes that overlap at least one of the IP-LCRs; such genes can be disrupted by the inversion rearrangement. Remarkably, 12.1% of the human genome is potentially susceptible to instability via NAHR-mediated inversions and 942 genes, 99 of which are on the X chromosome, are predicted to be disrupted secondary to such an inversion! Additionally, IP-LCRs larger than 800-bp with at least 98% sequence identity (DTIP-LCRs) were recently implicated in the formation of complex genomic structures consisting of DUP-TRP/INV-DUP by a replication-based mechanism involving a template switch between such inverted repeats. We identified 1,551 DTIP-LCRs in the human genome that could facilitate DUP-TRP/INV-DUP formation and contribute to genomic complexity. Remarkably, 1,445 disease associated genes are at risk of undergoing copy-number gain as a consequence of susceptibility to formation of a DUP-TRP/INV-DUP complex rearrangement. Our data implicate inverted LCRs as a human genome architectural feature that could potentially be responsible for genomic instability associated with many human disease traits.

486W

DNA from 1D to 3D: Effect of single nucleotide change on DNA conformational structure. *P.W. Yim, C. Liu, C. Woo, Y. Shi, J. Stevens, S. Dong.* Life Technologies, Foster City, CA.

Single Nucleotide Polymorphisms (SNPs) have been used as effective biomarkers for disease association and susceptibility. However, not all SNPs occur in exons, and the effect of a SNP in an intron or other non-coding regions of the DNA is unclear. We hypothesize that a noncoding or intronic SNP alters chromatin organization, which may then lead to changes in gene regulation and expression. In this work, we screened seven cell lines with 12 SNP assays in the PLD1 gene for heterozygosity. Breast cancer cell line MCF-7 exhibited heterozygosity for five of the 12 SNPs. Clustering between each of these SNPs and the rest of the PLD1 gene were analyzed using the TaqMan® 3C Chromosome Conformation Kit. For qPCR, two TaqMan genotyping probes were included in the same reaction and each probe was used to monitor allele-specific interaction. Our results revealed divergent interaction profiles for each allele of the SNPs. Although the interaction to some sites showed no difference, difference in the interaction frequency was observed between the two alleles for many sites. A comparison of the expression of the PLD1 gene showed lower expression for the heterozygous MCF-7 cell line when compared to the homozygous MDA-MB-231 cell line. Such analyses of the changes in the primary DNA sequence and their effects on the three dimensional DNA conformational structure can lead to new insights into the genetic cause of disease and open up new therapeutic approaches.

487F

Analysis of Dinucleotide Mutation Rates and Composition of Human Genome. *G. Zhang[1], D.W. Nebert[1,2].* 1) Human Genetics Division, Cincinnati Children's Hospital Medical Center, Cincinnati, OH; 2) Department of Environmental Health and Center for Environmental Genetics (CEG), University of Cincinnati Medical Center, Cincinnati, OH.

Mutation is the ultimate source of genetic variation; knowledge about the magnitude and mechanism(s) of nucleotide mutation is fundamental to our understanding of evolution of the human genome. It has been known for decades that nucleotide mutation is not random but rather highly dependent on neighboring nucleotides. A well-established example is the high mutation rate of 5-methylcytosine (5mC) to thymine (T) in the methylated CpG dinucleotide and consequent genome-wide underrepresentation of CpG. In this study we systematically investigated mutation rates and genome-wide relative abundance of dinucleotides. Specifically, we estimated dinucleotide mutation rates by two different approaches: 1) comparison between the non-polymorphic human reference sequence and ancestral sequence inferred from non-human primates; and 2) substitution patterns of almost 40 million SNPs from The 1000 Genomes Project. These two methods produced consistent (R^2=0.98) estimates on dinucleotide mutation rates; the observed pattern of dinucleotide substitutions was similar among intergenic, intronic and exonic regions - except for a significantly lower mutation rate of CpG in exonic regions. Detailed examination revealed a transition-to-transversion ratio (Ts/Tv=3.6) and substantially greater transition (24-fold) and transversion (4-fold) rates in CpG than in non-CpG dinucleotides. Also, we found higher rates of transversion between non-complementary nucleotides (A/C, G/T) than transversion between complementary nucleotides (A/T, G/C), and higher mutation rates of purine/pyrimidine or pyrimidine/purine dinucleotides than purine/purine or pyrimidine/pyrimidine dinucleotides. Under a steady-state Markov chain model, we calculated the expected dinucleotide composition, based on estimated dinucleotide mutation rates. The expected frequencies of dinucleotides agreed well with the observed frequencies for intergenic (R^2=0.92) and intronic regions (R^2=0.88) but not for exonic regions (R^2=0.37), which exhibited a much higher-than-expected CpC, GpG, GpC and CpG content. These results demonstrate that, in addition to the hypermutability of CpG dinucleotides, nucleotide mutation rates are specific for different dinucleotides. This phenomenon primarily shapes the observed genome-wide frequencies of dinucleotides - except for exonic regions, in which molecular mechanisms other than mutation appear to be responsible for maintaining the higher-than-expected G/C content.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

488W

The L141R and G153D variations associated with exfoliation glaucoma do not affect enzymatic functions of the lysyl oxidase-like protein 1 (LOXL1. *Y. Kim, S. Kim.* Dept Biochem, Wonkwang Univ Sch Med, Iksan, South Korea.

Lysyl oxidase-like gene 1 (LOXL1) is a copper-dependant amine oxidase that plays an essential role in biogenesis of the extracellular matrix through regulating formation of collagen and elastin fibers in connective tissues. LOXL1 is secreted into the extracellular matrix and become enzymatically active after proteolytic cleavage of the N-terminal propeptide region by BMP-1. Recently, two nonsynonymous SNPs of LOXL1, L141R and G153D that are located in the propeptide domain of LOXL1, have been reported to significantly increase susceptibility to exfoliation glaucoma (XFG). XFG is characterized by progressive accumulation of abnormal microfibrillar deposits in the anterior segments of the eye. To assess the effects of the L141R and G153D variations, we generated 4 different haplotype variants of LOXL1 across L141R and G153D, using oligonucleotide-directed mutagenesis. In amine oxidase assays, the four different haplotype variants of LOXL1 did not show statistically significant differences in amine oxidase activity toward collagen I and elastin. In addition, the L141R and G153D variations did not affect the proteolytic processing of LOXL1 in cultured cells. These results are consistent with recent genetic findings on the reversal of risk alleles of R141L and G153D in different ethnic backgrounds, suggesting that other unknown genetic factors or molecular mechanisms may be more relevant to the development of XFG.

489F

Exploration of HIV susceptibility by exome sequencing. *S. Liu[1], W. Scott[1], D. Dykxhoorn[1,2]*. 1) Hussman Institue of Human Genomics, University of Miami, Miami, FL; 2) Department of Microbiology and Immunology University of Miami Miller School of Medicine.

Human Immunodeficiency Virus (HIV) is a highly lethal lentivirus which is responsible for a tremendous amount of suffering and death in the United States and throughout the world. By attacking the host's immune system, HIV leaves the infected individual susceptible to a wide variety of opportunistic infections. However, population risk of HIV varies. Why and how individuals differ in their susceptibility to HIV remains largely unanswered. Understanding what provides natural protection against HIV is critical in the effort to develop effective strategies to combat HIV. For this purpose, we performed whole exome sequencing on a cohort of highly exposed seronegative (HESN) individuals, long-term nonprogressors (LTNP), and rapid progressors (RP) using the Illumina HiSeq 2000. The resulting data was processed and annotated with BWA, GATK, and SeattleSeq. SNP Variants and indels were filtered and a comparison was performed between different groups. Comparing the HESN individuals with the HIV+ individuals (LTNP and RP)will facilitate the identification of protective factors that inhibit HIV infection, while comparing between LTNP and RP reveals genes that regulate viral progression in host cells. Our initial study shows that rs78424385 and rs56000326 (both on pseudogene POLR2J4 which encoding variant of hRPB11 subunit of RNA polymerase II) occurs at a much higher frequency among LTNPs than they do in the RP group. RNA polymerase II is known to transcribe proviral DNA into RNA which will be packaged into virions. Nucleotide polymorphisms which affect the function of RNA polymerase II are likely to be related to the rate of viral replication. Studying the relationship between HIV susceptibility and gene variants will not only increase our understanding of the mechanism underlying HIV infection and progression but will also help to identify therapeutic targets that could help to prevent viral transmission.

490W

Linkage disequilibrium analysis of polymorphisms of FRAS1 Related Extracellular Matrix 1 (FREM1). *J.F. Tuff[1], N. Kaplonski[1], V. Ly[1], B. Liang[1], S. Tyler[1], F.A. Plummer[1,2], M. Luo[1,2]*. 1) National Microbiology Laboratory, Winnipeg, MB, Canada; 2) University of Manitoba, Medical Microbiology, Winnipeg, Canada.

Objective: A low-resolution genome wide SNPs analysis of the Pumwani Sex Worker (ML) cohort identified an intronic single nucleotide polymorphism (SNP) in FREM1 (rs1552896) to be highly correlated with the HIV-exposed seronegative (HESN) phenotype. To identify putative causal SNPs, we conducted comprehensive gene sequencing, and targeted re-sequencing, of FREM1. Design: FREM1 (176,330bp) was PCR-amplified with 22 overlapping amplicons for 69 patients selected to enrich for rs1552896 genotypes. Amplicons were pooled and sequenced with Genome Sequencer FLX™ (Roche). SNPs showing significant LD with rs1552896 were identified based on a Logarithm of the Odds (LOD) score ≥ 2.0 as determined by Haploview v4.2. Targeted PCR and sequencing was then carried out on a larger sample set (n=643) to confirm the LD of SNPs identified by pyrosequencing. Results: Full-gene pyrosequencing of FREM1 identified 2198 SNPs. Eighty of these SNPs were found to be in LD with rs1552896 (LOD score range: 2.05–54.92). The targeted re-sequencing confirmed LD between rs1552896 and two exonic SNPs: rs2779500 (LOD: 4.99, V439L, 4649bp 5' to rs1552896) and rs1353223 (LOD: 3.30, I499V, 1170bp 5' to rs1552896). The SNP showing the strongest LD with rs1552896 was rs2642404 (LOD: 54.92), an intronic SNP located 4735bp 5' to rs1552896. SNP rs1353223 is a particularly interesting candidate as it is predicted to introduce a novel 5' splice site (Alamut) and to cause a functionally damaging amino acid substitution to FREM1 (SIFT). Conclusions: A combination of full gene sequencing and targeted re-sequencing of FREM1 identified two exonic SNPs and one intronic SNPs in close proximity to, and in LD with, rs1552896. The function of these SNPs will be further investigated.

491F

Exome sequencing of subjects with familial 1q21.1 CNV and variable phenotype. *C. Harvard[1,2], F. Mo[3], F. Tang[1,2], Y. Qiao[1,2], S. Hamilton[4], S. Marles[5], B. McGillivary[4], C. Colins[3], M.E.S. Lewis[4], P. Pavlidis[6], E. Rajcan-Separovic[1,2]*. 1) Pathology & Laboratory Medicine, University of British Columbia, Vancouver, BC, Canada; 2) Child & Family Research Institute, Vancouver, BC, Canada; 3) Vancouver Prostate Centre, Vancouver, BC, Canada; 4) Department of Medical Genetics, Vancouver, BC, Canada; 5) Department of Biochemistry and Medical Genetics, University of Manitoba, Winnipeg, MN, Canada; 6) Centre for High-throughput Biology Michael Smith Laboratories, UBC, Vancouver, BC, Canada.

We performed whole exome sequencing (WES) for two families with a familial 1q21.1 CNV and variable degree of learning difficulty and phenotypic abnormalities in search for underlying genetic causes of the clinical heterogeneity seen in individuals with this CNV. We hypothesize that pathogenic mutation(s) in genes from the 1q21.1 CNV are present on the intact allele of the more symptomatic individuals. Family 1 consists of a trio (proband with 1q21.1 duplication, carrier father, and mother) while family 2 consists of three individuals with a 1q21.1 CNV deletion (proband, mother and grandmother) and two unaffected family members (father and sibling). Methods: Agilent's Sure Select exome capture kit (50 MB) was used to generate libraries sequenced on a Hiseq2000. Reads were mapped using BWA v.0.5.9 and variant calls made using SOAPsnp and SAMtools. Novel SNPs and indels with potential functional consequence, high conservation, and read depth were considered as putatively pathogenic, particularly if they were associated with changes in the expression of the associated gene obtained by whole genome expression analysis for five subjects with the 1q21.1 CNV. Results: None of the candidate pathogenic variants overlapped with the 1q21.1 CNV, ruling out the role of a secondary exonic mutation within this CNV in a more severe phenotype in the two probands. There were more than 60 novel putatively pathogenic mutations occurring de novo in the probands who had a 1q21.1 deletion and 1q21.1 duplication respectively, or shared between the probands and their less affected relatives. The comparison of these mutations with gene expression data identified a mutation within a gene from 1q23.3 with a role in stress response, which was shared between two 1q21.1 dup CNV carriers and associated with abnormal gene expression in both. Selection of additional candidate mutations is ongoing as well as their confirmation. Although the role of copy number changes of genes within 1q21.1 CNV in phenotypic variability can not be excluded, WES offers opportunities to identify other genetic changes that may be involved in modifying the phenotype of 1q21.1 CNV carriers, particularly if they are involved in response to environment.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

492W

Deep resequencing of the NOS2 gene suggests multiple potential mechanisms underlying the association of several complex diseases with a common synonymous SNP. *L. Wang[1,2], J. Santos[1], A. Dressen[1], A. Mehta[1], M. Pericak-Vance[1,2], J. Vance[1,2], W. Scott[1,2].* 1) Hussman Institute for Human Genomics, Univ Miami,Miller Sch Med, Miami, FL; 2) 2. Dr. John T. Macdonald Foundation Department of Human Genetics, University of Miami, Miller School of Medicine.

Three reports have confirmed an association between Parkinson disease (PD) risk and a common synonymous SNP rs1060826 (minor allele frequency (MAF) = 45%) in the NOS2 gene. The SNP has also been associated with other diseases, including age-related macular degeneration (AMD), autism, and tuberculosis. Furthermore, rs1060826 modifies the effect of cigarette smoking on both PD and AMD. We sought to identify the genetic variant(s) underlying the association and gene-smoking interaction at rs1060826. We performed deep sequencing of the 53 kb genomic region of NOS2 including introns, exons and promoter region. To increase our chance of finding the causative variant(s) accounting for the association at rs1060826, we selected individuals who are homozygotes for the risk and protective allele of rs1060826 (N=42 for each group). To overcome the substantial sequence homology between NOS2 and its segmental duplication (SD) regions, we used long-range PCR with primers outside the SD to capture NOS2 genomic region for massively parallel sequencing using the Illumina Hi-Seq2000. More than 300 variants, including small insertions/deletions, were found. Half of these variants were novel. The results to date suggest several potential functional mechanisms that could explain previous observations at rs1060826. First, we discovered that rare and less common functional variants are enriched on the background of rs1060826 G allele, including nonsynoymous variants, 3'UTR variants affecting miRNA binding, and a splicing site variant. The MAF of these variants ranges from 0 to 20% in the Caucasian population surveyed in the NHLBI Exome Sequencing Project. Second, we found a novel 12 bp intronic substitution (GTCCCGTACCTG to GCCTCCTGAG) that is in perfect linkage disequilibrium with rs1060826 and is 300 bp away from the closest exon. Finally, comparison of RNA and DNA sequence revealed that rs1060826 is an active RNA-editing site. To test whether the functional variants appeared in the rs1060826 G allele background collectively account for the association at rs1060826, we are genotyping them in large PD and AMD data sets and conducting individual and burden tests of association. Future studies will determine whether the intronic substitution directs allelic-specific splicing or transcription and whether the RNA editing at rs1060826 affects translation efficiency, as well as how these modifications mediate genetic association and gene-smoking interaction in complex diseases.

493F

Characteristics and analysis of exonic SNPs and indels discovered by low-pass and exome sequencing. *J. Schmidt, J. Gollub, J. Barts, C. Davies, B. Wong, R. Shigeta, G. Hsiao, M. Nitzberg, E. Shell, J. Peete, S. Hsu, A. Kohli, T. Ha, D. Brown, L. Bellon, T. Webster.* Informatics, Affymetrix, Santa Clara, CA.

Exome sequencing of over 12,000 individuals by the Exome Chip Design Consortium and the 1000 Genomes exome sequencing project (including both low coverage and exome sequencing) has revealed a large number of low-frequency SNPs and indels in or near the exome. While SNPs discovered by low-pass sequencing are becoming increasingly reliable, much work remains to be done on indels. Distinguishing true indels from false discoveries remains a challenge. We developed the Affymetrix® Axiom® Exome Genotyping Array to interrogate nearly 300,000 SNPs and 30,000 indels and have analyzed these variants by genotyping 1,249 individuals from four major populations (East Asian, European, African, and Latino) on the array and comparing the results to genotypes published by the 1000 Genomes Project. Genotyping array technologies are mature and have been employed to screen populations for the presence of known SNPs and (more recently) indels, with overall accuracy of >99.5%. They can therefore be employed as an accurate and inexpensive high-throughput analysis and validation tool to be used in conjunction with sequencing. We found excellent concordance to sequencing-derived calls of heterozygote SNP genotypes down to very low minor allele frequency. In the current state of the art, indels on the other hand appear to be prone to false discovery through sequencing. Interestingly, we found that even in the absence of detailed sequencing data, the number of independent sequencing efforts in which an indel was discovered was highly predictive of array validation (and hence indicating a much lower false positive rate), with a much lesser dependence on the frequency of the indel. We also analyzed the distribution of types of SNPs and indels in the various populations. We observed that non-synonymous variants (as compared to synonymous, UTR and intronic variants) have a significantly higher occurrence in variants private to a given population. This observation remains valid even after accounting for the fact that variants present in all four populations generally have a higher allele frequency than those private to just one of the populations. These results demonstrate the value of genotyping arrays for efficient validation of sequencing-derived variant discoveries and subsequent analysis. Genotypes are available from the Affymetrix website as well as the 1000 Genomes ftp site. We thank 1000 Genomes Project and the Exome Chip Design Consortium for access to their data.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

494W

Somatic mosaicism in the blood compartment of a 115-year-old woman. *H. Holstege[1], D. Sie[2], W. Pfeiffer[3], T. Nicholas[4], M.A. Miller[3], T. Harkins[5], C. Lee[5], T. Ross[5], B. Ylstra[2], H. Meijers-Heijboer[1], M. Reinders[6], G. Holstege[7], E. Sistermans[1], S. Levy[4]*. 1) Dept of Clinical Genetics, VU University Medical Center, Amsterdam, Netherlands; 2) Dept of Pathology, VU University Medical Center, Amsterdam, Netherlands; 3) San Diego Supercomputer Center, UCSD, La Jolla, USA; 4) Scripps Translational Science Institute, San Diego, USA; 5) Life Technologies, Beverly, MA, USA; 6) Delft Bioinformatics Lab, Delft University of Technology, Delft, The Netherlands; 7) Center for Uroneurology, University Medical Center Groningen, Groningen, The Netherlands.

The somatic mutation rate in human tissues is currently unknown. Using DNA sequencing technologies, we sought to detect somatic mutations acquired and retained in the blood compartment during a human lifetime. Therefore, we compared the blood and brain genomes (both >60x read depth) of a 115-year-old woman. This individual provides a unique opportunity to study the number, type, and location of somatic mutations exhibited in distinct tissues during a long and healthy life.

In a first analysis, we used a set of sequence filtering procedures to detect SNVs and indels that had a different genotype in blood and brain. We detected ~200 SNVs and ~20 indels that were heterozygous in blood, but not present in brain, consistent with our hypothesis that the higher rate of cell division in blood results in more blood-specific somatic mutations. Further experimental validation of these SNVs and indels by a secondary sequencing technology (>1000 × read depth) in blood, brain, and 9 other tissue types, confirmed 98% of these somatic mutations in blood but not in brain. Notably, lung tissue possessed a low variant allele frequency in these loci, presumably due to vast lymphocyte infiltration in this tissue.

The somatic origin for these variants is supported by their absence in dbSNP, by their heterozygosity, by their occurrence only in blood and not in brain, and by their lower mean variant allele frequency compared with heterozygous SNVs shared between blood and brain. This somatic mosaicism suggests that these loci emanate from a subset of the cells in the blood compartment. Several biological causes for the origin of these somatic mutations are currently being investigated.

Even though there were many heterozygous insertions shared between blood and brain, the putative somatic indels were nearly all deletions. This suggests that deletions are more readily acquired than insertions during DNA replication or DNA repair.

In this first effort to detect somatic mutations we analyzed only non-repetitive genomic regions. Extrapolation of the validated variants to the whole genome implies at least 800 somatic SNVs in the healthy blood compartment, which is almost 30-fold higher than the *de novo* mutation rate between two generations and immeasurably higher than the mutation rate in the healthy brain. Such a variable mutation rate between healthy tissues has implications for future genomic studies.

495F

High-Resolution Melting Analysis of *STAT3* Involved in Autosomal Dominant Hyper IgE Syndrome (HIES) using Bulge-Inducing Primers to Eliminate Polymorphisms and Snapback Primers for Genotyping. *M.T. Seipp[1], N.H. Augustine[2], A. Kumanovics[1,2], H.R. Hill[1,2,3,4], C.T. Wittwer[1,2]*. 1) Advanced Technology Group, ARUP Institute for Clinical and Experimental Pathology, Salt Lake City, UT; 2) Department of Pathology, University of Utah, UT; 3) Department of Pediatrics, University of Utah, UT; 4) Department of Medicine, University of Utah, UT.

Introduction: Autosomal dominant HIES is caused by mutations in *STAT3*, a 7.5 kb gene located on chromosome 17. Mutations can occur in any of the 23 coding exons as well as the associated splice sites. Gene scanning by high resolution (HR) melting analysis can be used to target the critical sequence areas and allow the detection of any variation from the expected DNA sequence. Exons with variant melting curves can be identified and further evaluated by genotyping or sequence analysis. **Methods:** Initial primer design for the 23 exons was done using Idaho Technology LightScanner Primer Design software. When known single nucleotide variants (SNVs) were likely to contribute to melting curve variations, primers were modified to either exclude polymorphisms by inducing a loop bulge, or explicitly genotyped by adding on 5'-genotyping tails to form snapback primers. PCR was performed on the BioRad C1000 followed by high resolution melting analysis on the Idaho Technology LightScanner. After this initial variant scanning, the samples were diluted 1:10 with water so that intra-molecular hairpins were favored for snapback genotyping. A second melting analysis of the diluted products was performed to specifically genotype any variants targeted by the snapback primers. Sequencing was only performed in the rare case of a melting curve variation not identified by snapback genotyping. Sixteen patients with autosomal dominant HIES along with 96 non-HIES DNA samples were scanned for mutations and genotyped for common variants. **Results:** Amplicons with polymorphisms that were eliminated by bulge-inducing primers had melting curves identical to the wild type sequence. When variants were detected by scanning in amplicons with snapback primers, the second melting analysis of the hairpin identified common heterozygous and homozygous variants, drastically reducing the need for sequencing. *STAT3* mutations were found in all affected individuals. The mutations were found in exon 13, exon 20 and exon 21. All other exons were found to match the expected wild type HR melting profiles. All 96 non-HIES samples lacked *STAT3* disease associated mutations. **Conclusion:** All HIES mutations from 16 HIES patients were identified. The combination of HR scanning with bulge-inducing primers and snapback genotyping decreases the sequencing burden of by over 99%. Sequencing was required in <0.5% of all amplicons and the assay could be completed in <4 hours.

496W

PON-P: Integrated Predictor for Pathogenicity of Missense Variants. *M. Vihinen[1,2,3], A. Olatubosun[2,3], J. Väliaho[2,3], J. Härkönen[2,3], J. Thusberg[2,3]*. 1) Department of Experimental Medical Science, Lund University, Lund, Sweden; 2) Institute of Biomedical Technology, University of Tampere, Finland; 3) BioMediTech, Tampere, Finland.

High throughput sequencing data generation demands methods for interpreting the effects of genomic variants. Numerous computational methods have been developed to assess the impact of variations because experimental methods are unable to cope with both the speed and volume of data generation. To harness the strength of currently available predictors, the Pathogenic-or-Not-Pipeline (PON-P) integrates 5 predictors to predict the probability that non-synonymous variations affect protein function and may consequently be disease-related. Random forest methodology-based PON-P shows consistently improved performance in cross-validation tests and on independent test sets, providing ternary classification and statistical reliability estimate of results. PON-P may be used as a first step in screening and prioritizing variants in order to determine deleterious ones for further experimentation. PON-P is freely available at http://bioinf.uta.fi/PON-P.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

497F

Data Submissions to EBI Variation Archives. *I. Lappalainen, V. Kumanduri, J. Almeida-King, L. Skipper, J.D. Spalding, M. Maguire, P. Flicek.* EMBL-EBI Wellcome Trust, Cambridge, United Kingdom.

The European Bioinformatics Institute (EBI) provides public archives that allow scientists to deposit genetic and phenotypic data from the peer reviewed or community research projects. The data, together with study protocols, are archived, indexed and made available to the research community in a consistent format using applicable data standards. The archives issue permanent identifiers for all submitted data allowing explicit referral in publications. These identifiers are also used for cross referencing data stored in different archives for the same project and to integrate archived data into other scientific resources at the EBI such as the Ensembl. The appropriate archive at the EBI depends on the type of informed consent.

For data with full consent, the Database of Genetic Variants archive (DGVa) is repository for all types of genomic structural variants (SV) or copy number variations (CNV) in all species. The resource has a data model designed to capture and describe the abundant complexity of current SV discovery and genotyping experiments. The archived studies are made available in GVF format and the variants can be accessed using Ensembl or Biomart. DGVa and the peer archive dbVar, at NCBI, exchange data monthly. The collaboration with the DGV resource at the University of Toronto will result in a high-quality curated reference set for normal human SV and CNV variation. The April 2012 DGVa release includes more than 90 studies and 10 million SV/CNV variants.

The European Genome-phenome Archive (EGA) provides a permanent archive for all genetic and phenotypic data that may be potentially identifiable. EGA contains data collected from individuals for the purpose of medical or genetic research and whose consent agreements prevents open, public data distribution. The data access decisions are made by a data access commitee and not by the EGA project. Once DAC authorization to the data has been granted, the EGA provides secure tools for downloading and decrypting the authorized files and general support for questions related to the data content. May 2012 data from more than 190 studies and 150 000 samples have been submitted to the EGA and made available to authorized researchers. These studies include array-based genotyping experiments as well as raw DNA sequence data from re-sequencing or transcriptomics projects.

498W

The power of large numbers: frequencies of rare pathogenic mutations in the 23andMe database. *E.R. Chang, D.A. Hinds, N. Eriksson, J.M. Macpherson, C.B. Do, B.T. Naughton.* 23andMe, Inc., Mountain View, CA.

23andMe, with over 150,000 genotyped customers, is in a unique position to accurately estimate the carrier frequencies of rare pathogenic mutations. Accurate, population-specific estimates of mutation frequencies facilitate effective and efficient carrier screening. We present population-specific carrier frequencies for several rare, autosomal recessive conditions. 23andMe's large database, consisting primarily of research participants not selected for any particular disease phenotype, reveals that some mutations that were thought to be associated only with certain populations are also present in additional populations at comparable rates. For instance, the mutation G269S, a late-onset Tay-Sachs mutation, was thought to be found predominantly in individuals of Ashkenazi ancestry. Our data suggest, however, that this mutation occurs in individuals of Eastern European ancestry at a higher frequency than in individuals of Ashkenazi ancestry. Such findings demonstrate the power of a large genetic database in which information about genetic ancestry, self-reported ancestry, and country of residence are all used to estimate population-specific and clinically meaningful statistics.

499F

Finding from CAGI: The Critical Assessment of Genome Interpretation, a community experiment to evaluate phenotype prediction. *S. Brenner[1], S. Repo[1,3], J. Moult[2], C. CAGI participants[2].* 1) Univ California, Berkeley, Berkeley, CA; 2) IBBR, University of Maryland, Rockville, MD; 3) EMBL-EBI, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SD, UK.

The Critical Assessment of Genome Interpretation (CAGI, ˋk-jēˊ) is a community experiment to objectively assess computational methods for predicting the phenotypic impacts of genomic variation. In this assessment, participants are provided genetic variants and make predictions of resulting phenotype. These predictions are evaluated against experimental characterizations by independent assessors. The CAGI experiment culminates with a community workshop and publications to disseminate results, assess our collective ability to make accurate and meaningful phenotypic predictions, and better understand progress in the field. A long-term goal for CAGI is to improve the accuracy of phenotype and disease predictions in clinical settings. The CAGI 2011 experiment consisted of 11 diverse challenges exploring the phenotypic consequences of genomic variation. In two challenges, CAGI predictors applied the state-of-the-art methods to identify the effects of variants in a metabolic enzyme and oncogenes. This revealed the relative strengths of each prediction approach and the necessity of customizing such methods to the individual genes in question; these challenges also offered insight into the appropriate use of such methods in basic and clinical research. CAGI also explored genome-scale data, showing unexpected successes in predicting Crohn's disease from exomes, as well as disappointing failures in using genome and transcriptome data to distinguish discordant monozygotic twins with asthma. Complementary approaches from two groups showed promising results in predicting distinct response of breast cancer cell lines to a panel of drugs. Predictors also made measurable progress in predicting a diversity of phenotypes present in the Personal Genome Project participants, as compared to the CAGI predictions from 2010. CAGI is planned again for 2012 and we welcome participation from the community. Current information will be available at the CAGI website at http://genomeinterpretation.org.

500W

Polymorphic Regulation of Gene Expression by the Histone Demethylase KDM4C. *B.L. Gregory[1], V.G. Cheung[2,3].* 1) VMD-PhD Combined Degree Program, University of Pennsylvania, Philadelphia, PA; 2) Departments of Pediatrics & Genetics, University of Pennsylvania, Philadelphia, PA; 3) Howard Hughes Medical Institute.

Histone demethylases are chromatin modifiers whose biochemical activities are characterized, yet their target genes remain largely unknown. In this study, we focus on KDM4C, a Jumonji demethylase that targets trimethylated H3K9 histone marks associated with transcriptional repression. Using genetic and molecular approaches, we identified and characterized target genes of KDM4C, allowing us to define its role in cell proliferation.

We measured gene expression in B cells from members of 45 large pedigrees, and then treated gene expression levels as quantitative traits in linkage and association analyses. We found extensive individual variation in the expression levels of *KDM4C*. Genetic analyses showed that *KDM4C* expression is regulated in cis by variants near the 3' UTR. These *KDM4C* alleles then influence the expression levels of nearly 100 genes in trans, including *CLTA*, *H3F3B*, *MEF2C*, *RNMT*, *RRP7A*, *TRA2B*, and *ZBTB25*. We followed up the genetic mapping results with functional studies. Using RNA interference, we found that knockdown of *KDM4C* led to a decrease in the expression levels of the target genes, confirming the regulator-target gene relationships. By chromatin immunoprecipitation, we found that KDM4C localized to the promoters of these genes, suggesting that KDM4C regulates the expression of these genes by direct binding. Individuals with higher *KDM4C* expression had more KDM4C promoter enrichment than individuals with lower *KDM4C* expression. By chromatin immunoprecipitation for di-methylated and tri-methylated H3K9, we showed that the demethylase activity of KDM4C is positively correlated with *KDM4C* expression level.

Among the targets of KDM4C are genes that play a key role in cell proliferation, including *RNMT* and *MYC*. To study the effect of *KDM4C* expression level on cellular phenotypes, we measured cell proliferation in individuals with high and low *KDM4C* levels. Growth curve measurements and BrdU assays show that cells from individuals with higher *KDM4C* expression grow faster than cells from individuals with lower *KDM4C* expression. In summary, our results reveal the interplay between DNA sequence variants and chromatin modifications in gene regulation. This is one of the first studies to identify common DNA sequence variants that affect epigenetic regulation of gene expression.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

501F

Mitochondrial Sequence Variation In Twins. *Y. Bouhlal[1], S. Martinez[1], H. Gong[1], K. Dumas[1], J.T.C. Shieh[1,2]*. 1) Division of Medical Genetics, Department of Pediatrics, University of California San Francisco, San Francisco, CA, USA; 2) Institute for Human Genetics, University of California San Francisco, San Francisco, CA, USA.

In applying genome-wide sequencing technologies to aid in disease investigation, it is increasingly important to resolve sequence variation in regions of the genome that may have homologous sequences. The human mitochondrial genome provides an opportunity to test sequencing and interpretation abilities given the potential for heteroplasmy, somatic variation, and homologous nuclear mitochondrial sequences or numts. Identical twins share the same mitochondrial DNA from early life, but whether the mitochondrial sequence remains the same is unclear. We compared adult monozygotic twins using high throughput-sequencing on blood-derived DNA and verified individual variants using fluorescent primer extension. We found 37 homoplasmic variants shared between the twin individuals. All shared variants were verified on pre-library genomic DNA samples. These studies supported highly identical genetic sequence in the absence of confirmed heteroplasmic somatic mutation. We detected a number of potential single nucleotide variants present at very low levels but with high sequence quality scores and read homology to the mitochondrial DNA. Such variants were assessed in pre-enriched mitochondrial DNA and suggest that numts can be effectively differentiated from mtDNA variation. We conclude that twin identity extends to the mitochondrial DNA, and it is critical to differentiate between numts and mtDNA in genome sequencing, particularly since heteroplasmy level can influence disease and phenotypes. Further studies on twin biology will aid in understanding how variation occurs and persists.

502W

Polymorphisms in primer binding site affect genotyping accuracy of *FCGR2B* rs1050501 in Caucasians. *J. M. Reynolds.* Clinical Pharmacogenomics Lab, Pfizer, Inc, Groton, CT.

A single nucleotide polymorphism (SNP) of *FCGR2B* called rs1050501 (c.695T>C, I232T) has been associated with susceptibility to Systemic Lupus Erythematosus (SLE) in Asians in several studies, but inconsistently in Caucasians. Many of the studies have used a forward primer that anneals to a site containing the SNPs rs2793082 and rs2793081. Since minor allele frequencies for rs2793082 and rs2793081 have not been reported in dbSNP or HapMap, they were estimated by testing HapMap01 (Caucasian), HapMap02 (Asian) and HapMap03 (African) gDNA samples. While the minor allele frequencies were estimated to be low in HapMap02 and HapMap03 samples, a minor allele frequency of 25% in the HapMap01 samples suggests the rs2793082 and rs2798301 SNPs are common in Caucasians. To determine if the accuracy of genotyping rs1050501 in Caucasians could be affected by using an amplification primer that anneals to the rs2793082 and rs2793081 region, Caucasian samples (n=84) were genotyped for rs1050501 with two Taqman® assays identical for all conditions, except the forward primer. The assays were designated rs1050501 Polymorphic to indicate the forward primer anneals to the rs2793082 and rs2793081 region, and rs1050501 Non-Polymorphic, to indicate the forward primer does not anneal to the rs2793082 and rs2793081 region. In comparing the two assays, it was found that the results were discordant for 14% of the samples. Six samples (7%) called T/C for rs1050501 in the Non-Polymorphic assay were called C/C in the Polymorphic Assay. All six samples were found to be heterozygous for rs2793082 and rs2793081. It appears that in the Polymorphic Assay there was amplification from only one allele. Six samples (7%) failed completely in the Polymorphic Assay, but were called by the Non-Polymorphic assay. All six samples were found to be homozygous for rs2793082 and rs2793081. In this case, the Polymorphic Assay was unable to amplify from either allele. It is possible that this polymorphic priming site has contributed to the inconsistent association of rs1050501 with SLE in Caucasian studies. Due to the high incidence of rs2793082 and rs2793081 in Caucasians and the evidence that using a primer annealing to this polymorphic region adversely affects rs1050501 genotyping accuracy, it is recommend that researchers discontinue the use of genotyping assays that include a forward primer that anneals to the rs2793082 and rs2793081 sites in Caucasian populations.

503F

Identification of rare variants contributing to pain sensitivity in the normal population: an exome sequencing study. *B. Zhang[5], F.M.K. Williams[1], S. Scollen[2], D. Coa[3], Y. Memari[4], C. Hyde[5], Y. Shi[3], J. Harris[1], I. Harrow[6], A. Malarstig[5], B. Dougherty[6], R. McEwen[6], J.C. Stephens[5], K. Patel[5], C. Menni[1], D. Hodgkiss[1], G. Surdulescu[1], W. He[5], B. Sidders[2], D. Ziemek[7], N. Soranzo[4], X. Jin[3,8], S.B. McMahon[9], S.L. John[5], J. Wang[3], T.D. Spector[1]*. 1) Dept Twin Research and Genetic Epidemiology, King's College London, UK; 2) Pfizer Limited, Neusentis, The Portway Building, Granta Park, Cambridge, UK; 3) BGI-Shenzhen, Shenzhen, China; 4) Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, UK; 5) Pfizer Research Laboratories, Groton, Connecticut, USA; 6) Pfizer Global Research and Development, Sandwich, Kent UK; 7) Computational Sciences CoE, Worldwide R&D, Pfizer Inc, Cambridge, MA, USA; 8) China and School of Bioscience and Biotechnology, South China University of Technology, Guangzhou, China; 9) Wolfson CARD, Kings College London, Guys Hospital, London UK.

Sensitivity to pain varies considerably between individuals. Twin studies have shown pain sensitivity to be heritable. Increased sensitivity to experimental pain is also known to be a risk factor for developing chronic pain; a common and debilitating but poorly understood symptom. To understand genetic mechanisms underlying pain sensitivity objectively we explored in detail the genetic variation in individuals' response to experimental heat pain. First, quantitative sensory testing (QST) to heat pain was performed in 2,500 volunteers from the TwinsUK cohort. Singletons lying in the high and low parts (100 each) of the QST distribution were selected for exome sequencing to a depth of 70X, in a discovery set, and the selection process was repeated to provide a replication set. A combination of collapsing and allele sharing methods was used to test the association between single nucleotide variants (SNVs) identified and pain sensitivity. The function of the SNVs identified was further explored using network analysis. We found the distribution of SNVs detected was different between the pain sensitive and insensitive categories, with the insensitive showing higher frequency of rare variants (p=0.033). After merging datasets, we identified 138 unique genes harbouring a SNV with a 2nd lowest p-value <0.01. First we examined the functional annotations of these 138 genes using the online functional annotation tool DAVID. Nine high level GO terms were nominally significantly enriched in the gene list. We applied causal reasoning to our data, which uses a large curated database of directed regulatory molecular interactions to identify the most plausible upstream regulators of a gene set. Of the 138 genes 86 were present in our database of causal interactions, from which we identified 6 nominally significant regulatory networks. One of the regulatory networks, angiotensin II, was highly significant after Bonferroni correction (p=4.9×10-6). The causal network analysis associated 12 genes identified in the present study with angiotensin II. This pathway has been implicated in animal models and human studies of pain. Our study suggests that the angiotensin pathway may provide fruitful new targets in pain management. The approach of sequencing extreme exome variation in normal individuals has provided important insights into gene networks mediating pain sensitivity in humans and will be applicable to other common complex traits.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

504W

Discovering Retrotransposon Insertion Polymorphisms (RIPs) Associated with Human Phenotypes by TIP-chip. *T.D. Babatz[1,2,3,5], D. Avramopoulos[1,2], J.D. Boeke[1,2,5], K.H. Burns[1,2,3,4,5]*. 1) Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD; 2) Department of Molecular Biology and Genetics, Johns Hopkins University School of Medicine, Baltimore, MD; 3) Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, MD; 4) Department of Oncology, Johns Hopkins University School of Medicine, Baltimore, MD; 5) HiT Center, Johns Hopkins University School of Medicine, Baltimore, MD.

Recent studies demonstrate that retroelements L1 and Alu are significantly more polymorphic between individual human genomes than previously appreciated. Retrotransposon insertion events can have functional impact, but the extent of this is not well understood, and consequences of relatively common polymorphisms are mostly unknown. Despite successes of genome-wide association studies (GWAS) in detecting highly significant and replicated associations of common genetic variants and common disease phenotypes, it is understood that the associated single nucleotide polymorphism (SNP) is typically not the causal variant. Rather, the trait-associated SNP (TAS) is tagging a haplotype on which resides a distinct functional variant conferring the phenotype association. Identifying these variants has been difficult, and most regions associated with disease have been intronic or intergenic, noncoding, and apparently nonfunctional. Our hypothesis is that a subset of GWAS SNPs tag a functional retrotransposon insertion polymorphism (RIP) common in the population conferring disease risk. To identify these variants, we are discovering L1 and Alu RIPs specifically in GWAS SNP intervals using transposon insertion profiling by microarray (TIP-chip). Our custom TIP-chip assays a 160 kb window spanning each of 1349 of the most significantly associated SNPs from 370 published GWAS. We expect to detect RIPs that are annotated in the human reference genome, RIPs reported in the database dbRIP and the 1000 genomes project, as well as RIPs that have not been previously described, all of which will be candidates for further study. Ten CEPH individuals from the HapMap project have been assayed using the targeted TIP-chip. The array includes 20 reference L1s and 32 reference Alu insertions for the purpose of evaluating array performance, all of which have been detected in the first four individuals studied. Of 49 reference L1s spanned by the intervals assayed on the array, 46 have been detected, suggesting robust array performance. We have gone on to PCR genotype candidate RIPs in 90 CEPH individuals and shown high LD with the respective TAS. Imputation of candidate RIP genotypes in larger populations and evaluation of linkage with trait-associated SNPs are ongoing.

505F

Fork Stalling and Template Switching (FoSTeS) mechanism generates a novel rearrangement associated with Incontinentia Pigmenti. *M.I. Conte, S. Raimo, M. Paciolla, A. Pescatore, E. Esposito, M.G. Miano, F. Fusco, M.V. Ursini.* Institute of Genetics and Biophysics ABT, IGB-CNR, Naples, Italy.

Incontinentia Pigmenti (IP, OMIM 308300) is a X-linked dominant neuroectodermal disease, lethal in males, belonging to pathological condition also known as *genomic disorder*. IP is caused by recurrent or non-recurrent DNA rearrangements in the Xq28 *NEMO/IKBKG* locus or by point mutation in the *NEMO* coding region. *NEMO* encodes for the NEMO/IKKgamma protein, which acts as a regulatory subunit of the Inhibitor of the KappaB Kinase (IKK) complex required for the NF-kB (Nuclear Factor kappaB) activation pathway. The *NEMO* gene has a unique genomic organization: in the centromeric direction, it partially overlaps the *G6PD* gene; in the telomeric direction, it is part of a Low Copy Repeat (LCR1), whose duplicated copy, the LCR2, located in opposite direction to gene, contains its non functional copy *pseudoNEMO*, known to be a *reservoir* for mutation causing IP. Moreover, the entire *NEMO* locus is characterized by a high frequency of repeated sequences, micro/macro-homologies and tandem repeats. We previously demonstrated that such genomic architecture enhances the vulnerability of the locus to produce de novo genomic rearrangements through different mechanisms (NAHR, Alu-Alu recombination, Gene conversion). Indeed, abnormal recombination in *NEMO* locus may cause either pathological deletions, which may occur during the parental germline meiosis, or benign variants (MER67Bdup and IKBKGPdel), predisposing to the disease. Here, we describe a novel pathologic genomic rearrangement in IP locus generated by replication-based mechanism, named Fork Stalling and Template Switching (FoSTeS), which caused a complex deletion (36bp) in *NEMO* exon 4 that alters the open reading frame of NEMO protein synthesis producing a premature stop codon (p.Q145X). According to the FoSTeS based mechanism, breakpoint sequence analyses revealed an additional T nucleotide at junction point. In addition, the presence of two micro-homology sequences (CCAG) located at breakpoint sites and one non-B DNA structure (tetraplex) in the deleted region due to stretches of G oligonucleotides, may induce DNA strand lesions and facilitate a FoSTeS event able to generate the complex rearrangement. Taken together these findings indicated that other mechanisms than recombination, based on replication events, might enhance the vulnerability of the IP locus to produce *de novo* pathological IP alleles.

506W

The ENCODE effort combining RNA-seq and RT-PCR-seq allows to catalog thousands of novel lncRNAs. *C. Howald[1], A. Tanzer[2], J. Chrast[1], F. Kokocinski[3], T. Derrien[2], N. Walters[1], J. Manuel Gonzalez[3], A. Frankish[3], B.L. Aken[3], T. Hourlier[3], J. Vogel[3], S. White[3], S.M.J. Searle[3], J. Harrow[3], T. Hubbard[3], R. Guigó[2], A. Reymond[1]*. 1) Center for integrative genomics, University of Lausanne, Lausanne, Switzerland; 2) Centre de Regulacio Genomica, Grup de Recerca en Informatica Biomedica, Barcelona, Spain; 3) Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge.

Within the ENCODE consortium, GENCODE aimed to accurately annotate all protein-coding genes, pseudogenes and non-coding transcribed loci in the human genome through manual curation and computational methods. The continuously evolving GENCODE gene set now contains 53,934 genes of which 20,110 are coding, and 183,086 transcripts (GENCODE freeze version 12). Lower confidence transcribed loci were systematically experimentally evaluated by RT-PCR amplification followed by highly multiplexed sequencing readout, a method we coined RT-PCR-seq. 79% of all assessed junctions are confirmed by this evaluation procedure demonstrating the high quality of the annotation reached by the GENCODE gene set. We further took advantage of the 5 billion sequences deep transcriptome profiling generated by the Illumina “Human Body Map” in 16 human tissues to uncover 5918 novel gene models that do not overlap any loci depicted in GENCODE. They potentially represent new non-coding RNA genes or alternatively unannotated 5' or 3'UTR portions of known genes, as the vast majority of these models was shown to have poor coding potential using comparative genomics and mass spectrometry. We experimentally validated using RT-PCR-seq, 73% of the new HBM models, de facto enriching the complexity of the human genome annotation of non-coding RNA genes by more than 4000 novel genes. Similarly, we uncovered and validated another 3300 transcripts and 1850 novel gene models using the 5.5 billion RNA-seq reads from 15 human cell lines produced by the ENCODE consortium. Our findings demonstrate the effectiveness of unbiased RNA-seq combined with targeted RT-PCR-seq to uncover new genome features. These two technologies were simultaneously similarly paired to unravel expressed pseudogenes by the GENCODE consortium. Our RT-PCR-seq targeted-approach also has the advantage of identifying novel exons of known genes, as we discovered unannotated exons in about 11% of assessed introns. We thus estimate that at least 18% of known loci have yet-unannotated exons. Our work demonstrates that the cataloging of all the genic elements encoded in the human genome will necessitate a coordinated effort between unbiased and targeted approaches, like RNA-seq and RT-PCR-seq, as the latter method reaches exquisite sensitivity. For example, we validate 41% of extremely rarely transcribed exons that are not represented within deep transcriptome profiling.

507F

Evolution of microRNA-3'UTR interactions and their impact on whole transcriptome regulation. *A.Z. Pietrzykowski, E. Kruse, N. Kinstlinger, L. Tejeda, S. Hamza, Y. Wang.* Animal Sciences, Genetics, Rutgers University, New Brunswick, NJ.

MicroRNAs (miRNAs) constitute a novel group of short RNA molecules functioning as major regulators of gene expression. It has been shown that they play a fundamental role in a majority of biological processes including development, aging, pathogenesis of cancer, development of addiction as well as many other processes. Despite an exponential growth of our knowledge in this young field many aspects of their action are still undiscovered. miRNAs decrease protein synthesis by interaction with the 3' untranslated regions (3'UTRs) of mRNA transcripts. Typically, each miRNA can have several binding sites on tens or hundreds of 3'UTRs. This allows miRNAs to have a very robust effect but also can lead to competition between 3'UTRs for miRNA binding. The effect of sequestration of miRNAs by a single 3'UTR on the rest of the transcriptome is unknown. We determined the impact of this regulation in six different species using KCNMA1 gene transcripts as a model. KCNMA1 is a highly conserved gene found in many species from worms to humans. It encodes a pore-forming subunit of a potassium channel of high conductance (BK channel) important for the proper activity of many cell types. Sequences of KCNMA1 3'UTRs were extracted from C. elegans, D. melanogaster, D. rerio, G. Gallus, R. norvegicus and H. sapiens databases using Ensembl and the UCSC Genome browser. Multiple sequence alignment was performed using ClustalW. We observed a persistent elongation and increase in the number of 3'UTRs correlated with the increased complexity of each species. We used miRNA target prediction software to generate a list of miRNAs interacting with the KCNMA1 3'UTRs in each species. Next, we used the same software to identify the other 3'UTR targets affected by those miRNA. Circos plots were made to map the network of 3'UTR-miRNA-target interactions. Surprisingly, we observed that a single 3'UTR could have a very potent effect on the entire transcriptome. In humans it can influence around ninety percent of UTRs. This effect is lessened with decreasing organismal complexity. Our data sheds new light on the evolution of miRNA regulatory networks and is important for understanding the genome-wide effects of miRNA regulation of gene expression.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

508W

Let-7 coordinately down-regulates genes on the amino acid sensing pathway to repress mTOR and activate neuronal autophagy. *A.N. Dubinsky[1], S. Djakovic[1], R. Zahra[1], C. Hsu[1], B.E. Morrison[1], A.R. La Spada[1,2,3].* 1) Pediatrics, University of California, San Diego, La Jolla, CA; 2) Cellular & Molecular Medicine, University of California, San Diego, La Jolla, CA; 3) Neurosciences, University of California, San Diego, La Jolla, CA.

Autophagy is a cellular recycling pathway that removes misfolded proteins and damaged organelles by sequestration into double membrane autophagosomes and lysosome fusion. To date, the genetic and molecular events responsible for regulating the autophagy pathway remain incompletely understood. MicroRNAs (miRNAs) are one class of regulatory molecule capable of modulating gene expression. After developing a method of nutrient deprivation for autophagy induction in primary cortical neurons, we found that Dicer activity is required for this autophagy activation. We therefore performed a combined miRNA - gene expression microarray analysis and identified a set of miRNAs that undergo significant alteration upon nutrient deprivation induced autophagy activation in neurons. Transfection of miRNA mimics into primary neurons resulted in robust activation of autophagy, as determined by GFP-LC3 puncta formation and a shift in LC3-I to LC3-II upon immunoblotting, thereby identifying let-7f as a positive activator of neuronal autophagy. As nutrient deprivation promotes autophagy activation via mTOR inhibition, we analyzed the PI3-kinase - Akt signaling pathway in neurons transfected with let-7f mimics, and noted that although mTOR and its targets are inactive, Akt was phosphorylated, ruling out a role for suppression of the insulin signaling pathway in let-7f-mediated autophagy induction. In silico analysis of our microarray expression data and of published argonaute HITS-CLIP targets of let-7f in the cortex, however, yielded a set of genes (Slc7a5, Map4k3, RragD) on the amino acid sensing pathway, upstream of mTOR, as predicted let-7 targets, which we confirmed by qRT-PCR analysis of let-7f transfected cortical neurons, and by evaluation of let-7 3'-UTR target sites in luciferase reporter assays. As knock-down of these let-7 targets is sufficient to induce autophagy in primary neurons, and injection of let-7f lentiviral vectors into the brains of GFP-LC3 mice is sufficient to induce neuronal autophagy in vivo, our results indicate that let-7f coordinately down-regulates positive activators of the amino acid sensing pathway to repress mTOR and thus promote autophagy activation in the CNS.

509F

Deep sequencing of hypoxic breast cancer reveals deregulation of non-coding RNAs, non ploy-A RNAs and natural anti-sense transcripts. *H. Choudhri[1, 2], C. Camps[1], J. Shoedel[3], S. Oikonomopoulos[1], L. Winchester[4], F. Buffa[4], D. Moralli[1], E. Volpi[1], P. Ratcliffe[3], A. Harris[4], D. Mole[3], J. Ragoussis[1].* 1) Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, Oxfordshire, United Kingdom; 2) Biochemistry Department, Faculty of Science, King Abdulaziz University, Jeddah, Saudi Arabia; 3) Henry Wellcome Building for Molecular Physiology, University of Oxford, Oxford, United Kingdom; 4) Medical Oncology Department, John Radcliffe Hospital, University of Oxford, Oxford, United Kingdom.

Hypoxia is associated with poor prognosis in breast cancer. In this study, we investigated the hypoxia transcription landscape, PolII activity and histone markers (H3K4me3 and H3K4me1) of an ER positive breast cancer cell line [MCF7] grown in hypoxia and normoxia using next generation of ribo- depleted RNA, mRNA, small RNA and ChIP sequencing. We found that about 7.3 percent of the annotated genome is differentially expressed in hypoxic stress, which includes 1138 up regulated and 240 down regulated Refseq transcripts (fold >1.5). We have established a computational pipeline of strand specific ribo-depleted RNA-seq data analysis and revealed deregulation of a range of non coding transcripts including piwiRNA, miRNA, sn/snoRNA, long noncoding RNA (lncRNA) under hypoxia. To define which hypoxic altered genes are dependent on, hypoxia-inducible factor (HIF), we suppressed HIF-alpha subunits in hypoxic MFC-7 cells using siRNA and subjected them to mRNA-seq. We identified 1585 non lnc-RNAs, including natural anti-sense transcripts (NATs), to be differentially expressed, and a substantial number of HIF dependent transcripts (n=59) that were not reported before as hypoxia regulated genes. The differential expression of selected regulatory transcripts, including lncRNA (NEAT1, MALAT1), NATs and non poly-A mRNA (H3J, H3C) were validated in panel of hypoxic breast cancer cell lines. We characterized two significant up regulated NATs (aSPAG4 and aCELSR2) in hypoxia and found that these transcripts are long (>200bp), lack of protein coding frames, localized in nucleus, and HIF dependent. A number (42) of novel un-annotated hypoxia induced transcripts were also detected in hypoxic MCF-7 cells. The presence of active histone markers (H3k4me3 and H3Kme1) at TSS of NATs and novel transcripts validates these findings. In addition, small RNA-seq revealed 107 differentially expressed miRNAs showing both direct regulation through HIF or through post transcriptional mechanisms. Four lnc-RNAs and one non polyA RNA were analysed further using 148 breast tumors and found to be associated with clinicopathological parameters such as triple negative or HER2 positive status. This is the first study to report genome-wide HIF dependent transcription, including non coding RNAs, at single base pair resolution. The combined dataset represents a resource in our efforts to functionally annotate the genome and identify novel pathways to develop therapies for breast cancer.

510W

Integrative annotation and functional characterization of human long non-coding RNAs in bladder cancer. *W. Tang, L. Prokunina-Olsson.* LTG, DCEG, NCI, NIH, Bethesda, MD.

Massively parallel transcriptome sequencing (RNA-seq) has allowed advances in the characterization and quantification of global transcription from simple model organisms to humans. This new strategy helps to explore the complexity of the transcriptome which encompasses a multitude of known and unknown protein-coding and non-coding RNA components. There is growing evidence of importance of long non-coding RNAs (lncRNAs), which are frequently aberrantly expressed in various human cancer types, suggesting potential role of lncRNAs in modulating the tumorigenesis. However, the biological functions of vast majority of these non-coding transcripts remain unknown. In our pilot study, we performed RNA-seq to investigate and characterize the entire transcriptome of three pairs of human muscle-invasive bladder tumors and adjacent normal tissues. We identified hundreds of differentially expressed coding genes between tumor and normal tissues, particularly, genes from pathways responsible for metabolism of xenobiotics, such as aryl hydrocarbon receptor (AHR) and the cytochrome of P450 (CYP1A1) pathways. We also explored expression profiles of 3468 annotated lncRNAs, and observed that some of these lncRNAs were differentially expressed between tumor and normal tissues. Specifically, we found a set of lncRNAs located 25kb downstream of AHR gene highly expressed in normal tissues (adjusted p-value < 0.01) and strongly correlated with AHR expression (r=0.94). AHR is a ligand-activated transcription factor that mediates the toxic actions of polycyclic aromatic and halogenated compounds. These chemicals are well-known carcinogens present in tobacco smoke and implicated in bladder cancer etiology. The activation of AHR pathway could enhance xenobiotic-induced toxicity and trigger carcinogenesis. Recent study also indicated AHR dependency may co-occur with MAP kinase activation in some NRAS-mutant cancer cells, and that elevated AHR may serve as a mechanistic biomarker for enhanced MEK inhibitor sensitivity. Future work will focus on confirming that lncRNAs close to AHR gene could modulate AHR transcription activity to affect the AHR signaling pathway, integrate annotation and functional characterization of these lncRNAs, and further address their effects on MEK inhibitor sensitivity.

511F

MicroRNA-xx regulates vascular smooth muscle cell phenotypes and prevents neointimal formation. *Y. Wang[1], H. Wang[2], Y. Liao[3, 4, 5], P. Tsai[1], K. Chen[1], H. Cheng[1], R. Lin[6, 7], S. Juo[1, 8].* 1) Department of Medical Genetics, Kaohsiung Medical University, Kaohsiung, Taiwan; 2) Department of Biological Sciences, National Sun Yat-Sen University, Taiwan; 3) Graduate Institute of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan; 4) Section of Neurology, Taichung Veterans General Hospital, Taiwan; 5) Department of Neurology, National Yang-Ming University School of Medicine, Taiwan; 6) Department of Neurology, Kaohsiung Medical University, Kaohsiung, Taiwan; 7) Department of Neurology, Kaohsiung Medical University Hospital, Taiwan; 8) Department of Medical Research, Kaohsiung Medical University Hospital, Taiwan.

Aims- Proliferation and migration of vascular smooth muscle cells (VSMCs) can cause atherosclerosis and neointimal formation. MicroRNAs have been shown to regulate cell proliferation and phenotype transformation. We discovered abundant expression of microRNA-xx in VSMCs and conducted a series of studies to identify its function in the cardiovascular system. Methods and Results- MicroRNA-xx expression was initially found to be altered when VSMCs were treated with oxLDL. Using cellular studies, we found that microRNA-xx reduced VSMC proliferation, migration and syntheses of interleukin(IL)-1β, IL-6 and IL-8. Using bioinformatic prediction and experimental studies, we showed that microRNA-xx could repress the expression of Cdc42, CCND1, and FGF1 genes. Using the rat model, we found that microRNA-xx gene carried by adenovirus substantially reduced the neointimal formation in the balloon-injured carotid artery. In situ hybridization confirmed the presence of microRNA-xx in the treated arteries but not in the placebo treated arteries. The experiment of immunohistochemistry for Cdc42 protein showed abundant Cdc42 in the neointima of treated arteries. Conclusions- We identified that microR-xx plays a role in the cardiovascular system by inhibiting VSMC proliferation, migration and proinflammatory biomarkers. MicroRNA-xx may have the potential to reduce the neointimal formation in patients receiving stenting or angioplasty.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

512W

Comparative analysis of strand-specific RNA sequencing approaches. *D. Munafo, P. Liu, C. Sumner, L. Apone, B. Langhorst, E. Yigit, L. Merrill, F. Stewart, E. Dimalanta, T. Davis.* New England Biolabs, Inc., Ipswich, MA.

RNA-Seq is a powerful technique that allows for sensitive digital quantification of transcript levels, enables the detection of noncanonical transcription start sites as well as termination sites, alternative splice isoforms and transcript mutation and edition. Standard “next-generation” RNA-sequencing approaches generally require double-stranded cDNA Synthesis, which erases RNA strand information. Synthesis of a randomly primed double-stranded cDNA followed by addition of adaptors for sequencing leads to the loss of information about which strand was present in the original mRNA template. The polarity of the transcript is important for correct annotation of novel genes, identification of antisense transcripts with potential regulatory roles, and for correct determination of gene expression levels in the presence of antisense transcripts. This work investigates the performance of different strategies for strand-specific (directional) RNA-Seq using next generation sequencing platforms. One is based on direct RNA-adaptor ligation and the second strategy uses the labeling of the second strand cDNA with dUTP and posterior excision. We also examine the effect of different RNA fragmentation methods (divalent cations plus heat versus enzymatic fragmentation) as well as strategies for ribosomal RNA removal. This work provides a comparative analysis (library complexity, continuity of gene coverage, strand specificity and 3'and 5'-end bias analysis) of different approaches for strand-specific RNA-sequencing.

513F

microRNAs modulating the expression of FMR1 mRNA: implications in the pathophysiology of FXTAS. *S. Zongaro[1], R. Hukema[2], L. Davidovic[1], S. D'Antoni[3], M.V. Catania[3], R. Willemsen[2], B. Mari[1], B. Bardoni[1]*. 1) CNRS UMR7275, Valbonne, PACA, France; 2) Erasmus MC Dept. of Clinical Genetics Rotterdam - The Netherlands; 3) Institute of Neurological Sciences National Research Council (CNR) - Catania - Italy.

FMR1 is the gene coding for FMRP (Fragile X Mental Retardation Protein), a RNA-binding protein modulating the expression of a subset of neuronal mRNAs. Silencing of FMRP expression in Fragile X (FXS) patients is due to the hypermethylated expansion of a CGG repeat (>200) localized in the 5' UTR and of the CpG island of *FMR1*, blocking its transcription. The CGG repeat number is polymorphic in the normal population (between 6 et 50 repeats) and it increases up to 55–200 in individuals carrying the premutation. While *FMR1* is silenced in FXS, its expression is elevated (2–8 fold) in individuals carrying the premutation that may develop a late onset neurodegenerative disorder, namely Fragile X Associated Tremor Ataxia (FXTAS). FXTAS patients develop ataxia, parkinsonism, peripheral neuropathy, essential tremor, progressive memory and executive functions deficits (that can be correlated to a reduced hippocampal volume), anxiety and autonomic dysfunction. Moreover, generalized brain atrophy and white matter disease was also observed in FXTAS patients. A KI mouse model for FXTAS exists that recapitulates most of the features of this disorder. The physiopathology of FXTAS remains largely not understood. Nuclear inclusions have been observed throughout the brain of patients as well as in the animal model. These inclusions contain *FMR1* mRNA. Due to the importance of fine tuning regulation of *FMR1* expression during neuronal development, we studied the molecular mechanisms regulating the level of its mRNA. We report here the analysis and the characterization of some microRNAs that target the 3'UTR of *FMR1* mRNA *in vivo*, modulating the expression of this mRNA in post-natal neuronal maturation and, in particular, at the synaptic level. Interestingly, the expression of one of these microRNAs is reduced in brain of FXTAS mice, suggesting an implication in the development/evolution of this neurodegenerative disorder. Detailed characterization of these microRNAs will be presented. Furthermore, the possibility to use them as a treatment for FXTAS and to better characterize the physiopathology of this disorder will be discussed.

514W

DYSFUNCTION OF miRNAs BIOGENESIS IN ASSOCIATION WITH NEURONAL DIFFERENTIATION IN FOCAL CORTICAL DYSPLASIA. *S.H. Avansini[1], F.R. Torres[1], D.B. Dogini[1], F. Rogério[3], A.C. Coan[2], R. Secolin[1], C.S. Rocha[1], A.F. Costa[3], A.L.F. Costa[2], A.C.S Piaza[3], L.A.M. Reis[3], E.P.L Oliveira[2], H. Tedeschi[2], L.S. Queiroz[3], F. Cendes[2], I. Lopes-Cendes[1]*. 1) Medical Genetics, UNICAMP, Campinas, SP., Brazil; 2) Neurology, UNICAMP, Campinas, SP., Brazil; 3) Anatomical Pathology, UNICAMP, Campinas, SP., Brazil.

MicroRNAs (miRNAs) are small noncoding RNAs which regulate post-transcriptional gene expression. Focal cortical dysplasia (FCD) is a malformation of cortical development which affects up to 36% of patients with drug-resistant epilepsy and its etiology is still not well understood. The objective of the present work is to investigate a possible role of miRNA regulation in FCD.We used brain tissue obtained after surgery for the treatment of medically refractory seizures from nine patients with FCD (four patients with FCD type IIa and five patients with FCD type IIb). In addition, we used cortical tissue from autopsy as controls (n=5). Total RNA was isolated with RecoverAllTM kit (Ambion) and RNA integrity was assessed by Agilent RNA Pico Chip Kit and Bio-Analyzer 2100.MiRNA expression profile was assessed by Affymetrix GeneChip platform miRNA array. Background correction, summarization and normalization were performed by RMA function. MiRNA expression was analyzed using RankProd (FDR $p < 0.05$). Results were validated using qPCR to evaluate miRNAs and their target genes. Our array analysis identified 39 miRNAs downregulated when patients and control group were compared and only one upregulated. Decreased expression of three miRNAs was confirmed by q-PCR, and when FCD type IIa and type IIb groups were compared we found that hsa-miR-31 was differentially expressed. Furthermore, we observed that DICER, key protein required for the biogenesis of microRNAs was found differently expressed and involved with neuronal differentiation. Thus our results show strong biological evidence that neuronal differentiation pathways are indeed involved in the pathophysiology of FCD. In addition, we identified a different miRNA expression signature in different FCD histological subtypes.

515F

Genome-wide miRNA signatures of human longevity. *A. ElSharawy[1], A. Keller[2, 5], F. Flachsbart[1], A. Wendschlag[3], G. Jacobs[4], N. Kefer[3], T. Brefort[3], P. Leidinger[5], C. Backes[5], E. Meese[5], S. Schreiber[1,6], P. Rosenstiel[1], A. Franke[1], A. Nebel[1]*. 1) sequencing/mutation detection, Institute of Clinical Molecular Biology, Kiel, Kiel, Germany; 2) Biomarker Discovery Center, Heidelberg, Germany; 3) Febit biomed gmbh, Heidelberg, Germany; 4) Biobank Popgen, Institute of Experimental Medicine, Christian-Albrechts-University, Kiel, Germany; 5) Department of Human Genetics, Saarland University, Medical Department, Homburg, Germany; 6) Clinic for Internal Medicine I, University Hospital of Schleswig-Holstein, Kiel, Germany.

Little is known about the functions of miRNAs in human longevity. Here, we present the first genome-wide miRNA study in long-lived individuals (LLI) who are considered a model for healthy aging. Using a microarray with 863 miRNAs we compared the expression profiles obtained from blood samples of 15 centenarians and nonagenarians (mean age 96.4 years) with those of 55 younger individuals (mean age 45.9 years). Eighty miRNAs showed aging-associated expression changes, with 16 miRNAs being up-regulated and 64 down-regulated in the LLI relative to the younger probands. Seven of the eight selected aging-related biomarkers were technically validated using quantitative RT-PCR, confirming the microarray data. Three of the eight miRNAs were further investigated in independent samples of 15 LLI and 17 younger participants (mean age 101.5 and 36.9 years, respectively). Our screening confirmed previously published miRNAs of human aging, thus reflecting the utility of the applied approach. The hierarchical clustering analysis of the miRNA microarray expression data revealed a distinct separation between the LLI and the younger controls (p-value <10-5). The down-regulated miRNAs appeared as a cluster and were more often reported in the context of diseases than the up-regulated miRNAs. Moreover, many of the differentially regulated miRNAs are known to exhibit contrasting expression patterns in major age-related diseases. Further in silico analyses showed enrichment of potential targets of the down-regulated miRNAs in p53 and other cancer pathways. Altogether, synchronized miRNA-p53 activities could be involved in the prevention of tumorigenesis and the maintenance of genomic integrity during aging. (Aging Cell, 2012).

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

516W

Peripheral Blood Mononuclear Cells MicroRNAs Predict Treatment Outcome of Chronic Hepatitis C. *E. Hsi[1,2], C. Huang[3,4], M. Yu[3,5,6], C. Dai[1,3,5,7], S. Juo[2,9], C. Chou[3], J. Huang[1,3,5,8], W. Chuang[1,3,5]*. 1) Dept Med, Kaohsiung Med UGraduate Institute of Medicine, College of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwaniv, Kaohsiung, Taiwan; 2) Department of Medical Research, Kaohsiung Medical University Hospital, Taiwan; 3) Hepatobiliary Division, Department of Internal Medicine, Kaohsiung Medical University Hospital, Kaohsiung, Taiwan; 4) Department of Occupational Medicine, Kaohsiung Municipal Ta-Tung Hospital, Kaohsiung Medical University Hospital, Kaohsiung Medical University, Kaohsiung, Taiwan; 5) Faculty of Internal Medicine, College of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan; 6) Department of Internal Medicine, Kaohsiung Municipal Hsiao-Kang Hospital, Kaohsiung, Taiwan; 7) Department of Preventive Medicine, Kaohsiung Medical University Hospital, Kaohsiung, Taiwan; 8) Department of Internal Medicine, Kaohsiung Municipal Hsiao-Kang Hospital, Kaohsiung Medical University, Kaohsiung, Taiwan; 9) Department of Medical Genetics, College of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan.

Background & Aims: The application of interferon-based therapy reduces the risk for cirrhosis and improves survival of chronic hepatitis C (CHC) patients particularly in those who achieve a sustained virological response (SVR). Of the host determinants, age, gender, race, insulin resistance, host immune responses and genetic predispositions might significantly affect the achievement of SVR. Several studies have shown that some specific MicroRNAs (miRNAs) sets are associated with the clinicopathological features of liver disease, such as cirrhosis, hepatitis B and C virus infection, hepatocellular carcinoma metastasis, recurrence, and prognosis. Recently, Disease-associated miRNA in peripheral blood mononuclear cells (PBMC) have been described for several diseases. Thus, PBMC-contained miRNAs have a great potential to serve as non-invasive markers. We aim to discover PBMC-contained miRNAs associated with interferon-based therapy response. Methods: We used a two steps method to find miRNA associated with interferon-based therapy response. 1) Performing expression profile of 754 miRNAs in PBMC of 13 typical CHC responders and non-responders by QPCR-based array; 2) Validating candidate miRNAs in an independent 48 CHC patients by QPCR. Results: Four miRNAs with fold change greater or less than four, which were selected for validation. miR-X was expressed significantly lower in non-SVR patients (p=0.0163). Considering IL28B rs8099917 genotype, the most powerful SVR baseline predictor, miR-X was also significantly associated with SVR (OR=2.74, p=0.0055). Combined miR-X and IL28B rs8099917 genotype can improve positive predictive value than IL28B genotype alone. Conclusions: PBMC-contained miRNAs, miR-X has a great potential to serve as non-invasive markers for predicting SVR of CHC patients.

517F

Copy number variation of miRNA in individuals with intellectual disability. *Y. Qiao[1, 2], C. Harvard[1], E. Mercier[3], S. Lewis[2], P. Pavlidis[3], E. Rajcan-Separovic[1]*. 1) Dept Pathology (Cytogenetics), BC Child and Family Res Inst, University of British Columbia, Vancouver, BC V5Z 4H4, Canada; 2) Dept Medical Genetics, BC Child and Family Res Inst, University of British Columbia, Vancouver, BC V6H 3N1, Canada; 3) Centre for High-throught Biology, 177 Michael Smith Laboratiories, University of British Columbia, Vancouver, BC V6T 1Z4, Canada.

MicroRNAs (miRNA) are a family of short non-coding RNAs regulating over 60% of human protein coding genes (miRNA target genes) by post-transcriptional regulation of gene expression. Over 1500 miRNAs have been identified in the human genome and numerous miRNAs are spatially and temporally expressed in the nervous system indicating their potential roles in brain functions. Recent evidence suggests that miRNA copy number change can cause aberrant miRNA expression and/or regulation of their target genes in subjects with intellectual disability (ID) and congenital abnormalities. In order to understand the correlation between miRNAs and copy number variants (CNVs) better, we searched for the prevalence of miRNAs in CNVs detected in a cohort of 206 ID probands and a cohort of 32 cognitively normal subjects. Using the Galaxy intersection web tool (https://main.g2.bx.psu.edu/) in our ID cohort, we found 50 miRNAs in 23 de novo CNVs covering 65 Mb of the genome (on average 3.13 miRNAs/CNV or ~1miRNA/Mb), whereas 52 miRNAs were found in 233 common CNVs covering 118.6 Mb of the genome (on average 0.22 miRNA/CNV, or 0.5 miRNAs/Mb). In the same ID cohort, 14 miRNAs were detected in 58 unique familial CNVs covering 38.1 Mb (0.24 miRNAs/CNV, or 0.5 miRNA/Mb). In cognitively normal controls, a total of 21 miRNAs were detected in the 67 common CNVs covering 35.4 Mb of the genome (on average 0.31 miRNAs/CNV, or 0.6miRNA/Mb). miRNA database searches (http://www.microrna.org/microrna/) showed that the expression of miRNA from de novo CNVs was highest in adult hippocampus, neuroblastomas and heart. One detected miRNA (hsa-mir-124-3), integral to a CNV on chromosome 20q13.33, is the most abundant microRNA expressed in neuronal cells, and is most highly expressed in adult hippocampus and cerebellum. For a comparable number of miRNAs from the common or familial CNVs, expression in the brain was not reported. Our findings revealed a two-fold increase in miRNA content in de novo versus either familial or common CNVs in subjects with ID.Their expression pattern supports a possible role of miRNA copy number change in cognition and/or CNV-mediated developmental delay.

518W

Novel Human Variation in MicroRNAs associated with Disease, Biomarkers, and Drug Metabolism. *R.Afi. Rawlings-Goss, S. Tishkoff.* Genetics, University of Pennsylvania, Philadelphia, USA.

MicroRNAs (miRNA) are evolutionarily conserved regulators of gene expression, which have a significant role in human development and disease. Under and over-expression of specific miRNA have been associated with a wide variety of diseases and disorders including but not limited to leukemia, myopathy, melanoma, lupus, Alzheimer's, neurodegenerative diseases and most forms of cancer. Additionally, there are several clinical trials pursuing the utility of miRNAs as biomarkers, gene therapy agents, and effectors of drug metabolism. As miRNA progress toward therapeutics, population specific variation in miRNA sequences have the potential to answer questions concerning viability of drug targeting, and variability of drug metabolism among different populations. Here we analyze a worldwide panel of participants representing 15 distinct populations from around the world, with a focus on African populations, to assess novel human variation in miRNAs of interest.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

519F

The tale about MIR147B regulate Catalase AEI by one SNP in the 5'-UTR. *Z. Wang[1], B. Wang[1], K. Zhang[1], L. Jin[2], W. Huang[1]*. 1) Department of Genetics, Shanghai-MOST Key Laboratory of Health and Disease Genomics, Chinese National Human Genome Center, Shanghai 201203, People's Republic of China; 2). MOE Key Laboratory of Contemporary Anthropology and Center for Evolutionary Biology, School of Life Sciences and Institutes of Biomedical Sciences, Fudan University, Shanghai, People's Republic of China.

miRNAs affect a variety of signaling pathways, and in traditonal, miRNAs bind through imperfect base pairing to the 3'- UTR of their target mRNAs and interfere with translational output. The allelic expression imbalance (AEI) can be measured in individuals heterozygous for a transcribed polymorphism. Here we show that miRNA miR-147b interacts with the 5' untranslated region of mRNAs encoding Catalase proteins.We have found that three variants (rs769214/ rs7943316/ rs1049982) in the promoter and 5'-UTR region of Catalase are predominant in Chinese Han, and they form two major haplotypes (CATH1: G-A-T; CATH2:A-T-C), and the CATH1 haplotype significantly decreased essential hypertension susceptibility in Chinese Han, and more interesting rs1049982 could serve as a marker for AEI. The expression levels of CATH1 and CATH2 individuals will be shifted in different oxidative stress conditions. miRNA target prediction showed that MIR147B can regulate CAT AEI through the SNP in the 5'- UTR. Luciferase reporter constructs carrying CATH1 (pGL3-CATH1) and CATH2 (pGL3-CATH2) were cotransfected with MIR147B precursor molecule/negative control or MIR147B inhibitor/negative control into Hela and ECV-304 cell lines, separately. Reporter expression was analyzed using the Dual-Luciferase Reporter Assay. We found the MIR147B precursor molecule can down regulated the luciferase expression level both pGL3-CATH1 and pGL3-CATH2. More interestingly, the transcriptional activity of pGL3-CATH2 was significantly reduced than pGL3-CATH1 in all cell lines ($p<0.05$). Compared with the negative control, when we cotransfected with MIR147B precursor molecule pGL3-CATH1 luciferase expression decreased to 44.7%, and pGL3-CATH2 decreased to 27.3%($p=0.019$)in Hela cell lines, and pGL3-CATH1 luciferase expression decreased to 56.1%, and pGL3-CATH2 decreased to 44.3%($p=0.005$). But compared with the negative control, there were no significantly change when we cotransfected with MIR147B inhibitor. We can draw the conclusion that the MIR147B interacts with the CAT 5'-UTR, rs1049982 indeed affects regulation of CAT expression by in vitro. And MIR147B had a higher binding affinity for the “C” genotype than for the “T” genotype. Thus, rs1049982 is a unique SNP that resides in a miRNA-gene regulatory loop. More in-depth understanding of the mechanisms by MIR147B regulate CAT AEI by rs1049982 will give us a charming story.

520W

Inactivation of the *miR-183/96/182* cluster results in syndromic retinal degeneration: The *miR-183/96/182* cluster is required for normal function of the retina and other sensory organs. *S. Xu[1], S. Lumayag[1], C. Haldin[1], C. Cowan[1], B. Kovacs[1], P. Larsen[2], D. Valle[3], P. D. Witmer[3]*. 1) Ophthalmology/Neurological Sci, Rush Univ Med Ctr, Chicago, IL; 2) Biosciences Division, Argonne National Lab, 9700 South Cass Ave, Lemont, IL; 3) McKusick-Nathans Institute of Genetic Medicine; Johns Hopkins University, School of Medicine, Baltimore, MD.

Previously, we identified a highly conserved, intergenic, sensory organ-specific, paralogous miRNA cluster, the *miR-183/96/182* cluster, located within ~4 kb on mouse chromosome 6qA3.3. Developmentally, expression of the *miR-183/96/182* cluster is minimal in embryonic retina, increases after birth and peaks in adult retina where it is specific to photoreceptors and follows a diurnal expression pattern. Although recent reports have demonstrated the importance of *miR-96* in hearing, the functions of this miRNA cluster in retina are still elusive. To explore this question, we first determined the genomic structure of the *miR-183/96/182* cluster gene. We provide evidence that the *miR-183/96/182* cluster gene spans more than 15 kb in the genome and is alternatively spliced with at least two transcripts, one of which has the potential to encode small peptides. The *miR-183/96/182* cluster resides in the intron of the cluster gene. We then characterized a mouse gene-trap embryonic stem cell clone, in which the gene-trap construct was inserted downstream of the first exon of the cluster gene, interrupting the expression of the *miR-183/96/182* cluster of the gene-trapped allele, designated as *miR-183C^{GT}*. Using this gene-trap clone, we generated a mouse model, designated as *miR-183C$^{GT/GT}$*. We show that the *miR-183/96/182* cluster gene is inactivated in homozygous *miR-183C$^{GT/GT}$* mice. These animals have early-onset defects in both scotopic and photopic ERGs and progressive retinal degeneration. Even before overt retinal degeneration, expression profiling revealed global gene-expression changes in the retina of *miR-183C$^{GT/GT}$* mice compared to their littermate controls. Genes known to be important in synaptogenesis and synaptic transmission, as well as phototransduction pathways are significantly enriched in dysregulated genes suggesting that the *miR-183/96/182* cluster is essential for the final functional differentiation, including synaptic connections of the photoreceptors. In addition to retinal defects, the *miR-183C$^{GT/GT}$* mice showed early onset of vestibular dysfunction, including a typical circling behavior and unstable gait. We therefore conclude that complete inactivation of the *miR-183/96/182* cluster results in syndromic retinal degeneration; and the *miR-183/96/182* cluster is required for normal function of the retina and other sensory organs.

521F

Micro-RNA Analysis in Human Ocular Scleral Tissue. *T. Young[1, 2], X. Luo[1, 2], F. Hawthorne[2], S. Feng[2], P. Gonzalez[1]*. 1) Ophthalmology, Duke Univ Eye Ctr, Durham, NC; 2) Center for Human Genetics, Duke University, Durham NC.

MicroRNAs (miRNAs) are expressed in a developmentally regulated and tissue-specific manner. Since miRNA target function studies for the eye will be performed in ocular tissue culture cells, it is important to profile them for miRNA expression and targeting. We profiled affected target genes for selected microRNAs important for regulation of extracellular matrix modeling in an adult human ocular scleral wall fibroblast cell line. Peri-limbal human scleral fibroblast cells (HSFCs) were cultured from postmortem human corneal rings. Mir- 21, mir-29b, mir-200c and mir-let-7F miRNAs were transfected into HSFCs. Whole genome gene expression profiling of transfected HSFCs was performed using the Human HT12-v4.0 Illumina Beadchip. Expression data was generated by using Illumina Genomestudio software. After background subtraction, data were log2 transformed. Sample outliers were removed using principle component analysis and the Hoteling T2 test with 95% confidence interval. The intensity data were normalized by Quantile normalization. The exact Wilcoxon rank sum test was performed to identify differentially expressed genes. The Benjamin and Hochberg false discovery rate procedure controlled at the 0.05 level was used to determine statistical significance. Selected genes with high fold changes were confirmed by quantitative real time PCR (qPCR). Gene expression profiling revealed 1,100 genes which were down- (482) or up-regulated (618)with at least a 2-fold change. qPCR confirmed down-regulated changes of Col4A1, Col3A1, SERPINA3, MFAP4, and TGFβ2 genes. Up-regulation was confirmed for the ANXA10, DUSP6, BMP2, FOS, SEMAF4, LUM, and ERG1 genes. Scleral extracellular matrix regulators showed high regulation from miRNA. This data can potentially be used for scleral wall remodeling alterations in ocular disorders of with scleral wall elasticity.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

522W

Founder mutation in a transposable element leads to the identification of a long noncoding RNA in a progressive infantile encephalopathy. *A. Henrion Caude[1], P. Munier[2], E. Benko[3], S. Hanein[1], E. Westhof[4], N. Boddaert[5], I. Desguerre[6], S. Bandiera[1], M. Girard[1], D. Rodriguez[7], A. Munnich[1], M. Fähling[3], S. Lyonnet[1], F. Cartault[1,2].* 1) Genetics, Inserm U781 - Necker Hospital, PARIS, France; 2) Department of Genetics, Université de La Réunion, Centre Hospitalier Régional de La Réunion, Saint-Denis, LA REUNION, France; 3) Institut für Vegetative Physiologie, Charité-Universitätsmedizin Berlin, BERLIN, Germany; 4) Institut de Biologie Moléculaire et Cellulaire du Centre National de la Recherche Scientifique, Université de Strasbourg, STRASBOURG, France; 5) Département de Radio-Pédiatrie, Necker Hospital, PARIS, France; 6) Département de Neurologie Pédiatrique, Necker Hospital, PARIS, France; 7) Service de Neuropédiatrie, Hôpital Armand Trousseau, PARIS, France.

Familial recurrence of an extreme phenotype of infantile anorexia in a Caucasian isolate in Reunion Island led us to suspect autosomal recessive inheritance of the disease. Here, we report homozygosity mapping of the locus of this syndrome, called Ravine. Complete DNA sequencing of the 400-kb linkage locus revealed a point mutation in a primate-specific retrotransposon that was transcribed as part of a unique long noncoding RNA, which was expressed in the brain. In vitro knockdown of this RNA increased neuronal apoptosis, consistent with the inappropriate dosage of this RNA in vivo and with the phenotype. Moreover, structural analysis of the sequence revealed a small RNA-like hairpin that was consistent with the putative gain of a functional site when mutated. We assessed the global impact of the mutation by transcriptomic analysis of brain samples, and discuss the different mechanisms that could underlie the effect of the mutation. We believe that this work could contribute in the understanding of the significance of subtle variations in retrotransposons.

523F

Human single nucleotide polymorphisms influence non-coding RNA expression. *V. Magadi Gopalaiah[1], H-J. Westra[1], J. Karjalainen[1], D.V. Zhernakova[1], B. Hrdlickova[1], R. Almeida[1], A. Zhernakova[1], M.H. Hofker[3], J. Fu[1,2], S. Withoff[1], L. Franke[1], C. Wijmenga[1].* 1) Department of Genetics, University Medical Center Groningen, University of Groningen, Groningen, Netherlands; 2) Department of Epidemiology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands; 3) Department of Pathology and Medical Biology, Molecular Genetics, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands.

Genome-wide association studies (GWAS) have been successful in identifying approximately 7,000 disease or trait predisposing single nucleotide polymorphisms (SNPs). Surprisingly, only 5% of these SNPs are located within protein coding regions and remaining 95% are located within non-coding regions. Approximately 50% of GWAS associated SNPs affect the expression of nearby protein coding genes. However, most of the human genome is transcribed to produce not only protein coding transcripts but also large numbers of non-coding RNAs. It is unknown whether GWAS associated SNPs can influence the expression levels of non-coding RNAs. Large intergenic non-coding RNAs (lincRNAs) are non-coding transcripts of more than 200 nucleotides long and are recognized as key regulators of many biological processes including gene expression. Recently Cabili et al. have described a catalogue containing more than 8,000 lincRNAs that makes these the largest subclass of the non-coding transcriptome. We hypothesized that GWAS associated SNPs can affect the expression of lincRNAs, thereby providing a novel disease mechanism. To test our hypothesis we performed eQTL mapping on 2,360 human lincRNA-probes using genome-wide gene expression and genotype data of 1,240 peripheral blood samples. We also investigated tissue-dependent eQTLs for lincRNAs using data from four different primary non-blood tissues obtained from 85 individuals. We tested whether SNPs that affect the levels of lincRNA expression are associated with diseases or traits. Finally, we predicted the function(s) of a subset of eQTL-affected lincRNAs using GeneNetwork database that is based on co-expression data, extracted from approximately 80,000 Affymetrix microarray experiments. We obtained 5,106 unique SNPs that affected the expression of 124 lincRNA probes mapping to 117 unique lincRNA genes in peripheral blood. We observed 4, 9 and 117 specific lincRNA eQTLs in liver, saturated adipose tissue and blood, respectively. In total we found 142 eQTL-probes from 5 different tissues. Our results clearly indicate that there is a strong genotype-lincRNA expression correlation in a tissue dependent manner ($P < 1.4E-04$ to 3.5E-198). A considerable number of these eQTLs are GWAS associated SNPs (12% of the eQTL-affected lincRNAs). Our study provides future direction to refine GWAS associated regions and we expect that RNA-sequencing data that is expected to become available in the near future will confirm our observations.

524W

Identifying Source Locations for Alu Transcripts within Human Testes. *S. Linker[1], P. Deininger[2], D. Hedges[1].* 1) HIHG, University of Miami, Miami, FL; 2) Tulane University, New Orleans, LA.

Alu retrotransposons comprise approximately 10% of the human genome. While their transcription has been associated with the cellular stress response, and they have been demonstrated to serve as promoters for functional miRNAs, relatively little is known concerning which specific Alu loci are transcribing and what factors regulate their activity. Previous work in this area has been hindered by technical obstacles, such as cross-hybridization of homologous templates from different sources during PCR and contamination from Pol-II sequences from gene transcripts. These technical hurdles have limited the ability to uniquely identify transcript sources for the purpose of studying their regulation. To circumvent these issues, we have developed a high-throughput method of RNA-Seq preparation that combines a unique oligo extension with plasmid-based amplification. Using this method, we uniquely identified 890 Alu transcription locations from human testes samples. The transcripts originating from these sources were consistent with Pol-III transcripts, as evidenced by depletion of Pol III terminators within the Alu-promoted transcript and enrichment of terminators proximal to the end. As expected, younger AluY was the most transcriptionally active subfamily, exhibiting a four-fold enrichment over expectation based on genomic density. AluY can be further subdivided by age, and within those subclasses, the youngest elements are the most transcriptionally active. Specifically AluYd8, AluYb8, and AluYa5 are enriched 159X, 56X, and 46X over expectations respectively. Further consistent with the transcription of the youngest elements is the fact that transcript source locations exhibit reduced nucleotide divergence from their subfamily consensus sequence. Lastly, the regions surrounding the transcriptionally active Alu were analyzed to extract potential regulatory motifs. We found that transcribing Alu are depleted for several binding motifs. Most significantly, a depletion was observed in the transcription repressor REST, which recruits heterochromatin forming complexes. In summary, we have developed a high throughput next generation sequencing based method for identifying authentic Alu transcripts. We believe our method will greatly facilitate the study of Alu transcription, allowing for a better understanding of their regulation and enabling further examination of their potential functional and pathogenic roles.

525F

The genetic basis of human drug response variation. *M. Kaganovich, D. Spacek, M. Snyder.* Genetics, Stanford University, Stanford, CA.

We have developed a method to assess the functional variation in drug response in lymphoblastoid human cell lines (lcls) in a high-throughput and efficient manner. Four hundred LCLs from various ethnic backgrounds were exposed to chemotherapy agents and their growth response was assayed. The resulting functional variation was associated to genetic differences as measured by the 1000 Genomes Project (1KGP). Our technique and results contribute to nascent inquiries into the functional consequences of human genomic differences found by studies such at the 1KGP and PGP studies.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

526W

Dietary protein regulates gene expression of amino acid-degrading enzymes through PPARα. *A.V. Contreras[1], C. Rangel[1], V. Ortiz[2], G. Aleman[2], B. Palacios[2], E. Tejero[1], N. Torres[2], A.R. Tovar[2]*. 1) Instituto Nacional de Medicina Genómica (INMEGEN), México DF, México; 2) Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, México DF, México.

Peroxisome proliferator activated receptor alpha (PPARα), which is a member of the PPAR family, is a nuclear receptor that functions as a ligand-activated transcription factor in response to dietary signals. The effects of PPARα ligands, such as fatty acids, are have been demonstrated in several tissues, particularly in liver, modulating the expression of enzymes for fatty acid oxidation, but recent evidence suggests that also regulates amino acid oxidation. The amino acid catabolism is regulated by hepatic amino acid-degrading enzymes (AADE), however it is not known whether PPARα regulates the expression of AADE. Thus, the aim of this work is to study whether PPARα controls the gene expression of AADE. With this purpose, we use PPARα null mice and age-matched male C57BL6 mice that were fed diets with 3 different dietary protein concentrations (6, 20 or 50%) in a restricted feeding scheme for 8 days. RNA was extracted from liver to assess the whole genome expression using microarrays (Mouse Gene 1.0 ST Array, Affymetrix), and serum was used to measure glucose, free fatty acids, insulin and glucagon. We found that in WT mice the gene expression of AADE, such as *Histidine ammonia-lyase*, *Glutamic-oxaloacetic transaminase 1*, *Argininosuccinate synthase 1*, *Serine dehydratase* and *Urocanase domain containing 1*, was proportional to the increase in dietary protein concentration. In contrast, in PPARα null mice this effect was absent, even more, for these genes we did not observed significant differential expression between groups fed 6 and 50% of protein. Moreover, some AADE such as *Kynureninase*, *Ornithine aminotransferase* and *Aminocarboxymuconate semialdehyde decarboxylase*, and the *Hepatocyte nuclear factor 4 alpha* (*HNF4α*) showed higher gene expression on PPARα null mice than C57BL6 mice. Network analysis suggests that PPARα regulates the expression of AADE genes through HNF4α interaction. PPARα functions as a metabolic sensor to preserve amino acids by stimulating fatty oxidation. Amino acids may be spared differently in the body depending of the dietary fat, as well as the type of fatty acids mobilized from adipose tissue.

527F

Patterns of allele-specific gene expression in ten human tissues. *K.R. Kukurba[1,2], T. Nance[3], R. Piskol[1], M.H. Tan[1], J.B. Li[1], S.B. Montgomery[1,2]*. 1) Department of Genetics, Stanford University, Stanford, CA; 2) Department of Pathology, Stanford University, Stanford, CA; 3) Department of Mathematics, Stanford University, Stanford, CA.

Allele-specific gene expression (ASE) is the product of gene regulatory variation, chromatin state and nuclear environment. ASE effects have been recently described as important markers of causal regulatory variants and capable of modifying protein-coding loss-of-function mutation. However, studies of ASE have been limited to few cell types leaving many questions regarding the tissue-specificity and overall systemic impact of ASE unanswered. In order to address these questions, we have investigated ASE patterns in ten divergent tissues (cerebellum, colon, frontal lobe, heart, small intestine, liver, lung, pancreas, skeletal muscle and stomach) from a single individual collected post-mortem. ASE was measured using RNA-sequencing (RNA-Seq) data where near-complete ascertainment of coding heterozygous positions was obtained from exome sequencing. Using this data, we report the prevalence and effects of allelic imbalance on tissue differentiation and its evolutionary implications in cell-type identity. We also report essential quality control statistics aimed at eliminating common sources of error when identifying ASE from RNA-Seq and exome sequencing data. We demonstrate the efficacy of filtering by removing potential PCR duplicates, heterozygous loci in intergenic and indel regions, sites with low base quality scores (<30) and with high exome sequencing allele bias. As hypothesized, the shared allelic effects between tissues reflect the embryonic lineage map. In addition, we detected 36 genes across the ten tissues which display monoallelic expression. For two genes in particular, NDN and MAP2K3, monoallelic expression was detected in all 10 tissues. NDN is a known imprinted gene that is only expressed from the paternal allele. MAP2K3 has evidence of allelic expression in humans and mice and allelic methylation in humans. Genes that may selectively escape inactivation in one or more tissues are also observed and may suggest novel escape mechanisms. Furthermore, we also report the tissue- specificity and direction of allelic effects influencing loss-of-function mutation. These findings highlight how regulatory variation influences tissue phenotypes and, ultimately, human phenotypic complexity.

528W

Gene expression variation in the response of human monocyte-derived dendritic cells to LPS and influenza dNS1 stimulation. *M.N. Lee[1,2,3], C. Ye[1], A.C. Villani[1], W. Li[1,3], P.I. Chipendo[4], M.H. Lee[4], P.L. De Jager[1,2,4], C.O. Benoist[2,4], A. Regev[1,5,6], N. Hacohen[1,2,3]*. 1) Broad Institute, Cambridge, MA; 2) Harvard Medical School, Boston, MA; 3) Massachusetts General Hospital, Boston, MA; 4) Brigham and Women's Hospital, Boston, MA; 5) Massachusetts Institute of Technology, Cambridge, MA; 6) Howard Hughes Medical Institute.

Substantial baseline gene expression variation observed in human populations, in both cell lines and primary cells, have been associated with genetic variation in cis and in trans. To fully understand the functional role of genetic variation on gene expression, several recent studies have begun comparing the association patterns observed in stimulated conditions, such as radiation-treated lymphoblastoid cell lines and tuberculosis-infected primary human dendritic cells, to their corresponding baseline profiles. Like other expression quantitative trait loci (eQTL) studies, these response QTL studies could be confounded by environmental and technical factors, which limit their utility. We designed and conducted a serial replication experiment to assess the reproducibility of stimulating primary human monocyte-derived dendritic cells. Primary cells were isolated from a healthy cohort and stimulated with two different innate immune ligands: LPS (receptor: TLR4) and influenza dNS1 (receptors: RIG-I and TLR3). Genome-wide expression profiles before and after stimulation were assayed using the Affymetrix GeneChip Gene 1.0 ST array. Using a linear mixed model comparing the interindividual variation between 30 individuals versus the intraindividual variation within 12 individuals collected several months apart, we found 572, 314 and 240 genes that have significant interindividual variation (FDR less than 0.2) at baseline, after LPS stimulation, and under flu infection, respectively. Taking the difference in expression between the baseline and stimulated matched samples as a quantitative trait, we further identified 601 and 642 genes that have significant interindividual variation (FDR less than 0.2) in the LPS and flu responses. These results point to a remarkably high reproducibility of human innate immune responses, and suggest that there could be a substantial genetic component underlying them. We compared our set of genes to two known eQTL studies. We show statistically significant overlap with eQTLs identified in both studies suggesting that these genes are likely to have a genetic component. We further genotyped our cohort using the Illumina HumanOmni1-Quad and HumanExome at over 1M SNPs. Even with limited sample number, we were able to map expression quantitative traits to cis SNPs at genome-wide significance. A statistically significant number of these cis eQTLs overlap those previously identified, further validating our approach and results.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

529F

Gene expression profiling in a large cohort of Europeans with Sjögren's syndrome reveals viral, immune, and interferon-related pathways. *J.A. Ice[1], H. Li[1,2], J.A. Kelly[1], I. Adrianto[1], S.B. Glenn[1], K.S. Hefner[3], E.S. Vista[4], D.U. Stone[2], R. Gopalakrishnan[5], G.D. Houston[2], D.M. Lewis[2], M.D. Rohrer[5], P. Hughes[5], J.B. Harley[6,7], C.G. Montgomery[1], J. Chodosh[8], J.A. Lessard[9], J-M. Anaya[10], B.M. Segal[11], N.L. Rhodus[5], L. Radfar[2], M.B. Frank[1], R.H. Scofield[1], C.J. Lessard[1,2], K. Moser Sivils[1]*. 1) Arthritis & Clinical Immunology, Oklahoma Medical Research Foundation, Oklahoma City, OK; 2) University of Oklahoma Health Sciences Center, Oklahoma City, OK; 3) Hefner Eye Care and Optical Center, Oklahoma City, OK; 4) University of Santo Tomas Hospital, Manila, Philippines; 5) University of Minnesota, Minneapolis, MN; 6) Cincinnati Children's Hospital Medical Center, Cincinnati, OH; 7) Department of Veterans Affairs Medical Center, Cincinnati OH; 8) Harvard Medical School, Boston, MA; 9) Valley Bone and Joint Clinic, Grand Forks, ND; 10) Universidad del Rosario, Bogotá, Colombia; 11) Hennepin County Medical Center, Minneapolis, MN.

Sjögren's syndrome (SS), characterized by symptoms of dry eyes and mouth, is a progressive autoimmune exocrinopathy present in 0.7–1% of Europeans. To better understand the molecular pathways involved in SS pathophysiology, we performed global gene expression profiling (GEP) using peripheral blood on 48803 probes from the Illumina HumanWG-6 v3.0 BeadChip in 201 SS cases and 76 healthy controls. Analyses were performed in the R Bioconductor suite. After quality control, 20342 probes (in 15607 genes) were quantile normalized and Welch's t-tests, q-values, and fold change (FC) calculated. Differentially expressed (DE) genes were selected by: q<0.05; and FC >1.25 or <0.87. Pathway analysis for DE genes was carried out in Genomatix. In total, 2380 genes were DE (0.05<q<5.5×10E-21). The most significant canonical pathways in DE genes were apoptotic DNA-fragmentation/tissue homeostasis (6/11 genes; p=5.1×10E-4) and regulation of Wnt-mediated beta catenin signaling/target gene transcription (18/32 genes; p=3×10E-3). DE genes were involved in several biological processes, including translational termination (40/93 genes; p=4.3×10E-17), viral transcription and genome expression (49/140 genes; p=3.3×10E-16), innate immune response (93/418 genes; p=1.3×10E-14), and type I interferon-mediated signaling (31/70 genes; p=5.6×10E-14). *OTOF* was the most upregulated gene in SS cases (FC=86.57), while HS.567460 was the most downregulated (FC=0.36). Nearly all DE genes in viral pathways were downregulated with the notable exception of *SP100* (q=7.852×10E-8, FC=1.43), an interferon-inducible protein believed to play a role in transcriptional regulation and the innate immune response following viral infection toward which antibodies are made in primary biliary cirrhosis. Several DE genes overlap with genetic associations identified through genome-wide association studies in SS, including the MHC genes *TRIM38* (q=7.44×10E-12, FC=1.38), *TAP1* (q=5.62×10E-11, FC=1.41), and *TAP2* (q=7.15×10E-16, FC=1.52), in addition to those identified previously by candidate gene approaches, including *IL1RN* (q=4.26×10E-6, FC=1.41), Fas (q=1.23×10E-3, FC=1.27), and *EBF1* (q=1.69×10E-3, FC=1.42). These results highlight alterations in immunologically-relevant pathways in SS, potential to reveal several new candidate expression quantitative trait loci, and provide focus for the development of novel hypotheses for further studies of this complex autoimmune disorder.

530W

A normalization procedure for removal of residual multiplex PCR amplification bias from ultra-deep sequencing of the TCR repertoire. *M.J. Rieder[1], C. Carlson[1,2], A. Sherwood[1], R. Emerson[1], C. Desmaris[1], M. Chung[1], J. Parsons[1], M. LaMadrid-Herrmannsfeldt[1], D. Williamson[1], R. Livingston[1], H. Robins[1,2]*. 1) Adaptive Biotechnologies, Seattle, WA; 2) Fred Hutchinson Cancer Research Center, Seattle, WA.

The human adaptive immune system is a primary defense against the vast world of potential pathogens. To counter this enormous pathogenic diversity, B and T cells constituting the adaptive immune system rearrange surface receptors such that each rearrangement has the potential to code for a different amino acid sequence. For the T cell receptor (TCR), the CDR3 regions are formed by rearrangements of Variable and Joining (VJ) segments for the α chain and Variable, Diversity, and Joining (VDJ) segments for the β chain, in addition to nucleotide insertions/deletions at these junctions, creating the vast diversity in the TCR repertoire. As the adaptive immune system functions in part by clonal expansion, accurately measuring the changes in total abundance of each clone is vital for understanding the dynamics of an adaptive immune response. We developed an optimized multiplex PCR assay for gDNA rearrangements at the TCRβ locus that allows for quantitative assessment of each clonotype using ultra-deep DNA sequencing. This multiplex PCR consisted of 54 V segment and 13 J segment primers designed against all V-J gene segments of TCRβ. To produce an accurate quantitative assay, we need our assay to faithfully output the correct abundance of each clone from an arbitrary sample of T cells. We created a controlled, gold standard input template by synthesizing a single set of 702 oligonucleotides spanning all V-J gene pairs and mixing these at equal ratios. Using this standard as the template input, we first assessed the amplification bias using a PCR reaction containing equimolar ratios of TCRβ V-J primers. We performed additional experiments to assess 1) the independence of each specific V and J primer in the presence of a pools of complimentary primers (i.e. pools of all V or J primers) 2) cross-amplification due to variation in primer concentration (10-fold range) and 3) detection of specific spike-in oligonucleotides. Our results showed that V and J primers showed near complete independence with little interaction. We iterated through different primer rebalancing formulations producing nearly equal amplification across a large dynamic range as assessed by Illumina sequencing. Residual amplification bias following PCR optimization was computationally corrected to further improve the quantitation of the assay. Quantification of clones using this assay allows for accurate and robust tracking of TCRβ repertoire dynamics over time and between tissue types.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

531F

Expression QTL analysis of a gene expression signature which predicts advanced non-alcoholic fatty liver disease. *M. Garrett[1], C. Moylan[1], J. Gibson[1], H. Yang[3], H. Pang[3], A. Dellinger[3], A. Suzuki[1], H. Tillmann[1], C. Guy[1], M. Abdelmalek[1], S. Murphy[2], A. Diehl[1], M. Hauser[1], A. Ashley-Koch[1].* 1) Department of Medicine, Duke University Medical Center, Durham, NC; 2) Department of Obstetrics and Gynecology, Duke University Medical Center, Durham, NC; 3) Department of Biostatistics and Bioinformatics, Duke University Medical Center, Durham, NC.

Non-alcoholic fatty liver disease (NAFLD) is associated with obesity and the metabolic syndrome and affects over one third of American adults. A subset of NAFLD patients develop advanced disease associated with liver injury, scar tissue (fibrosis), and cirrhosis, increasing risk for liver-related mortality, including primary liver cancer. Several genes have been associated with NAFLD, but determination of their function in the disease process has been difficult. We hypothesized that expression quantitative trait loci (eQTL) analysis would improve the identification of causally related genetic variants by integrating SNPs with gene expression. Patients with biopsy-proven NAFLD were selected based on fibrosis stage (mild NAFLD: stage F0-F1 vs. advanced NAFLD: stage F3-F4). Genomic data, including genotype (Cardiometabolic SNP array, Illumina), methylation (Human Methylation450 BeadChip, Illumina) and expression (Human U133 Plus 2.0 GeneChip, Affymetrix) were available for 69 patients. We identified a novel 64-gene expression signature differentiating advanced and mild NAFLD, confirmed in a validation data set. Using ANOVA, cis-eQTL analysis was conducted with transcript levels from this signature and SNPs within 1 MB of the expression probe midpoint, while controlling for gender, age at biopsy, diabetes and BMI. SNPs falling directly within expression probes were excluded. Empirical p-values were generated from permutation testing. 28 SNPs significantly predicted transcript levels ($p<0.001$), including those for CD24P2 (3 independent SNPs), COL1A2 (1 SNP), CYBRD1 (1 SNP) and ID4 (23 SNPs in 3 haplotype blocks). CD24P2 is a pseudogene which does not appear to produce a functional protein. COL1A2 encodes the pro-alpha2 chain of type 1 collagen, a fibrilar collagen found in most connective tissues. CYBRD1 is a ferric-chelate reductase present at the brush border of duodenal enterocytes. All 23 SNPs predicting ID4 expression lie distally within or near the CDKAL1 gene, previously associated with type 2 diabetes. Both ID4 and CKDAL1 are involved in regulation of cell cycle transitions. Interestingly, COL1A2, CDKAL1 and ID4 also exhibit reduced methylation and increased gene expression in advanced versus mild NAFLD patients. The eQTLs that we identified provide important insight into pathologic processes contributing to advanced NAFLD and are significantly enriched for tissue repair, insulin resistance and cell fate decisions.

532W

Bead-based method for capturing specific microRNA in human urine: A non invasive way of detecting clinical conditions. *M. Rath, Tom. Xu, Chris. Trinh, Kathy. Tran, Junko. Stevens, Shi. Amy, G. Marie-Pierre.* Genetic Analysis Integration System R&D, Life Technologies, Foster city, CA.

Research shows that MicroRNA is much more stable and intact than mRNA in various extracellular body fluids such as plasma and urine. Due to this stability, microRNAs in urine samples can be a great promise as a non-invasive way to discover important new biomarkers for a wide range of diseases and biological processes. However, getting microRNA profiles from such samples can be challenging. Life Technology has developed a highly efficient and simplified bead based strategy (TaqMan® miRNA ABC Purification Kit; Cat# 4473087 and 4473088) for rapid purification of microRNA (miRNA) from small inputs of all human sample types including body fluids such as Urine, saliva, cerebro spinal Fluid, FFPE, solid tissues, and cell cultures sample type. In this study, we measured the presence of microRNAs in human urine samples using TaqMan® Array Human MicroRNA A & B cards. A total of 150miRNAs were detected using a Cq cutoff of 35 for TaqMan® Human MicroRNA Array A v3.0 and 79 miRNAs were detected using TaqMan® Human MicroRNA Array B v2.0 in Urine samples. In addition, our profiling data from two pregnant donors indicated gestational age-dependent changes in the expression for 13 miRNAs from TaqMan® Array Human MicroRNA panel A. The validation of these miRNAs with RT-qPCR was found in accordance with the screening result. In this study we have also investigated the expression of miR143 and miR145 in urine samples of patient with balloon injured arteries in comparison to normal artery. Scientific data published by Dr. Chunxiang Zhang showed down regulation of miR-145 in balloon- injured arteries and demonstrated that miR-145 is a critical modulator of vascular smooth muscle cell (VSMC) phenotype.Our result confirmed significant down regulation of these two microRNAs in urine samples of patient with balloon injured arteries in comparison to normal artery. Our method demonstrates the successful purification of specific microRNAs from urine samples without using hazardous organic extractions such as phenol and chloroform. Our result indicated an excellent run to run reproducibility. TaqMan® miRNA ABC Purification Kit; Cat# 4473087 and 4473088 Kit can be used to profile miRNA expression and characterize variances in miRNA species by various downstream applications such as Real Time PCR and miRNA sequence analysis. Profiling urinary microRNAs offers a non-invasive approach and therefore may hold the promise for the development of valuable clinical tools.

533F

Experimental Depletion and Regeneration of Human Mitochondrial DNA to Investigate Its Role in Nuclear Gene Regulation. *S. Kumar, C. Bellis, M.P. Johnson, H.H.H. Goring, T.D. Dyer, J. Blangero, J.E. Curran.* Genetics, Texas Biomedical Research Institute, San Antonio, TX.

Research on mitochondria has evolved from bioenergetics to biogenesis, the genetic functions of mitochondrial DNA (mtDNA), and diseases associated with mitochondrial dysfunction. Although these areas continue to be investigated vigorously, a new era in mitochondrial research has emerged that concerns the role of mitochondria in intracellular signaling - a process that is likely to have far-reaching implications in development, aging, disease, and environmental adaptation. Retrograde regulation is the general term for mitochondrial signaling, and is broadly defined as cellular responses to changes in the functional state of mitochondria. Previously, using complete mitochondrial genome sequences and genome-wide transcriptional profiles of the San Antonio Family Heart Study (SAFHS) individuals we identified approximately 350 transcripts that are highly likely influenced by mitochondrial sequence variation. Among these were several expected cis-effects on mitochondrial transcripts such as *MTCO1* (mitochondrial genome-wide $p = 4\times10^{-34}$) and many dramatic novel trans-effects on the nuclear genes *CENPN* ($p = 1\times10^{-70}$), *PRPH* ($p = 2\times10^{-64}$), *CBX1* ($p = 1\times10^{-61}$), and *TRIM65* ($p = 2\times10^{-43}$) among others. To further investigate the cellular responses to the changes in the functional state of mitochondria, we generated 20 different functional states of mitochondria by depleting and regenerating mitochondrial DNA (mtDNA) experimentally in two lymphoblastoid cell lines from our SAFHS individuals. Real-time quantitative PCR analysis of mtDNA content in these 20 stages of generated cells revealed that their mtDNA content varies from 0.01 to 160 per cent of the original cells content, depending on their functional state as defined by the stage of mtDNA depletion/regeneration. Thus, using this method we can experimentally model the functional state of mitochondria and its influence on nuclear genome. We have performed whole genome gene expression analysis of these 20 stages of generated cells in two cell lines to confirm our original genetically driven findings and to identify novel nuclear genes influenced by the various experimentally generated functional states of mitochondria, to better understand retrograde regulation and inter-individual variations in response. Our results suggest that this experimental approach represents a valuable tool for detecting the impact of mitochondrial genes on nuclear gene regulation.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

534W

Primary Mitochondrial Diseases Dysregulate the Transcriptional and Post-transcriptional Regulatory Systems in a Tissue-specific Pattern. *Z. Zhang[1], M. Tsukikawa[1], E. Polyak[1], J. Ostrovsky[1], C. Clarke[1], E. Place[1], E. Rappaport[1], G. Reiner[2], M. Yudkoff[1,3], R. Naviaux[2], R. Haas[2], J. Baur[3], M. Falk[1,3]*. 1) The Children's Hospital of Philadelphia, Philadelphia, PA; 2) University of California San Diego, San Diego, CA 92103; 3) University of Pennsylvania Perelman School of Medicine, Philadelphia, PA 19104.

Primary mitochondrial respiratory chain (RC) diseases are an increasingly recognized but poorly understood class of genetic disease that impair basic cellular energy metabolism. Due to the highly heterogeneous etiology, severity, and range of multi-systemic manifestations that typify RC disorders, therapies are generally empirically based and focused on supplementing residual mitochondrial capacity. **METHODS**: To identify a common cellular response to primary RC that might improve mechanistic understanding and therapeutic opportunities for RC disease sequelae, we systematically analyzed collective transcriptomic profiles from skeletal muscle biopsies and fibroblast cell lines (FCLs) from a diverse sample of human mitochondrial disease subjects relative to controls. **RESULTS**: Our data showed that individuals with diverse causes of primary RC dysfunction share a distinctive transcriptomic pattern from that of controls. Interestingly, global changes of patient transcriptomes have the opposite direction in muscle biopsy and FCLs, suggesting that different cell types have unique responses to a deficiency of energy metabolism. Exon-level investigations identified a large number of alternative 5'-UTRs, 3'-UTRs, and transcript isoforms in the RC disease group. A particularly strong association existed between differential 3'-UTR levels and the presence of AU-rich elements (AREs), which are protein-binding motifs within 3'-UTRs that regulate mRNA stability and translation efficiency. Genes with AREs were more likely to have 3'UTRs elevated in patient muscle and reduced in FCLs. The ribosomal protein (RP) genes also illustrated a strikingly distinctive pattern in RC disease. Patient muscles had higher cytosolic RP transcription and lower mitochondrial RP transcription, while FCLs showed the opposite trend. Our data collectively suggest that an mTOR-centered regulatory network plays an essential role in mediating the cellular response to mitochondrial diseases. Indeed, the most significantly up-regulated gene in patient muscle, *RHEB*, is a direct regulator of mTORC1 function. **CONCLUSIONS**: Our systems biology approach strongly demonstrates that transcriptional, post-transcriptional and translational dysregulation occurs in primary mitochondrial disease in a tissue-specific fashion. Recognizing the central role of nutrient sensing signaling pathways in the cellular response to mitochondrial RC disease may offer novel therapeutic opportunities.

535F

Genetic Variation in Expression Response to Endoplasmic Reticulum Stress in Human B-cells. *W. Bernal[1,2], M. Morley[2], V. Cheung[3]*. 1) The Children's Hospital of Philadelphia, Philadelphia, PA; 2) University of Pennsylvania, Philadelphia, PA; 3) Howard Hughes Medical Institute, Chevy Chase, MD.

Endoplasmic recticulum (ER) stress is caused by excessive demands on the protein-processing capacity of the ER; inefficiencies in response to ER stress lead to various human diseases. In this project, to study the role of ER stress in disease susceptibility, we identified the genes and pathways involved in ER stress response in human cells, assessed individual variation in these pathways, and mapped DNA sequence variants that influence ER stress response. We exposed cultured human B-cells from 131 normal individuals to tunicamycin, a chemical inducer of ER stress, for 8 hours. Then, using microarrays, we measured expression levels of genes and identified 1,523 ER-stress responsive genes whose expression levels changed at least 1.5-fold on average among 131 individuals. The results also revealed extensive individual differences in ER stress response. For example, across our subjects, the expression responses of canonical ER stress sensors IRE1 and PERK vary by 4-fold and 3-fold, respectively. Similarly, the expression responses of *XBP1* and *DDIT3* (CHOP), which encode key ER stress transcription factors, vary by 5-fold and 9-fold. These results suggest that variation exists not only in sensing of ER stress but also in the extent of ER stress response. To gain insight into pathways that underlie individual differences in expression response to ER stress, we constructed a coexpression network using pairwise correlations. *KDELR2*, which encodes a receptor that retrieves ER proteins from the Golgi, is the most highly connected gene in the network. The *KDELR2* subnetwork is enriched for ER-associated degradation (ERAD) and rRNA transcription, which suggests a role for *KDELR2* in ER stress response beyond ER protein retention.

To understand the basis of individual differences in ER stress response, we carried out a genetic mapping study in B-cells from members of 15 large families. We measured expression responses to ER stress and treated them as quantitative phenotypes in genetic linkage and association studies. We focused our analysis on 778 ER stress-responsive genes and identified DNA sequence variants that regulate the response of 497 of the genes, including *KDELR2*. Of the 497 regulatory relationships identified, 349 were informative when combined with the coexpression network. In this presentation, I will discuss our findings with respect to individual variation in response to ER stress and its underlying genetic regulation.

536W

Dynamic network connectivity mapping of T-cell activation in Caucasians and African Americans identifies a common set of asthma-associated hub genes. *J. Chu[1], W. Qiu[1], V. Carey[1], K. Barnes[2], C. Ober[2], F. Martinez[2], R. Lemanske[2], A. Liu[2], S. London[2], S. Weiss[1], B. Raby[1]*. 1) Channing Lab, Brigham & Women's Hosp, Boston, MA; 2) on behalf of the Asthma BRIDGE consortium.

CD4+ T-cell activation represents a sentinel event in the expression of allergic and asthmatic responses. Though individual molecular components of this process have been characterized, more holistic, system-wide descriptions using transcriptome network mapping in large populations has not been performed. We set out to characterize and compare the gene networks in resting and stimulated human CD4+ peripheral blood lymphocytes, and define those focal modules most radically altered upon activation. A subset of subjects participating in the Asthma BRIDGE initiative from six clinical centers provided whole blood samples, from which PBMCs were isolated and plated in split samples. 25ul of PHA (@5ug/ml) was added to 50% of wells (stimulated samples). At 24 hours, cells were harvested and CD4+ cells and RNA were isolated. Expression profiling was performed using the Illumina Human HT12 array. Following data QC and preprocessing procedures, we used differential connectivity mapping procedures (Chu et al. BMC Systems Biology 2011) and compared the transcriptome networks of the stimulated and unstimulated samples to define regulatory modules that undergo restructuring during T-cell activation. Expression data from 98 non-Hispanic whites and 157 African Americans were available for analysis. We noted widespread differences in the topographical landscape of the CD4+ lymphocyte transcriptome following stimulation, identifying 216 gene hubs that each demonstrated differential connectivity to no fewer than 50 genes. 86 genes were identified in both the white and African American samples, representing a reproducible set of regulatory modules with altered topography (i.e. altered gene-gene co-expression patterns) upon T-cell activation. These 86 genes are enriched for genes involved in antigen processing and presentation (p=3.0E-04) and those with genetic variants associated with immune-related diseases. Among the most differentially connected genes were UTS2, GSTM1, HLA-C, ARG1 and HLA-DRB1, all of which have been implicated in the genetics of asthma. In conclusion, we have described the network topography underlying T-cell activation, and have identified a reproducible set of 86 gene hubs with substantially altered connectivity upon stimulation. As many of the genes on this list are known asthma genes, the remainder should be prioritized for further evaluation as therapeutic targets in the treatment of allergic and autoimmune disease.

537F

Immunity genes and pathways are associated with human microbiome composition at host body sites. *R. Blekhman[1,2], J. Goodrich[1,3], K. Huang[4], R. Bukowski[5], Q. Sun[5], R.E. Ley[1,3], D. Gevers[4], A.G. Clark[1]*. 1) Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY; 2) Department of Neurology and Neuroscience, Weill Cornell Medical College, New York, NY; 3) Department of Microbiology, Cornell University, Ithaca, NY; 4) Genome Sequencing and Analysis Program, The Broad Institute of MIT and Harvard, Cambridge, MA; 5) Computational Biology Service Unit, Cornell University, Ithaca, NY.

Recent advancements in high-throughput sequencing technologies have unveiled wide variability in the communities of microbial species that inhabit the human body. Shifts in the species composition of the human microbiome have been associated with multiple chronic conditions, including diabetes, inflammatory bowel disease, and obesity. Although the human microbiome is influenced by environmental factors, a strong host genetic influence is also expected in the interaction between humans and bacteria through the immune system and metabolic pathways. However, to date, there are no comprehensive studies examining the interactions between human genetic variation and the composition of this microbial ecosystem on a genome- and microbiome-wide scale. Here, we report a genome-wide association study aimed at identifying human genomic loci correlated with levels of different bacteria in various body sites. To do so, we used shotgun metagenomic sequencing data from the Human Microbiome Project, and extracted non-bacterial sequence reads that were labeled as “host contamination.” By mapping these reads to the human genome, we achieved a 10x mean coverage and identified 4 million SNPs in 93 human individuals for which extensive metagenomic data is also available. Despite the small sample size, a genome-wide association approach successfully identified a number of loci in the human genome that show significant associations with microbiome composition at 15 host body sites. These loci are enriched with genes and pathways involved in host immune response, which likely play a role in promoting inflammation in the host as a response to bacterial infection. In addition, we show that many of these pathways are also enriched in a replication study examining gut microbial composition in a separate population. Our results highlight the role of common genetic variation in controlling the composition of the human microbiome and the importance of host immunity in determining bacteria levels across the body. Furthermore, the associations discovered here might be a starting point toward understanding the complex interaction between host genetic variation and the microbiome in susceptibility to common autoimmune diseases.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

538W

Ribosome footprints on Williams syndrome region genes in the mouse brain. *R. Weiss[1], M. Howard[1], L. Dai[2], J. Korenberg[2]*. 1) Human Genetics, Univ Utah, Salt Lake City, UT; 2) Center for Integrated Neuroscience and Human Behavior, Brain Institute, Univ Utah, Salt Lake City, UT.

The advent of quantitative methods to measure global gene expression levels, most recently RNA sequencing, has greatly increased our ability to create a global picture of cellular processes. However, until recently no comparable technology existed to examine genome-wide protein synthesis (translational) activity. Translation is also a level at which a general response to distress caused by genomic imbalance may be regulated to maintain cellular homeostasis, and may be a novel point for therapeutic intervention. Williams syndrome (WS) is a CNV disease caused by a *de novo* hemizygous deletion of 25–28 genes on Chr7q11.23 and is characterized by a distinct pattern of cognitive effects. Global transcriptome profiling in WS has shown a *cis* effect at the transcript level in the deleted region, as well as transcript-level perturbations in a proximal gene network comprised of physically interacting gene products and their associated pathways. But since perturbations in protein expression can arise due to combined changes in mRNA abundance and translational activity on individual mRNAs, we approached the problem by applying the novel technique of ribosome profiling in a multi-tissue comparison of adult tissues from C57BL/6J mice and assessed the transcriptional and translational state of genes in skeletal muscle, heart, liver and brain. Ribosome-mRNA complexes were isolated from snap frozen adult mouse tissues, and after nuclease (RNase) digestion, the ~30 nt ribosome-protected mRNA fragment "footprints" were characterized by deep sequencing using an Illumina sequencer. We observed regions of active translation in WS transcripts that were characterized by dense footprints in CDS regions, but also specific footprints in a subset of 5' UTRs. The location of ribosome footprints in the 5' UTR of Cyln2 were conserved across tissues and located at an AUG codon initiating the translation of short upstream open reading frame prior to the main AUG start codon. This pattern is characteristic of other genes and may be related to the translation efficiency of the main CDS. These results suggest an additional level of gene expression control in the WS region may be susceptible to alteration by copy number changes. Ribosome profiling may have wider utility for interpreting how CNVs alter the abundance of their encoded protein(s) and their secondary effects on the abundance of proteins expressed from other regions of the genome.

539F

Regulatory mutations in the 5'UTR of *NMNAT1*, encoding the nuclear isoform of nicotinamide nucleotide adenylyltransferase 1, cause Leber Congenital Amaurosis. *F. Coppieters[1], A. Baert[1], C. Van Cauwenbergh[1], M. Bauwens[1], S. De Jaegere[1], T. de Ravel[2], F. Meire[3], N. Abdelmoula Bouayed[4], L. Florentin-Arar[5], B.P. Leroy[1,6], E. De Baere[1]*. 1) Center for Medical Genetics Ghent, Ghent University, Ghent, Belgium; 2) Center for Human Genetics, Leuven University Hospitals, Leuven, Belgium; 3) Hôpital Des Enfants Reine Fabiola, Brussels, Belgium; 4) Department of Histology, Medical University of Sfax, Sfax, Tunis; 5) Molecular biology and cytogenetics, Lito Maternity Hospital, Athens, Greece; 6) Department of Ophthalmology, Ghent University Hospital, Ghent, Belgium.

Leber Congenital Amaurosis (LCA) is the earliest and most severe inherited retinal dystrophy (RD). Recently, coding mutations in *NMNAT1* uncovered this gene as the causal disease gene for the LCA9 locus (Pierce *et al.*, ARVO Meeting 2012), mapped a decade ago (Keen *et al.*, 2003). *NMNAT1* encodes the nuclear isoform of the nicotinamide nucleotide adenylyltransferase 1. The goal of this study was twofold. First, we aimed to identify the causal genetic defect in an LCA9-linked consanguineous Sub-Saharan African family. Second, we wanted to determine the contribution of *NMNAT1* mutations in a pre-screened LCA population.

Starting point of this study was a consanguineous family with 3 affected and 1 unaffected sibling in which genomewide IBD mapping identified an IBD region of 14 Mb encompassing LCA9. Sanger sequencing of the *NMNAT1* transcription unit identified a homozygous 5'UTR variant, c.-70A>T, which segregated with disease in the family and was absent in 265 healthy controls. This variant was predicted to alter secondary structure formation of the 5'UTR. Sequencing of c.-70A>T on cDNA of leukocytes revealed loss of heterozygosity in the parents and the healthy sib, suggesting *NMNAT1* mRNA degradation. Subsequent sequencing of *NMNAT1* in 76 unrelated probands with LCA or early-onset RD revealed mutations in 5 additional probands. Of note, a second 5'UTR variant, c.-69C>T, was found in homozygous state in a Moroccan LCA patient. Interestingly, it is located 1 nucleotide downstream of c.-70A>T and predicted to cause similar aberrant 5'UTR folding. A phenotypic evaluation of both families revealed LCA with evolutive macular involvement. In addition, two probands were compound heterozygous for the common missense variant p.Glu257Lys and a novel missense mutation predicted to affect protein function (p.Met69Val and p.Tyr181Cys). Two other probands were found to carry a single heterozygous missense variant (p.Arg188Trp and p.Asn18Ser). Resequencing of the entire genomic region is currently ongoing to exclude a deep intronic mutation.

In conclusion, this study sustained the role of coding *NMNAT1* mutations in LCA. Moreover, the identification of two neighboring, potentially pathogenic 5'UTR variants in *NMNAT1* makes this the first study to link 5'UTR regulatory variants to LCA with macular involvement. Overall, this study may impact upon the role of 5'UTR variations in other RDs in general.

540W

Disruption of a cis-acting microRNA site in FMR1 results in reduced translation of an FMRP reporter. *J. Suhl[1], R. Muddashetty[2], G. Bassell[2], S. Warren[1,3]*. 1) Human Genetics, Emory University, Atlanta, GA; 2) Cell Biology, Emory University, Atlanta, GA; 3) Biochemistry and Pediatrics, Emory University, Atlanta, GA.

Fragile X (FX) syndrome is the most common inherited form of intellectual disability (ID) and the leading known genetic cause of autism spectrum disorder. In a vast majority of FX cases, the gene responsible for the syndrome, FMR1, is silenced by hypermethylation that is triggered by expansion of an unstable CGG repeat in the promoter. However, few FX patients have been identified who exhibit normal repeat length, but harbor conventional mutations. Recently, we performed high throughput sequencing of FMR1 from 963 developmentally delayed and/or autistic males with normal repeat length, revealing a number of novel variants. Here, we report the functional analysis of a variant, identified in 6 patients but absent in controls (locally tested and databases), located in the 3'UTR of FMR1 (c.*746T>C) that is predicted to disrupt the interaction of miR-548p by interfering with seed region binding. A patient 3'UTR containing this variant causes a 25–40% decrease in reporter expression compared to a control, while mRNA levels remain unchanged, suggesting FMR1 is post-transcriptionally regulated via the 3'UTR. In support of the microRNA hypothesis, we find that the patient transcript co-immunoprecipitates poorly with AGO2 as compared to the control. Strikingly, we have found that the control reporter expression in cultured rat cortical neurons show a marked increase following synaptic stimulation that is entirely lost with the mutant reporter. Thus, activity-dependent translation of FMR1 in the postsynaptic space, recognized for a number of years but poorly understood, may be regulated by the microRNA pathways and that the c.*746T>C mutation appears to abrogate the local synthesis of FMRP following synaptic activity. Since this variant was seen in 0.6% of the patients studied, it may represent a new, clinically significant, variant of FMR1 and suggests that FMR1 sequencing in addition to repeat length determination could be warranted to rule out FMR1 involvement.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

541F

Is the biosynthesis of the major phospholipid class phosphatidylcoline in mammalian cells regulated by micro-RNA? *K. Hokynar, S. Hänninen, M. Hermansson, P. Somerharju.* Institute of biomedicine, Department of biochemistry and developmental biology, P.O.Box 63 (Haartmaninkatu 8), 00014 University of Helsinki, Helsinki, Finland.

Mammalian cells and cell organelles are surrounded by lipid bilayers formed mainly of phospholipids (PL), of which over 50 % are species of phosphatidylcolines (PC) and the rest include phosphatidylethanolamines, -inositols and -serines. In higher eukaryotes different cell membranes and membrane leaflets have specific PL compositions that are essential for proper cellular functioning and vitality and thus kept within narrow limits. Homeostasis of PL is maintained by biosynthesis, degradation, remodeling of acyl chains and interorganelle transport. Dozens of enzymes are involved in synthesis and degradation of PLs, but only little is known about their regulation at the transcriptional and post-transcriptional/translational level. MicroRNAs (miRNA) regulate gene expression post-transcriptionally by repression of translation or/and decay of the mRNA. miRNAs have been shown to be involved in many biological processes such as development and differentiation, recently also in regulation of cholesterol, fatty acid and glucose homeostasis. Our aim is to study whether miRNAs participate in the homeostasis of PLs in mammalian cells. HeLa cells were depleted of microRNAs by silencing the expression of enzymes Dicer1or Drosha by siRNA, after which the rate of PC biosynthesis was assessed by a method using deuterium labeled choline precursors and mass spectrometric (MS) analysis. Silencing of miRNA biogenesis resulted in an increase of PC synthesis (16 % in Dicer and 14 % in Drosha silenced cells). We also examined the expression levels of enzymes involved in PL biosynthesis and degradation after silencing of Dicer1 or Drosha. There was no significant change in the expression of PL degrading enzymes iPLA2β and iPLA2δ, neither at the mRNA nor protein level. However, silencing of miRNA synthesis resulted in an increase of expression of certain enzymes involved in PC synthesis (CDS1, CDS2 and SGMS2). In conclusion the biosynthesis of PC and the expression of certain PC biosynthetic enzymes increased in HeLa cells after silencing of miRNA biogenesis. These results suggest that miRNAs might be involved in the regulation of homeostasis of mammalian cell membrane PL content by effecting - directly or indirectly - the enzymes involved the biosynthesis of phosphatidylcholine.

542W

Poly(A) binding protein nuclear 1 (PABPN1) levels affect alternative polyadenylation. *E. de Klerk[1], A. Venema[1], S.Y. Anvar[1], J.J. Goeman[2], O. Hu[1], J.T. den Dunnen[1], S.M. van der Maarel[1], V. Raz[1], P.A.C. `t Hoen[1].* 1) Human Genetics, Leiden University Medical Center, Leiden, Netherlands; 2) Medical Statistics and Bioinformatics, Leiden University Medical Center, Leiden, Netherlands.

The choice for a polyadenylation site (PAS) determines the length of the 3'-untranslated region (3'-UTRs) of an mRNA. Inclusion or exclusion of regulatory sequences in the 3'-UTR may ultimately affect gene expression levels. Poly(A) binding protein nuclear 1 (PABPN1) is involved in polyadenylation of pre-mRNAs. A repeat expansion in PABPN1 causes oculopharyngeal muscular dystrophy (OPMD), a late onset and progressive muscle disorder. Microarray expression profiling of mice overexpressing expanded-PABPN1 showed alternative PAS usage in 8% of the interrogated genes. We hypothesized that previously observed disturbed gene expression patterns in OPMD muscles may have been the result of an effect of PABPN1 on alternative PAS usage. To explore PAS usage on a genome-wide level, we developed a single molecule PAS sequencing method. We identified 2,012 transcripts with altered PAS usage. In the far majority, alternative more proximal PAS were used. This resulted in overall shortening of 3'-UTRs and increased expression for most of the transcripts. We recapitulated these changes in PAS usage in myogenic cells by low overexpression of expanded but not wild-type PABPN1. Since expanded-PABPN1 is known to be trapped in intranuclear inclusions, we investigated the effect of shRNA mediated downregulation of PABPN1. We found that reduced Pabpn1 levels also resulted in shortening of 3'-UTRs. Our data suggest that PABPN1 is involved in PAS selection. We propose therefore that reduced availability of functional PABPN1 in OPMD muscles results in use of alternative, proximal PAS, leading to large-scale deregulation of gene expression.

543F

Identification of human RNA editing sites using publically available RNA-sequencing data. *G. Ramaswami, J.B. Li.* Department of Genetics, Stanford University, Stanford, CA.

Background: RNA editing plays an essential role in normal functioning of the nervous system. Post-transcriptional modification of adenosine to inosine (A-to-I) is the most common form of editing in mammals. The emergence of high-throughput RNA sequencing (RNA-seq) has resulted in a plethora of publically available RNA-seq datasets from various human tissues. This is a powerful resource for identification of A-to-I editing sites. Recently, we developed a strategy to accurately identify RNA editing sites by comparing the sequence differences between RNA-seq and coupled genomic DNA sequencing. Although this method was effective, its major drawback is the requirement of DNA sequencing data from the same individual to discriminate RNA editing sites from SNPs. This prohibits its wide application to the multitude of publically available RNA-seq datasets, because these data samples often do not have matched DNA sequence available. **Methods:** In this study, we developed two separate, but related methods to detect RNA editing sites using RNA-seq data by itself. Both of our methods rely on RNA-seq data from multiple individuals. In the first method, RNA variants are called in each sample, and all common SNPs are removed. To distinguish RNA editing sites from rare SNPs in the remaining RNA variants, we require the variants to be shared among multiple individuals because RNA editing is often biologically conserved. In the second method, the RNA-seq data are pooled to achieve higher read coverage, enhancing sensitivity for calling RNA variants. We eliminate rare SNPs by requiring a minimum frequency for RNA variants. We applied our two methods to identify human RNA editing sites using RNA-seq data from 40 lymphocyte cell lines and 50 brains. **Results:** In lymphocytes, we detected over 100,000 RNA editing sites in repetitive elements and over 1,000 editing sites in non-repetitive sequences. In brains, we detected over 300,000 RNA editing sites in repetitive elements and over 5,000 editing sites in non-repetitive sequences. Overall, we identify over 300,000 novel A-to-I editing sites, including many edits that modify protein-coding sequence. **Conclusions:** In this study, we were able to identify numerous A-to-I RNA editing sites in an extremely cost-effective manner by utilizing publically available RNA-seq data. Identification of novel A-to-I editing sites will be a significant step towards understanding how RNA editing regulates critical biological processes.

544W

Human ADAR proteins regulate RNA editing and gene expression as revealed in systems study. *I. Wang, J. Toung, A. Richards, E. So, Y. Zhao, W. Ankener, J. Devlin, Y. Liu, V. Cheung.* V Cheung Lab, Univ Pennsylvania, Philadelphia, PA.

ADAR family of adenosine deaminases mediates A-to-G RNA editing. Previously, it was thought that RNA editing is rare. However, recent deep sequencing of DNA and RNA has revealed that such events may be more common than previously known. In this study, we carried out DNA and RNA deep sequencing, RNA interference, RNA-immunoprecipitation (RNA-IP) and mass spectrometry in human B-cells to examine the functions of ADAR proteins. By coupling DNA and RNA sequence analyses with ADAR RNA-IP, we uncovered over 60,000 sites where the adenosines (A) are edited to guanosines (G); over 90% of these were previously unknown editing sites. These editing sites are found in mRNA and long non-coding RNAs (lncRNAs). Most of them are found in Alu repeats and in 3' UTRs of genes. Hyperediting is rather common; over 200 genes contain more than 50 A-to-G editing sites within each gene. ADAR gene knockdowns showed that in addition to editing, ADAR affects expression levels of several thousand genes and a number of microRNAs. The effects on gene expression occur in an editing-independent manner, suggesting ADAR proteins play important roles in cellular processes other than editing. Using proteomic approach, we identified more than 100 proteins in the ADAR protein complex. These ADAR-interacting proteins include splicing factors, heterogeneous ribonucleoproteins, several members of the dynactin protein family and other proteins involved in RNA processing. Our data support that ADAR functions together with other proteins and regulates multiple cellular processes including gene expression. Previously, we identified RNA-DNA sequence differences beyond the traditional A-to-G editing events. By ADAR gene knockdowns, we showed that while ADAR proteins mediated A-to-G editing, they do not play a role in other types of RNA-DNA sequence differences. In this presentation, I will describe results from genomic, transcriptomic and proteomic analyses which allowed us to identify tens of thousands of unknown A-to-G editing sites and to characterize ADAR's roles in gene regulation. The presentation will focus on findings that underscore the complexity of RNA processing and suggest that there are co- or post-transcriptional mechanisms yet to be determined.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

545F

Integrating functional and computational genomics to develop neuronal regulatory vocabularies. *X. Reed[1], L. Taher[2], G.M. Burzynski[1], C. Fletez-Brant[1], D. Lee[1,3], D.S. Gary[4], M.A. Beer[1,3], I. Ovcharenko[2], A.S. McCallion[1]*. 1) McKusick-Nathans Institute of Genetic Medicine, The Johns Hopkins University School of Medicine, Baltimore, MD; 2) Computational Biology Branch, National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD; 3) Johns Hopkins University, Department of Biomedical Engineering, Baltimore, MD; 4) International Center for Spinal Cord Injury (ICSCI), Kennedy Krieger Institute, Baltimore, MD.

The brain is a complex tissue comprising an array of cell types that express thousands of gene products. Disregulation of gene expression can give rise to phenotypic variation that may manifest as disease/risk. The genetics of many brain disorders, including Parkinson's disease, are poorly understood. By elucidating the encryption of transcriptional regulatory control we will gain insight into the molecular basis of cognitive variation and disease. To address the need for a neuronal regulatory vocabulary we initially used a machine learning approach to create a classifier that predicts sequences that may act as enhancers in the hindbrain. The classifier is trained on conserved enhancers previously shown to direct expression in the hindbrain, and looks for motifs enriched therein. Using the classifier we identified 40,000 sequences predicted to act as enhancers in the hindbrain. 51/55 (93%) predicted enhancers directed expression within the CNS of mosaic embryos, and 30/34 (88%) of stable lines show consistent expression in the hindbrain. Encouraged by this and other genome-wide enhancer classifier progress in our lab, we analyzed CNS ChIP-seq data available on the Human Epigenome Atlas. We called peaks for H3K4me1, a histone modification associated with active enhancers, in various sub-populations of the human brain (*substantia nigra*, temporal lobe, hippocampus, and anterior caudate). To identify those features that are specifically important in the substantia nigra we removed all sequences that were found within any of the other datasets. The remaining H3K4me1 flanked regions were run through an SVM classifier (Lee, Karchin, Beer 2011) to identify k-mers (motifs) that were overrepresented in the dataset. The ROC AUC is 65%, and the top k-mers returned correspond to binding sites for NR4A2 and FOXA2, transcription factors known to be important in the development of dopaminergic neurons. We selected ten regions which were located near genes characteristically expressed in dopaminergic neurons, or near SNPs which have been associated with Parkinson's disease, to be tested for enhancer function in zebrafish. These sequences are the focus of ongoing *in vivo* functional analyses and contribute to the larger search for neuron-specific vocabularies. To this end we have also initiated ChIP-seq analyses on *ex vivo* isolated and purified neuronal subpopulations, including cortical, striatal, cerebellar and hippocampal. We will discuss our progress in these efforts.

546W

Transcriptome genomic analysis using RNA-seq in three tissues of a twin cohort. *A. Buil[1,2], A. Brown[3], M. Davies[4], A. Viñuela[4], M. Gallardo[5], D. Glass[4], M. Blasco[5], R. Durbin[3], T.D. Spector[4], E.T. Dermitzakis[1,2] for the EUROBATS Consortium.* 1) Department of Genetic Medicine and Development, University of Geneva, Geneva, Switzerland; 2) Institute of Genetics and Genomics in Geneva, University of Geneva, Geneva, Switzerland; 3) Wellcome Trust Sanger Institute, Hinxton, United Kingdom; 4) Department of Twin Research and Genetic Epidemiology, King's College London, London, United Kingdom; 5) Telomeres and Telomerase Group, Molecular Oncology Program, Spanish National Cancer Centre (CNIO), Madrid, Spain.

We used RNA-seq data from a sample of twins to address the question of tissue specificity and genetic architecture of the regulation of gene expression. We used the EUROBATS project sample which consists of around 400 twin pairs with RNA extracted from fat, LCLs and skin. For each sample we sequenced the mRNA using 49bp paired-end sequencing in an Illumina HiSeq (2330 samples in total). Currently, we have analyzed a pilot set of samples from 200 individuals in the three tissues (600 samples). We obtained 18M exonic reads per sample on average. We quantified reads to known exons and used imputed genotypes from Hapmap II to discover cis eQTL. We observed 1901 genes with at least a significant cis eQTL in fat, 1776 in LCL and 1816 in skin (FDR=10%). That is an increase of 15% compared to eQTLs found with the same samples but using microarray gene expression measures. We also found that between 35% and 50% of the eQTL are shared among two tissues and that 25% of them are shared by the three tissues. RNA-seq data allows the quantification of allele specific expression (ASE). We performed a binomial test for each individual at each heterozygous site controlling for various confounders such as reference mapping bias. At a 10% FDR we observed that 9.5% of the heterozygous sites had a significant ASE effect in fat, 9.3% in LCL and 9.1% in skin. We also found that 8.8% of the significant ASE sites were present in the same individual in all three tissues and that the direction of the ASE ratio was consistent. Finally, using the ASE ratios of each sample at every heterozygous site, we defined an ASE distance among samples. We observed that the ASE distance is smaller for samples from the same tissue than for samples from different tissue. We also found that the ASE distance is smaller for MZ twins than for DZ twins and for DZ twins than for unrelated individuals. Given the useful twin structure of the data we plan to develop this ASE distance metric further in order to define the relative contribution of genetic to genetically independent epigenetic effects on allele specific expression. With the analysis of the whole dataset we will obtain more precise results for cis eQTL discovery and ASE analysis. At the same time, the larger dataset will allow us to look for trans eQTL in the three tissues. Furthermore we will determine the transcript structure of each tissue and the tissue specific genetic determinants of alternative splicing.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

547F

The genetics of gene expression, splicing and transcriptional efficiency through Poly-A selection and ribosomal depletion RNA sequencing in humans. *M. Pala[1,2,6], M. Marongiu[2], A. Mulas[2], R. Cusano[2], F. Crobu[2], F. Reinier[3], R. Berutti[3,6], M.G. Piras[2], C. Joens[3], D. Schlessinger[5], G. Abecasis[4], A. Angius[2], S. Sanna[2], F. Cucca[2,6,7], S.B. Montgomery[1,7].* 1) Stanford University, Stanford, CA; 2) Istituto di Ricerca Genetica e Biomedica (IRGB), CNR, Monserrato, 09042, Italy; 3) CRS4, Advanced Genomic Computing Technology, Pula, Italy; 4) Center for Statistical Genetics, University of Michigan, Ann Arbor, MI; 5) Laboratory of Genetics, National Institute on Aging, Baltimore, Maryland, USA; 6) Dipartimento di Scienze Biomediche, Università di Sassari, Sassari, Italy; 7) co-senior authors.

Most genome-wide association study (GWAS) signals fall within non-coding regions and are thus hard to interpret. Recent RNA-seq studies have demonstrated multiple modes of impact of GWAS variants on expressed transcripts. Notably, expression quantitative trait loci (eQTLs) for many human genes are emerging as informative intermediate phenotypes and it has been shown that genetic variants can affect not only overall gene expression but may alter individual transcripts. These isoform eQTLs can function at the level of alternative splicing, alternative polyadenylation sites or regulatory variation in alternative promoters, resulting in a combination of transcriptional, co-transcriptional (e.g., splicing), and post-transcriptional (e.g., RNA stability) mechanisms. In order to study in detail both transcriptional and post-transcriptional mechanisms regulating eQTL targets, we sequenced RNA from peripheral blood mononuclear cells for 80 Sardinian individuals enrolled in the SardiNIA cohort study with both poly-A selection and ribosomal depletion protocols. For each individual, an average of 50M of pair-end reads (51+51 nucleotides) was produced. We report the diversity of eQTLs observed in the Sardinian founder population, their relationship to other European populations and their implications on existing GWAS data. Furthermore, from the ribominus depletion protocol, we get significant amounts of unspliced, pre-mRNA, allowing us to profile, in combination with the poly-A selection protocol, the levels of both exonic and intronic sequences and uniquely report genetic effects on transcriptional efficiency. The same individuals have also been whole genome sequenced and characterized longitudinally for >200 quantitative traits, with measurements repeated in follow-up visits giving the possibility to cross compare different sources of data. Overall, this study offers unique opportunities to refine eQTLs in the founder Sardinian population, but also can improve our understanding of the genetic basis of transcriptional regulation and its impact on human phenotypic variation.

548W

Quantification of chitinase transcripts in human tissues. *M. Ohno, K. Tsuda, M. Sakaguti, Y. Sugahara, F. Oyama.* Dept Applied Chemistry, Kogakuin Univ, Hachioji, Tokyo, Japan.

Chitin is the second abundant polysaccharide in nature. It is an integral component of the fungal cell walls, the exoskeletons of crustaceans and insects, and the microfilarial sheaths of parasites. Although mammals do not have chitin and its synthase, genes encoding chitotriosidase and acidic mammalian chitinase (AMCase) and their translation products have been found in both human and mouse. Recent studies have shown an association between the mammalian chitinase and inflammatory lung diseases. For instance, chitotriosidase levels were elevated in the plasma of Gaucher disease and the bronchoalveolar lavage fluid of smokers with chronic obstructive pulmonary disease (COPD). AMCase expression and activity is upregulated in allergic airway responses in mouse models of asthma. Little is known, however, about the mutual regulation of the chitinases in human. We quantified expression levels of the chitinases and reference genes on the same scale using quantitative real-time RT-PCR system. Both chitinases were expressed in all tissues. Chitotriosidase mRNA was predominantly expressed in lung. AMCase mRNA was also highly expressed in lung, but their expression level was one third of the chitotriosidase. Our study shows that chitotriosidase plays a central role in human lung.

549F

Two functional variants in the promoter region of UBE2L3 confer risk for Systemic Lupus Erythematosus. *S. Wang[1], G.B. Wiley[1], F. Wen[1], J. Warner[1,2], J. Robertson[1], J.M. Guthridge[1], M. Dozmorov[1], J.D. Wren[1], J.A. James[1,2], P.M. Gaffney[1].* 1) Oklahoma Medical Research Foundation, Oklahoma city, OK; 2) College of Medicine, University of Oklahoma Health Sciences Center, Oklahoma City, OK.

Systemic lupus erythematosus (SLE) is an autoimmune disease with diverse clinical manifestations characterized by the development of pathogenic autoantibodies and target tissue damage. Genome-wide association (GWA) studies have identified genetic variants in the UBE2L3 region associated with SLE in Europeans and Asians, and we recently reported the association of a functional risk haplotype in the region of UBE2L3 with SLE that increases mRNA expression of UBE2L3. To discover and characterize functional variants carried on the SLE associated UBE2L3 risk haplotype, electrophoretic mobility shift assay (EMSA) was performed to test the differential binding affinity of variants in and around the promoter region of UBE2L3 to the nuclear proteins prepared from five different types of cell lines, including, U937, THP1, HEK293T, HeLa, and EBV-transformed B cell lines. SNPs were prioritized for functional testing using Genomatix and Genome Runner. A proportion of the cells were subjected to the preparation of nuclear proteins and the isolation of total RNAs. The rate of growth of EBV transformed B cell line was calculated by counting the total number of cells every 24 hours. Messenger RNA expression of UBE2L3 was determined using quantitative RT-PCR and normalized to a control gene (HMBS). Data obtained by EMSA demonstrates that two variants (rs140490 and rs2266959) function in regulating transcription factor binding. The variant rs140490 binds EBV cell-type dependent nuclear protein complex with enhanced affinity (P=0.0031). The binding of the protein complex to the variant region increases mRNA transcription of UBE2L3 in a time-dependent manner. In contrast, rs2266959 results in a reduced affinity of binding (P= 0.0014) to a common nuclear protein complex present in multiple cell lines. Here, we observed evidence for two functional alleles of UBE2L3: rs140490 is located in the promoter region that enhances the affinity of binding to a EBV cell-type specific protein complex in a time-dependent manner; rs2266959 is located in the first intron of UBE2L3 that results in a reduced affinity of binding to a common nuclear protein complex from different cell lines. Our data suggest that functional variants on the UBE2L3 risk haplotype are responsible for the gene expression phenotype. Further characterization of the functional variants of UBE2L3 will help to understand the role this risk haplotype plays in SLE pathogenesis.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

550W

A 660 kb deletion centromeric to the lamin B1 gene (*LMNB1*) mimics LMNB1 duplication and causes adult-onset autosomal dominant leukodystrophy. *E. Giorgio[1], D. Robyr[2], E. Di Gregorio[1,3], D. Lacerenza[1], G. Vaula[4], A. Brusco[1,3], S.E. Antonarakis[2], A. Brussino[1].* 1) University of Torino - Dpt of Genetics, Biology and Biochemistry, Torino, Torino, Italy; 2) Department of Genetic Medicine and Development, University of Geneva Medical School and University Hospitals of Geneva, Geneva, Switzerland; 3) SCDU Medical Genetics, AOU S.Giovanni Battista, Torino, Italy; 4) Department of Neuroscience, AOU S.Giovanni Battista, Torino, Italy.

We report a position effect mutation in the lamin B1 locus due to a genomic deletion centromeric to the gene that moves a cryptic enhancer closer to the *LMNB1* gene, resulting in its overexpression, and mimicking Adult-onset Autosomal Dominant LeukoDystrophy (ADLD). ADLD is a rare central nervous system demyelinating disease presently associated to lamin B1 (*LMNB1*) gene duplication. We collected a large Italian pedigree segregating an autosomal dominant form of late onset leukodystrophy clinically overlapping ADLD. Linkage analysis mapped the disease to the LMNB1 locus, although *LMNB1* was not duplicated or mutated. However, *LMNB1* was overexpressed in fibroblasts both at mRNA and protein level as in reported ADLD patients. Using array-CGH analysis, we identified a 660 kb deletion located approximately 66 kb centromeric to the *LMNB1* gene, that we hypothesised alters *LMNB1* gene expression. The deleted segment spanned three genes, *GRAMD3*, *ALDH7A1* and *PHAX*, and its boundaries were defined within two Alu repeats. shRNA silencing assays on HeLa cells showed *GRAMD3*, *ALDH7A1* and *PHAX* knockdown did not affect *LMNB1* expression. To search for non-coding regulatory elements, we performed circular chromosome conformation capture (4C) on fibroblasts from a patient and a healthy control, using lamin B1 promoter as a bait. We identified four potential regulatory regions: region A, located within the deletion (0.12 Mb centromeric to *LMNB1* gene), which interacted with the promoter both in the patient and the control DNA, and regions B, C and D (0.77, 1.9, and 2.0 Mb centromeric to the *LMNB1* gene and upstream of the deletion), whose interaction was present in the patient only. Using a dual-luciferase assay, each putative regulatory region was cloned into the reporter vector (a modified pGL4.10) upstream of the lamin B1 promoter. Two regions behaved as enhancers: region A and B: both were not previously reported as *LMNB1* enhancers and are not evolutionary conserved. Therefore, in our patients, the presence of the deletion takes away region A in one of the two alleles, bringing region B closer to *LMNB1* and substituting region A in its enhancer activity. Given that dual luciferase assay demonstrates that region B is stronger than region A (luciferase signal changes from 2.8 to 3.4 fold vs. control) we speculate this is causing lamin B1 overexpression. Our results confirms that mutations in regulatory elements may be important determinants of Mendelian phenotypes.

551F

Variation in Transcription Factor Binding Among Humans. *F. Grubert[1,2], M. Kasowski[1,2], C. Heffelfinger[2], M. Hariharan[1,2], A. Asabere[1,2], S. Waszak[3], L. Habegger[4], J. Rozowsky[5], M. Shi[1,2], A. Urban[2,6], M. Hong[2], K. Karczewski[1], W. Huber[3], S. Weissman[6], M. Gerstein[4,5,7], J. Korbel[3,8], M. Snyder[1,2].* 1) Dept Genetics, Stanford University, Stanford, CA; 2) Department of Molecular, Cellular, and Developmental Biology, Yale University, New Haven, CT; 3) Genome Biology Research Unit, European Molecular Biology Laboratory, Heidelberg, Germany; 4) Program in Computational Biology and Bioinformatics, Yale University, New Haven, CT; 5) Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT; 6) Department of Genetics, Yale University School of Medicine, New Haven, CT; 7) Department of Computer Science, Yale University, New Haven, CT; 8) European Molecular Biology Laboratory-European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton CB10 1SD, UK.

Differences in gene expression may play a major role in speciation and phenotypic diversity. We examined genome-wide differences in transcription factor (TF) binding in several humans and a single chimpanzee by using chromatin immunoprecipitation followed by massively parallel sequencing. The binding sites of RNA polymerase II (Pol II) and a key regulator of immune responses, nuclear factor kB (p65), were mapped in 10 lymphoblastoid cell lines, and 25 and 7.5% of the respective binding regions were found to differ between individuals. Binding differences were frequently associated with single-nucleotide polymorphisms and genomic structural variants, and these differences were often correlated with differences in gene expression, suggesting functional consequences of binding variation. Furthermore, comparing Pol II binding between humans and chimpanzee suggests extensive divergence in TF binding. Our results indicate that many differences in individuals and species occur at the level of TF binding, and they provide insight into the genetic events responsible for these differences.

552W

Widespread R-loop formation in the human genome: implications for epigenetic control and beyond. *Y. Lim, P. Ginno, P. Lott, I. Korf, F. Chedin.* Molecular and Cellular Biology, University of California, Davis, CA.

Long R-loop structures can form upon transcription through GC-skewed sequences when a G-rich RNA is transcribed. As we recently showed, unmethylated CpG islands (CGI) promoters - the main promoter class in the human genome - are characterized by a sharp rise in GC skew immediately downstream of the transcription start site and represent prominent hotspots of R-loop formation. Co-transcriptional R-loop formation was further proposed to directly contribute to the protection of CGI promoters from DNA methylation, a prevalent epigenetic mark associated with transcriptional silencing. We have now computationally profiled GC skew patterns using a high-sensitivity iteration of the SkewR Hidden Markov Model and measured R-loop formation genome-wide using DNA:RNA immunoprecipitation coupled to high throughput sequencing (DRIP-seq) in human fibroblasts. This analysis reveals the following novel characteristics. (1) GC skew patterns at the 5'-end of genes can distinguish between four broad promoter classes with distinct, genetic, epigenetic, and ontology signatures, including a class of promoters highly associated with developmentally-regulated transcription factors. (2) The X chromosome shows evidence for depletion of promoter GC skew over its evolutionary history and R-loop formation is accordingly reduced on the X, consistent with this chromosome undergoing broad silencing during X-inactivation. (3) In addition to the 5'-end of genes, GC skew and R-loop formation are prevalent at the 3'-end of thousands of genes, particularly genes located in gene-dense regions. This observation is consistent with recent evidence suggesting that R-loop formation is required for efficient transcription termination at the 3'-end of genes. Overall this works indicates that GC skew patterns and R-loop formation play important roles in defining the boundaries of human genes, ensuring proper transcription initiation and termination, and molding gene neighborhoods along human chromosomes.

553F

Genome-wide transcriptome profiling reveals the functional impact of rare de novo and recurrent CNVs in autism spectrum disorders. *R. Luo[1,2], S. Sanders[3], Y. Tian[4], I. Voineagu[5], N. Huang[6], S. Chu[7], L. Klei[9], C. Cai[1,8], J. Ou[5], J. Lowe[5], M. Hurles[6], B. Devlin[9], M. State[3], D. Geschwind[1,2,5], ACE network consortium.* 1) Human Genetics, UCLA, Los Angeles, CA; 2) Center for Autism Research and Treatment, Semel Insittute, University of California Los Angeles, Los Angeles, CA 90095, USA; 3) Programs in Neurogenetics and Human Genetics and Genomics, Child Study Center and Departments of Psychiatry and Genetics, Yale University School of Medicine; 4) Interdepartmental Ph.D. Program in Bioinformatics, University of California Los Angeles; 5) Neurogenetics Program, University of California, Los Angeles; 6) Wellcome Trust Sanger Institute, Hinxton, Cambridge; 7) Department of Statistics, Carnegie Mellon University; 8) Department of Biostatistics, David Geffen School of Medicine, University of California; 9) Department of Psychiatry, University of Pittsburgh School of Medicine.

Autism Spectrum Disorder (ASD) is a disorder of neural development that is characterized by deficits in social interaction, communication, and the presence of restricted and repetitive behavior. Recent studies have found that rare Copy Number Variants (CNVs) are a major contributor to ASD pathophysiology. However, the functional impact of these rare variants is largely unexplored. We reasoned that if autism is caused in some cases by CNVs, genes within the CNVs may exhibit significant up- or down- regulated expression, and such genes could be identified by microarray analysis of gene expression. To highlight pathogenic CNVs and functional alterations in gene expression, we performed functional genomic profiling in lymphoblast cell lines from 439 discordant siblings from 244 Simons Simplex families analyzed on the Illumina Human Ref8 version3 chip. The distribution of gene expression was analyzed in autistic probands and unaffected siblings. Genes that were two or three standard deviations (SDs) from mean population expression levels were deemed potential outlier (dysregulated) genes. Our results show that outlier genes identified in probands, but not in unaffected siblings, fall into neural-related pathways. By combining expression data with CNV data from the same population (Sanders et al. 2011), we demonstrate that outlier genes are significantly enriched within the most pathogenic CNVs (rare de novo CNVs). This analysis suggests that deletions at 3q27, 3p13 and 3p26 and duplications at 2p15 and 3q14, all of which show a significant enrichment in outlier genes compared to genome background, are likely new ASD susceptibility loci. We analyzed gene expression in multiple individuals harboring one of three recurrent CNVs (16p11.2 microdeletions, 16p11.2 microduplications, and 7q11.23 duplications) and find that each harbors a distinct set of dysregulated genes and pathways. Our analyses also show that specific genes, including within the 16p CNV interval TAOK2, CORO1A, KCTD13 and QPRT, are correlated with differences in head circumference. This study is the first to provide evidence that pathogenic structural variants lead to transcriptome alterations in ASD at a genome-wide level and demonstrates the utility of this approach for prioritization of genes for further downstream functional analysis.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

554W

Transcriptional alterations underlying variable phenotypic expressivity associated with 16p11.2 microdeletion/microduplication syndrome within multiplex autism families. *I. Blumenthal[1], J.Z. Levin[3], Y. Shen[1,5,6], C. Ernst[1], C.C. Morton[3,5,7], J.F. Gusella[1,2,3], M.E. Talkowski[1,3,4]*. 1) Center for Human Genetic Research, Massachusetts General Hospital, Boston, MA; 2) Department of Genetics, Harvard Medical School, Boston, MA; 3) Program in Medical and Population Genetics, Broad Institute, Cambridge, MA; 4) Department of Neurology, Harvard Medical School, Boston, MA; 5) Department of Pathology, Massachusetts General Hospital, Boston, MA; 6) Department of Laboratory Medicine, Children's Hospital Boston, Boston, MA; 7) Departments of Obstetrics, Gynecology and Reproductive Biology, Brigham and Women's Hospital, MA.

Autism spectrum disorder (ASD) is a complex, heterogeneous developmental disorder affecting approximately 1% of the population. Collectively, recurrent microdeletions and duplications represent the largest known risk factor for ASD. Of these, copy number variation (CNV) of a 593kb region of 16p11.2 is among the highest contributors to ASD, accounting for approximately 1% of all cases. However, the penetrance and phenotypic expressivity associated with deletion or duplication of the region are highly variable with deletion presenting a stronger effect than duplication. In this study, we describe the precise consequences of 16p11.2 dosage imbalance on mRNA expression within the 16p11.2 region, and its regulatory impact on other transcripts both in cis and in trans. We performed strand-specific RNA sequencing on lymphoblasts of a unique cohort with 35 individuals from 7 families sourced from the Autism Genetic Resource Exchange (AGRE). Each family harbors a segregating 16p11.2 microdeletion or microduplication and heterogeneity of both genotype and phenotype, *e.g.* some individuals with ASD are discordant for the CNV, while others with 16p11.2 dosage imbalance are discordant for phenotype. Pair-wise comparisons of sequencing data within and between families uncovered significant differential gene expression in cases compared to individuals without a 16p11.2 CNV. Notably, while deletions were associated with a uniform mRNA reduction of genes within the 16p11.2 region (~0.5X normal expression), duplications were associated with a more heterogeneous outcome, ranging from 1.2-2X normal expression. CNVs within the 16p11.2 region also lead to significant genome-wide expression changes; 379 genes were significantly dysregulated in concordant affected individuals from deletion families while only 135 genes had altered expression in duplication families, though there was meaningful overlap in loci between groups. Our analyses show broad alterations in gene expression regulation resulting from recurrent, reciprocal dosage imbalance of 16p11.2 with impact on a spectrum of functional loci from chromatin modifiers, transcriptional regulators and long non-coding RNAs. Analyses of patient-specific induced pluripotent stem cell lines, neural progenitor cells and mature neurons are ongoing, however these studies provide insight into the transcriptomic alterations that may underlie variable phenotypic expression resulting from structural alterations of 16p11.2.

555F

Differential reduced expression of IRS-1 in visceral and subcutaneous adipose depots in morbidly obese and type2 diabetic subjects and their correlation with metabolic profiles with 2 novel mutations sequenced in Asian Indians. *M. Sharma[1], K. Luthra[1], S. Aggarwal[2], N. Vikram[3], A. Misra[4]*. 1) BIOCHEMISTRY, ALL INDIA INSTITUTE OF MEDICAL SCIENCES, NEW DELHI, DELHI, India; 2) SURGERY, ALL INDIA INSTITUTE OF MEDICAL SCIENCES, NEW DELHI, DELHI, India; 3) MEDICINE, ALL INDIA INSTITUTE OF MEDICAL SCIENCES, NEW DELHI, DELHI, India; 4) Fortis-C-DOC Centre of Excellence for Diabetes, Obesity, Metabolic Diseases and Endocrinology, Centre of Internal Medicine (CIM), Fortis Hospital, New Delhi, India.

We looked for differential expression of IRS-1 at regional fat depots (visceral and subcutaneous) in morbidly obese diabetic patients, and correlate genotype-phenotype traits if any, in the above subjects. A total of 35 morbidly obese (BMI>40Kg/mSuperscript Text2) in presence of co-morbidities and 15 non-obese as controls (BMI<25Kg/m2) adipose tissue were obtained from subjects undergoing Bariatric surgery and elective abdominal surgery respectively. Total RNA isolated from RNeasy lipid extraction kitand expression checked in Real-Time qPCR. Biochemical parameters were assessed in the subjects. Genotyping of IRS-1 Gly972Arg was performed by PCR-RFLP. Sequencing of whole exon of IRS-1 with respective primers was carried out to scan for any novel variations. Informed consent was signed from each subject.The morbidly obese subjects showed a marked increase in serum insulin levels and the lipoprotein fractions except, HDL-C. We observed a decrease in the mRNA expression of IRS-1 in the adipose tissue of the visceral and subcutaneous depots in the morbidly obese subjects than controls (non-obese), with a marked reduction in the visceral adipose tissue as compared to the subcutaneous adipose tissue of the morbid obese subjects. We did not observe any correlation between the IRS-1 expression in the adipose tissue and the IRS-1 Gly972Arg polymorphism. A salient finding of this study was the presence of two novel variations in IRS-1 gene, not reported previously: one, in the codon 1102 with GAG>AAG, a non-synonymous mutation, encoding lysine in place of glutamate in the morbid obese subjects with insulin resistance. The other novel variation, a deletion at codon 658, at 1979bp was observed (submitted to Genbank). Our observation of a substantial reduction in the expression of IRS-1 in the visceral and subcutaneous adipose tissues of the morbidly obese and diabetic subjects strongly suggest that the IRS-1 expression may serve as a novel marker for predicting the insulin responsiveness in the adipose tissue. This may further be confirmed by performing functional studies. Such studies may be of importance in unraveling the effect of alterations in the distribution of localized fat in the subcutaneous vs. visceral adipose tissues and its influence on insulin action, particularly at the level of IRS-1. Besides, the novel mutations sequenced, alone or in combination with other reported mutations in the IRS-1 may contribute substantially to develop obesity and/or T2DM.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

556W

Quantitative analysis of the human beta cell transcriptome. *A.C. Nica, H. Ongen, J. Irminger, S. Dupuis, A. Planchon, L. Romano, S.E. Antonarakis, P.A. Halban, E.T. Dermitzakis.* Genetic Medicine & Development, University of Geneva, Geneva, Switzerland.

We present here the first detailed description of the human beta cell transcriptome using deep coverage RNA sequencing (RNA-seq), perform the first comparison of beta cell - islet expression profiles and offer initial insights into the genetic control of beta cell gene expression. We collected islets from 11 cadaveric pancreata of individuals without documented diabetes and FACS sorted islet cells to obtain “beta cell” (approx. 90% pure) and beta-cell depleted “non-beta” (N=5) populations. All RNA samples were deep-sequenced, attaining an unprecedented transcriptome resolution (~ 200 million quality filtered reads per sample). As expected, insulin (INS) is by far the most abundantly transcribed gene, followed by INS-IGF2 and IGF2, making up ~ 39%, 9% and 3% of the total nuclear beta cell transcriptome respectively. While the relative percentages are lower, this ranking is maintained in the islet (INS 27%, INS-IGF2 5%, IGF2 1.6%). The ranked correlation between beta cell and islet-expressed genes is high (Spearman rho = 0.934) and we estimate that 85% of the variance in beta cell gene expression can be explained by using islet expression as proxy. Beta cells and islets separate markedly from non-beta cells, with insulin, IGF-2, glucagon, transthyretin, pancreatic polypeptide and somatostatin driving most of this separation. In addition to RNA-seq, we have performed medium depth (16 ×) whole-genome sequencing for 7 of the 11 samples, followed by variant calling and 1000 Genomes based imputation. We detect an abundance of significant allelic differences in transcript levels at heterozygous sites (9% of total sites tested, corresponding to 55.42% of the genes), highly overlapping in beta and islet cells (89.11% estimated median enrichment of significant beta cell ASE pvalues in islets and 88.11% vice-versa). A subset of significant ASE sites maps to diabetes-associated genes, with a significant enrichment (Fisher's pvalue: 0.015) for genes associated with fasting glucose/insulin levels. Furthermore, we note a significant ASE enrichment at heterozygous expression quantitative trait loci (eQTLs) discovered in other tissues compared to homozygous eQTLs, suggesting that a proportion of previously documented regulatory variants are also active in beta cells. Given these, we provide here a valuable resource to the community, helping our understanding of the genetics of both type 1 and type 2 diabetes.

557F

Novel sequence motif regulates pluripotency. *Y. Song[1], S. Bao[1], G. Niu[1], Y. Zhao[2], H. Zhu[3].* 1) Dept Biochemistry, Univ Hong Kong, Hong Kong, China; 2) 6th Affiliated Hospital, Sun Yat Sen Univeristy; 3) Dept microbiology, Univ Hong Kong, Hong Kong, China.

The pluripotency of embryonic stem cells (ESCs), combined with theoretically unlimited self-renewal characteristic, has greatly promoted the development of ESCs applying in regenerative medicine and organ transplantation fields. Deciphering the mechanisms of pluripotency would offer a much deeper understanding of the specific reprogramming process of ESCs. In our study, we tried to understand the networked regulations on the genomic level. Thanks to recently developed next-generation sequencing technologies, genome-wide maps of chromatin state has been established. The patterns of mammalian chromatin can offer insights into the locations and functions of underlying regulatory elements and genes. Here, we tried to recognize the specific regulatory sequence characteristics from the open chromatin signals on ES cells' genome. After obtaining `informative' ESC-specific open chromatin (OC) bins, which were open chromatin sequences with length around 500 bp and expressed significantly higher in ESCs than other tissues, we applied MEME's `discriminative motif discovery' model to extract ESC-specific OC motifs. Then we filtered motifs that were conserved with each other and significantly more enriched in ESC OC bins compared with non-ESC OC bins. Around 60 sequence motifs were disclosed and we found that there is bias of nucleotide occurrence in some of the identified ESC open chromatin regions. This knowledge spreads us a much clearer image about the short and specific sites, which would introduce TF-DNA interactions in vivo and maintain pluripotent stability, on the human genome. Furthermore, such knowledge may also help us design novel functional elements that could be in response to the inside environmental signals in the pluripotent cells.

558W

Quantitative expression analysis of mouse chitinases: acidic mammalian chitinase is the major transcript in stomach. *F. Oyama, M. Ohno, K. Tsuda, M. Sakaguchi, Y. Sugahara.* Dept Applied Chemistry, Kogakuin Univ, Hachioji, Tokyo, Japan.

Two active chitinolytic enzymes, chitotriosidase and acidic mammalian chitinase (AMCase), have been identified in mouse and human, although mammals do not have chitin and chitin synthase. Since these chitinolytic enzymes are upregulated in Gaucher disease or mouse models of asthma, both chitinases have been regarded primarily as a host-defense mechanism against chitin-containing pathogens and/or play important roles in the pathophysiology. Little is known, however, about mutual regulation of gene expression between chitotriosidase and AMCase in vivo. Here we set up a real-time PCR system capable of quantifying expression levels among the chitinases and reference genes. We found that both chitinases mRNAs are predominantly expressed in mouse stomach. Mouse tissues express more AMCase mRNA than chitotriosidase mRNA, whereas chitotriosidase mRNA was prevalent only in eyes. In stomach, AMCase mRNA is expressed 7–10 fold higher than those of housekeeping genes and expressed comparative level with pepsinogen C (progastricsin) mRNA, a major component of the gastric mucosa. Thus, AMCase mRNA is a major transcript in mouse stomach, suggesting that AMCase functions as a digestive enzyme breaking down polymeric chitin in gastric juice.

559F

The Japanese firefly, Luciola cruciata, gene: small intron and their insufficiency of splicing in CHO cells. *M. Ishii[1,3], R. Kojima[1], S. Fukuda[1], M. Sakaguchi[1], Y. Sugahara[1], M. Kamaya[2,3], F. Oyama[1,3].* 1) Biotechnology Laboratory, Department of Applied Chemistry, Faculty of Engineering, Kogakuin University, 2665-1 Nakano, Hachioji, Tokyo 192-0015, Japan; 2) Enviroment Analytical Chemistry Laboratory, Department of Environmental and Energy Chemistry, Faculty of Engineering, Kogakuin University, 2665-1 Nakano, Hachioji, Tokyo 192-0015, Japan; 3) The Firefly-Breeding Project at Hachioji Campus, Kogakuin University, 2665-1 Nakano, Hachioji, Tokyo 192-0015, Japan.

Most genes in eukaryotes are interrupted by non-coding sequences called intron. So far, they had been considered as unimportant sequences which are to be removed from an unspliced precursor RNA in order to generate the functional mRNA. Recently, it has been shown that intron also has important functions, to increase mRNA and protein levels in mammalian cells. In this study, we cloned and sequenced the genomic and cDNA clones of the Japanese firefly, L. cruciata. The luciferase gene contains six introns, all of which were very short; the longest (intron 1) was 98 bases in length, and the smallest (intron 5) was only 44 bases in length. Next, we examined the effect of the small introns on the expression levels of the luciferase in mammalian cells. The luciferase gene (with 6 introns) and cDNAs (without introns) were inserted into mammalian expression vectors and transiently expressed in CHO cells respectively. We found that the luciferase gene didn't give enzymatically active firefly luciferase, whereas cDNA normally did. RT-PCR analysis also indicated that luciferase pre-mRNA transcribed from the gene was detected without splicing in the CHO cells. These results suggest that firefly small introns are insufficient to be spliced in mammalian cells.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

560W

Expression QTL interactions in a cohort of 869 individuals reveals sex specific, cis-cis, trans-cis and trans-trans interactions. *J. Bryois[1], A. Buil[1], DM. Evans[2], DF. Conrad[3], SM. Ring[2], M. Hurles[3], P. Deloukas[3], GD. Smith[2], ET. Dermitzakis[1].* 1) University of Geneva, Geneva, Switzerland; 2) University of Bristol, Bristol, UK; 3) The Wellcome Trust Sanger Institute, Cambridge, UK.

In order to better understand how genetic variations affect gene expression we mapped 2'543'887 imputed SNPs and 3329 copy number variants (CNVs) to gene expression in LCLs in 869 individuals from the ALSPAC cohort. We find 3890 genes with a variant affecting gene expression in cis (FDR= 4.1%) and 167 genes with a variant affecing gene expression in trans (FDR= 9.9%). We find that cis-eQTLs are mainly driven by SNPs while trans-eQTLs are largely driven by CNVs. 47 genes with a cis-eQTLs, that is also a trans-eQTL, show a strong enrichement in immune GO terms but no enrichment in transcription regulation GO terms, suggesting that immunity and other traits might be under the same genetic control. The large sample size of the data set allows investigating how the effect of the discovered eQTLs can be conditional to other varaiables, such as other SNPs or the sex of the individuals. Using a linear model with an interaction term, we find that 176/51 (5%FDR) of the best cis/trans-eQTLs per gene act in a sex specific manner. We also looked for cis-cis interactions using SNPs marginally (pvalue ≤ 10-3) and independently associated to the same gene and found 133 cis-cis interactions (5% FDR). Trans-cis interactions were also investigated, which lead to the discovery of 6 trans-cis interactions (5% FDR) when both SNPs are affecting the same gene, 31 interactions (5% FDR) when the model explains the expression of the cis regulated gene and 4 interactions (5% FDR) when the model explains the expression of the trans regulated gene. We also looked for trans-trans interactions and found 9 interactions (5% FDR) when both SNP are independently associated to the same gene, no interaction for genes regulated by a trans-eQTL (excluding trans-eQTLs affecting the same gene) and 20 interactions (5% FDR) for all genes not regulated by a trans-eQTL. Finally we are investigating if cis/trans eQTLs are also GWAS SNPs and if they interact with GWAS SNPs, as this could give clues on the possible mechanism leading the SNPs to the diseases.

561F

Tissue-specific effects of genetic variation and DNA methylation on gene regulation. *E.T. Dermitzakis[1,2], M. Gutierrez-Arcelus[1,2], S.B. Montgomery[3], T. Lappalainen[1,2], H. Ongen[1,2], A. Yurovsky[1,2], J. Byrois[1,2], A. Buil[1,2], T. Giger[1,2], L. Romano[1,2], A. Planchon[1,2], E. Falconnet[1,2], I. Padioleau[1,2], C. Borrel[1,2], A. Letourneau[1,2], P. Makrythanasis[1,2], M. Guipponi[1,2], C. Gehrig[1,2], S.E. Antonarakis[1,2].* 1) Dept of Genetic Medicine and Development, University of Geneva, Geneva, Switzerland; 2) Institute of Genetics and Genomics in Geneva (iGE3), Geneva, Switzerland; 3) Dept of Pathology, Stanford University, Stanford, California, United States of America.

Multiple studies have demonstrated the importance of genetic variants affecting gene expression (eQTLs) and initial limited studies have identified genetic variants affecting local DNA methylation levels (mQTLs). Moreover, DNA methylation is known to be associated with gene silencing, and more recently with active gene expression when present within the gene body. Yet the mechanisms by which genetic variation and DNA methylation affect gene expression and tissue specificity are not well understood. In this study, we use a cohort of approx. 180 newborn Caucasian individuals who we genotyped (2.5 million SNPs) and for whom we measured mRNA levels through RNA-seq in three cell-types derived from cord blood and umbilical cord: lymphoblastoid cell lines, T-cells and fibroblasts. For approx. 100 samples of each of the three cell-types, we have assayed DNA methylation levels in more than 400,000 CpG sites through bisulfite conversion and bead chips. We characterized the tissue-specificity of > 4000 eQTLs discovered in the three tissues, allele-specific expression, mQTLs and eQTMs (methylation-expression associations). We also identified frequent differences in alternative splicing between different tissues, as well as genetic variants and, intriguingly, a modest amount of methylation sites affecting splicing patterns. Additionally, we tested the contribution of different mechanistic models for the interplay between genetic, epigenetic effects and gene expression. We provide evidence for synergistic interactions between SNPs and methylation sites to expression. We also found a that in the majority of cases a SNP affects gene expression and DNA methylation independently, but also a number of cases in which the SNP effect on gene expression is mediated through methylation, or the SNP affects gene expression which in turn modulates DNA methylation. Finally, we describe the mechanistic differences between patterns of methylation associated with increase and decrease of gene expression. Overall, our results provide insights into the genetics and epigenetics involved in gene regulation in several cell-types and in samples coming from a unique time point in development. We highlight the importance of studying and integrating different dimensions of information for better understanding complex traits.

562W

Multi-scale Integration Systematically Connects Disease-Associated Variants to Molecular Pathology. *K.J. Karczewski[1,2], J.T. Dudley[3], R. Chen[3], A.J. Butte[3], M. Snyder[2].* 1) Biomedical Informatics Training Program, Stanford University School of Medicine, Stanford, CA; 2) Department of Genetics, Stanford University School of Medicine, Stanford, CA; 3) Division of Systems Medicine, Department of Pediatrics, Stanford University School of Medicine, Stanford, CA.

Genome-wide association studies have discovered many genetic loci associated with disease, but the molecular basis of these associations is often unresolved. Genome-wide regulatory and gene expression profiles measured across individuals and diseases reflect downstream effects of genetic variation, and may allow for functional annotation of disease-associated loci. Here, we present the first systematic approach effort to combine disease association, transcription factor binding, and gene expression data to assess the functional consequences of variants associated with hundreds of human diseases. In an analysis of genome-wide binding profiles of NFκB, we find that disease-associated SNPs are enriched in NFκB binding regions overall, and specifically for inflammatory mediated diseases, such as asthma, rheumatoid arthritis, and coronary artery disease. Using genome-wide binding variation information for 8 fully sequenced individuals, we find regions of NFκB binding correlated with disease-associated variants in a genotype-specific and allele-specific manner. Furthermore, we show that this binding variation is often correlated with expression of nearby genes, which are also found to have altered expression in independent profiling of the variant-associated disease condition. In this systematic approach, we close a major loop in biological context-free association studies and assign putative function to many disease-associated SNPs.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

563F

Analysis of primary Natural Killer cells reveals multiple novel cell-specific eQTL. *S. Makino[1], B.P. Fairfax[1,2], E. Ng[1], J.C. Knight[1].* 1) Wellcome Trust Centre for Human Genetics, Oxford, University of Oxford, OX3 7BN, United Kingdom; 2) Oxford Cancer Centre, Churchill Hospital, Oxford, OX3 7LJ, United Kingdom.

We have previously demonstrated that the majority of eQTL in freshly purified primary $CD19^+$ B-cells and $CD14^+$ monocytes are specific to these cell types. Here we further investigate the specificity of eQTL in the different cellular subsets that compose peripheral blood mononuclear cells by analysing gene expression in $CD56^+$ Natural Killer (NK) cells. NK cells form a vital arm of the innate immune response with a crucial role in immune surveillance, demonstrating immunoreactivity to virally infected and malignantly transformed cells. Identification of loci that regulate gene expression in these cells should further inform our understanding of how common genetic variants impact on innate immune responses and disease susceptibility.

NK cells were purified from freshly prepared peripheral blood mononuclear cells using a negative selection protocol. Individuals were genotyped at >730,000 loci using Illumina OmniExpress beadchips. Gene expression profiles from NK-cells, in addition to monocytes and B-cells concurrently purified from the same individuals, were assayed using Illumina HT-12 whole genome expression arrays (116 individuals, 348 arrays). To perform eQTL analyses we used ANOVA methodology, incorporating principal components from expression profiles as covariates.

Using this approach we find that at a false discovery rate of 0.01, we observe a total of 41,177 SNP-probe pairs in *cis* across all 3 cell types, 71% of which were unique to a specific cell-type and 11.7% of which were shared amongst all 3 cell-types. NK-cells demonstrated a greater number of eQTL than the other 2 cell-types and showed marginally greater overlap in-terms of shared eQTL with myeloid derived monocytes than lymphoid derived B-cells. Examples of the most significant novel NK cell specific eQTL include those forming to the basic helix-loop-helix protein encoding gene *TCF25*, rs1007932, $P=7\times10^{-38}$; the golgi associated gene *COG5*, rs34325395, $P=5.3\times10^{-40}$ and the vaccinia H1-related phosphatase encoding DUSP3, rs1234612, $P=2.7\times10^{-36}$.

This analyses further informs our understanding of how common genetic variation can modulate gene expression in defined cell populations and thus impact upon human health and disease susceptibility.

564W

The impact of polymorphisms within probe sequences on expression QTL studies. *A. Ramasamy[1,2], D. Trabzuni[2,3], R.J. Gibbs[2,4], A. Dillman[4], D.G. Hernandez[2,4], S. Arepalli[4], R. Walker[5], C. Smith[5], A.A. Shabalin[6], Y. Li[6,7], A.B. Singleton[4], M.R. Cookson[4], J. Hardy[2], M. Ryten[2], M.E. Weale[1], North American Brain Expression Consortium and UK Brain Expression Consortium.* 1) Medical and Molecular Genetics, King's College London, London, London SE1 9RT, United Kingdom; 2) Reta Lila Weston Institute and Department of Molecular Neuroscience, UCL Institute of Neurology, Queen Square, London WC1N 3BG, UK; 3) Department of Genetics, King Faisal Specialist Hospital and Research Centre, PO Box 3354, Riyadh 11211, Saudi Arabia; 4) Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA; 5) MRC Sudden Death Brain Bank Project, University of Edinburgh, Department of Neuropathology, Wilkie Building, Teviot Place, Edinburgh, EH8 9AG; 6) Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA 27599; 7) Department of Genetics, Department of Biostatistics, Department of Computer Science, University of North Carolina, Chapel Hill, NC, USA 27599.

BACKGROUND: Polymorphisms in the target mRNA sequences can greatly affect the binding affinity of microarray probe sequences, leading to false positive and false negative expression QTL signals (eQTLs) with any other polymorphisms in linkage disequilibrium. **METHOD:** We provide the most complete solution to date to this polymorphism-in-probe problem, by using SNPs and indels from the 1000 Genomes project (March 2012) and SNPs from the NHLBI Exome Sequencing Project to identify common polymorphisms (frequency >1% in Europeans) in probe sequences using two commonly-used microarray panels (the gene-based Illumina Human HT12 array and the exon-based Affymetrix Human Exon 1.0 ST array). We assessed the impact of probes containing polymorphisms in cerebellum and frontal cortex tissues of 438 neuropathologically normal individuals (curated by the UK and North American Brain Expression Consortia) on the number of declared cis-QTLs (defined as within 1Mb of transcription start site of the associated transcript). **RESULTS:** Only 5.3% of the probes in the Affymetrix Human Exon array (25-mer probe design) contained SNPs or indels but these accounted for 50–90% of the declared cis-eQTLs (depending on the significance threshold and tissue type). Similarly, 11.4% of the probes in the Illumina array (50-mer probe design) contained polymorphisms and these accounted for 20–52%. The strongest cis-eQTLs observed were those arising from probes with polymorphism, which implies that choosing a more stringent significance threshold would simply enrich for a higher proportion of false positives. About half of these false positives would not have been discovered if we had used a less complete variation panel (HapMap release 28). Finally, we confirm by conditional analysis on polymorphism-in-probe genotype that the majority (>93%) of these suspect cis-eQTLs are indeed false positives. This is true even if the polymorphism is not in high linkage disequilibrium with polymorphism in probe, so LD filtering should not be used to try and recover suspect cis-eQTLs. **CONCLUSIONS:** Researchers working on eQTLs should be aware of this important technical artefact and use the latest variation panel available to identify suspect eQTLs. LD filtering should not be used. Published eQTL studies with insufficient correction for this problem could contain a significant proportion of false positives, and this motivates a re-analysis of earlier studies using more stringent criteria.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

565F

PGC-1α regulates TFEB to promote autophagy-lysosome mediated turnover of aggregate-prone neurodegenerative disease proteins. *T. Tsunemi[1], B.E. Morrison[1], T.D. Ashe[1], A.R. La Spada[1,2,3]*. 1) Pediatrics, University of California, San Diego, La Jolla, CA; 2) Cellular & Molecular Medicine, University of California, San Diego, La Jolla, CA; 3) Neurosciences, University of California, San Diego, La Jolla, CA.

A number of neurodegenerative disorders result from misfolded proteins that accumulate in the CNS and yield neuron dysfunction, degeneration, and cell death. We recently found that the transcription factor PPARγ co-activator 1α (PGC-1α) promotes the turnover of mutant huntingtin (htt) protein in Huntington's disease, and determined that a normal autophagy pathway is required for PGC-1α to prevent htt protein aggregation and neurotoxicity. As a normally functioning autophagy pathway is critical for maintaining protein quality control in the CNS, we hypothesized that a potential link might exist between PGC-1α transactivation and enhanced autophagy-lysosome pathway activity. A transcription factor, known as TFEB, promotes the expression of genes in the autophagy pathway as well as genes required for lysosome biogenesis and function, identifying TFEB as a major node in the regulation of the entire autophagy-lysosome pathway. To evaluate the role of TFEB in PGC-1α-mediated htt protein turnover, we measured TFEB expression levels in HD in vitro and in vivo models, and found that polyQ-expanded htt repressed the TFEB gene expression - an effect that was rescued by PGC-1α. Chromatin immunoprecipitation analysis of the TFEB promoter localized PGC-1α occupancy to a proximal region, enabling us to derive a TFEB promoter-reporter construct. When we studied TFEB transactivation in ST-Hdh striatal-like cells with this promoter-reporter construct, we observed significant repression of TFEB promoter activity in Q111/Q111 homozygous cells, and again documented that PGC-1α expression could dramatically rescue this repression. We then tested if TFEB expression could prevent htt protein aggregation in Neuro2a cells expressing htt-104Q protein, and observed marked TFEB-dependent reductions in htt aggregate formation. To clarify the pathway relationship between PGC-1α and TFEB, we tested if PGC-1α required TFEB to limit htt aggregate formation or if TFEB required PGC-1α to limit htt aggregate formation. Importantly, we found that TFEB is capable of reducing htt aggregation when PGC-1α is knocked down, but that PGC-1α in the presence of TFEB knock-down no longer reduced htt aggregation. Our findings indicate that PGC-1α promotes htt protein turnover and aggregate suppression by co-activating the expression of TFEB, and places PGC-1α upstream of TFEB in the transcriptional regulation of the autophagy-lysosome pathway.

566W

The SLE associated TT>A polymorphism downstream of TNFAIP3 demonstrates lower affinity for NF-κB and SATB1 and reduced enhancer activity. *F. Wen[1], A.K. Templeton[2,3], S. Wang[1], J.M. Guthridge[1], M.B. Humphrey[3], P.M. Gaffney[1]*. 1) Arthritis $ Clinical Immunology, Oklahoma City Medical Research Foundation, Oklahoma City, OK; 2) Microbiology & Immunology, University of Oklahoma Health Science Center, Oklahoma City, OK; 3) College of Medicine, University of Oklahoma Health Science Center, Oklahoma City, OK.

Systemic lupus erythematosus (SLE) is a complex human autoimmune disease influenced by both genetic and environmental factors. The tumor necrosis factor alpha inducible protein 3 (TNFAIP3) gene encodes the ubiquitin-modifying enzyme A20, which negatively regulates NF-κB activity and genetic variants in the region of TNFAIP3 are associated with SLE in multiple populations. We recently described a functional TT>A polymorphic dinucleotide, 42kb downstream of the TNFAIP3 promoter that demonstrates reduced affinity for a nuclear protein complex that includes NF-κB. To further characterize the nuclear protein complex that engages the TT>A polymorphic region we used electrophoretic mobility shift assay supershift (EMSA-SS) with antibodies predicted to be in the complex. EMAS-SS was performed in both monocytoid derived (THP1) and B-lymphocyte derived (EBV transformed lymphocytes) cell lines under various stimulation conditions. EMSA-SS results were confirmed using chromatin immunoprecipitation followed by quantitative-PCR. To evaluate enhancer function we cloned a 340bp DNA sequence surrounding the TT>A polymorphism into a luciferase mini-reporter plasmid and measured luciferase expression in HEK293 and THP1 cells following stimulation. EMSA-SS with antibodies to NF-κB subunits determined that p50, p65, and cRel along with the protein, SATB1 were present in the complex. Binding patterns were also cell-type specific and differed between cells of monoctyoid and B lymphocyte origin. To confirm the EMSA-SS results with an orthogonal method, we performed ChIP with anti-p50 antibody, followed by qPCR with primers flanking the TT>A polymorphism. Our data demonstrated a statistically significant decrease ($p<0.05$) in the amount of protein bound to the TT>A variant region and confirmed the EMSA-SS results. The TT>A region demonstrated evidence of enhancer function with relative luciferase activity 5.1 fold higher than the empty vector for the reference construct (TT) while the polymorphic construct (-A) demonstrated a statistically significant ($p<0.001$) loss of enhancer activity (2.4 fold). In summary, these results demonstrate that the region containing the TT>A polymorphism functions as an enhancer may predispose to SLE through attenuation of recruitment of NF-κB subunits and SATB1 that modulate long-range transcriptional regulation of A20.

567F

Transcriptome and genome sequencing uncovers functional variation in human populations. *T. Lappalainen[1], M. Sammeth[2], N. Kurbatova[3], J. Monlong[2], M. Friedländer[4], M.A. Rivas[5], T. Wieland[6], O. Karlberg[7], M. Barann[8], P.A.C. 't Hoen[9], T. Griebel[2], M. Sultan[10], D. MacArthur[11], E. Lizano[2], T. Strom[6], S. Schreiber[8], H. Lerach[10], S.E. Antonarakis[1], G.J. van Ommen[9], R. Sudbrak[10], R. Häsler[8], A. Brazma[3], A.C. Syvänen[7], P. Rosenstiel[8], T. Meitinger[6], R. Guigo[2], I. Gut[2], X. Estivill[4], E.T. Dermitzakis[1], The Geuvadis Consortium.* 1) Department of Genetic Medicine and Development, University of Geneva Medical School, Geneva,, Switzerland; 2) Centro Nacional de Analisis Genomico, Barcelona, Spain; 3) European Bioinformatics Institute, Hinxton, UK; 4) Center for Genomic Regulation and UPF, Barcelona, Spain; 5) Wellcome Trust Centre for Human Genetics, Oxford, UK; 6) Institute of Human Genetics, Helmholtz Zentrum München, Munich, Germany; 7) Department of Medical Sciences, Uppsala University, Sweden; 8) Institute of Clinical Molecular Biology, University of Kiel, Germany; 9) Center for Human and Clinical Genetics, Leiden University Medical Center, The Netherlands; 10) Department of Vertebrate Genomics, Max Planck Institute for Molecular Genetics, Berlin, Germany; 11) Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA.

Functional interpretation of genetic variants discovered in human genome sequencing is essential to understand human phenotypic variation. We sequenced mRNA and small RNA of LCLs from 465 individuals from CEU, TSI, GBR, FIN and YRI populations of the 1000 Genomes samples, with a median 52M mapped reads per sample for mRNA and 1.5M for miRNA. Our 203 replicate samples allowed us to measure technical variation in high precision. In our analysis, we first characterized human transcriptome variation in unprecedented depth. We show how the total number of expressed genes increased from a median of 400 miRNA and 13.3K protein-coding genes in a single individual to a total of 1651 and 20.1K in the entire sample, with nearly every individual expressing some unique genes. We discovered almost 200 novel miRNA genes, unannotated splice junctions and novel transcriptionally active regions, many of which were rare in the population. In most genes, individual differences in splicing contributed more (63%) to the total variation than differences in total gene expression levels. Furthermore, the combination of high-quality transcriptome and genome sequencing data gave us the opportunity to characterize both common and rare regulatory variants. We discovered over 5000 genes with common eQTLs and hundreds with splicing QTLs (FDR < 10%), and show that indel variants are more likely to affect gene expression than SNPs ($p < 10^{-4}$). The enrichment of eQTL variants in regions with regulatory annotation ($p < 10^{-10}$) points to causal regulatory variants and sheds light on functional mechanisms underlying eQTL effects. We uncovered rare regulatory effects by allele-specific expression analysis, showing that they have higher effect sizes ($p = 10^{-15}$) and account for a large proportion of variation between individuals. Finally, we characterized transcriptome effects of hundreds of loss-of-function variants in our dataset. In addition to functionally validating variants predicted to cause e.g. nonsense-mediated decay or disrupt splice-sites, we developed models to improve functionality predictions of previously unseen variants base on their properties. Altogether, this study takes us beyond cataloguing putative functional variants towards understanding and predicting the cellular effects of variants in the human genome.

568W

Impact of private variants on human transcriptome architecture. *X. Li[1], K. Karczewski[2], K. Smith[1], K. Kukurba[2], S. Montgomery[1,2]*. 1) Pathology, Stanford University, Palo Alto, CA; 2) Genetics, Stanford University, Palo Alto, CA.

Recent explosion of human population size has created abundances of variants segregating only in small groups of people. Without the test of natural selection, these rare variants can potentially have stronger functional consequences and may underlie susceptibility of many complex diseases. However the actual effect sizes of these variants are still poorly understood due to their very low recurrence rate in a population. To systematically evaluate the functional effects of rare variants, we build a deep catalog of variants private to an extended three-generation family of 17 individuals from whole genome sequencing data and have linked them to transcriptome architecture. We use whole transcriptome (mRNA) sequencing to construct a comprehensive cellular landscape of transcriptional activities across multiple hierarchies including expression levels of an entire gene, abundances of different isoforms, novel transcripts and allele specific expression. We yield a regulatory map between different types of variants (SNPs and indels) to their cis-acting and trans-acting targets and quantify their enrichment in each functional category. We also conduct a comprehensive assessment of heritability of these effects through their transmission from parents to children. Estimation of degree of heritability and segregation patterns of rare variants on a functional level provides important insights towards a thorough understanding of their phenotypic implications in human populations.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

569F

Predicting disease-relevant *cis*-regulatory SNPs (*cis*-rSNPs) via tissue-specificity of allelic expression and chromatin state. *N. Light[1], V. Adoue[1], P. Lundmark[2], B. Ge[1], T. Kwan[1], A. Syvänen[2], T. Pastinen[1].* 1) Human Genetics, McGill University, Montreal, QC, Canada; 2) Medical Sciences, Uppsala University, Uppsala, Sweden.

Over 80% of complex disease-associated variants, discovered by GWAS, are located in non-coding regulatory regions of the genome. Regulatory variants have largely been identified through eQTL mapping; however, direct assessment of *cis*-regulatory SNPs (*cis*-rSNPs), which affect up to 30% of RefSeq transcripts, requires an allele-specific (AS) approach. *cis*-rSNPs explain more than 50% of population variance in allelic expression (AE) in LCLs, with up to 70% of *cis*-rSNPs being shared between tissues. Recent mapping of human epigenomic data in disease-relevant cell-types, by the ENCODE and Roadmap Epigenomics Project consortiums, has allowed for chromatin context-dependent analyses of *cis*-regulatory variation. Categorizing *cis*-rSNPs, mapped in accessible cell-types, by tissue-specificity and chromatin state will be essential for optimally predicting their disease-relevance. In order to characterize *cis*-rSNP by cell-type specificity, we mapped *cis*-rSNPs in 70 fibroblast (FB) cell lines, 188 purified monocyte (MNC) samples, and 118 HapMap LCLs (55 CEU, 63 YRI), using phased genotypes and AE differences measured on Illumina 1M/2.5M BeadChips. To investigate chromatin state, we generated AS-ChIP data for 5 known regulatory-associated histone marks, including enhancer-associated H3k4me1 (me1) and active promoter-associated H3k4me3 (me3), in fibroblast and LCL trios, using Illumina HiSeq in addition to 1M/2.5M/5M BeadChips. Differentiating *cis*-rSNPs by tissue-specificity of AE association and chromatin state, reveals a 3–4 fold enrichment of MNC-specific *cis*-rSNPs in MNC-specific me1 or me3 peaks, which is not observed in shared peaks. Transcripts exhibiting similar RNA-seq expression across LCL, MNC and FB samples, show significant enrichment for shared associations. Filtering *cis*-rSNPs mapped in LCLs by shared overlap with a shared MNC chromatin state, similarity of LCL and MNC expression and/or overlap with transcription factor binding sites, significantly enriched for SNPs with stronger association to AE in FB. Intersecting AE-mapped *cis*-rSNPs with GWAS catalog SNPs shows a large enrichment of immune-related diseases/traits for LCL and MNC-specific *cis*-rSNPs. In contrast, FB-specific associations showed a depletion of immune system-related GWAS hits. Our data suggests leveraging functional genomic data from disease-relevant tissues with AE-mapped *cis*-rSNPs from accessible cell-lines, may be a potentially powerful tool for predicting disease-relevant variants.

570W

Characterization of the basic regulatory elements involved in the regulation of OTC gene expression in human. *O. Luksan[1], L. Dvorakova[2], M. Jirsa[1].* 1) Institute for Clinical and Experimental Medicine, Praha, Czech Republic; 2) Institute of Inherited Metabolic Disorders, First Faculty of Medicine, Charles University in Prague, and General University Hospital in Prague.

Ornithine carbamoyltransferase deficiency (OTCD; OMIM 311250), X-linked disease caused by the lack of enzyme catalyzing the synthesis of citrulline, is the most common inherited defect of the urea cycle. Clinical diagnosis based on the mutation analysis of *OTC* coding regions fails in 20–25% of patients. In such cases, the disease may be caused by gross gene rearrangements or defects in regulatory regions of the gene. We previously reported the pathogenic effect of a single nucleotide substitution c.-366A>G found in the promoter of a patient manifesting OTCD. We also demonstrated that the liver-specific regulation of OTC expression is dependent on the interaction of a 793 bp promoter with a strong enhancer located 9 kb upstream the translation start side (TSS). The aim of our recent study is a detailed characterization of the promoter and enhancer regions. Using the Dual Luciferase Repoter Assay (Promega, Madison, Wi) with pGL3 constructs containing enhancer sequence and promoter variants truncated from the 5`-end, we specified the core promoter region. Promoter core, delimited from the 3'-end by the transcription start site located 95 bp upstream the initiation codon, is formed by a sequence of 173 base-pairs containing three positive and a negative *cis*-acting element. The Cis1+ positive regulatory element was found in a short sequence of 35 bp located close to the transcription origin and reaching 80% of the activity of complete promoter. Another regulatory elements Cis2+, Cis3- and Cis4+ were localized within the positions 128–179 bp, 180–219 bp and 220 –269 bp upstream the TSS. We set up an innovated DNase I footprinting method based on the capillary electrophoresis of fluorescently labeled fragments after DNase I cleavage of investigated DNA. The comparative analysis of electropherograms obtained for DNA samples incubated with Hep-G2 nuclear extract and controls (without nuclear extract) revealed four protected regions corresponding to Cis1+ and Cis2+ elements. Two of them showed a strong similarity with the sequence motifs recognized by HNF-4 transcription factor, which was previously reported to have a key role in the regulation of *OTC* expression in the rat. No protected regions were observed within the sequences encompassing Cis3- and Cis4+ elements so far. These regions are currently being studied in a set of gel shift assays. Financial support: GA UK 251168_99910.

571F

Regulatory impact of alternative polyadenylation in human B-lymphoblastoid cells. *O.K. Yoon, T.Y. Hsu, J.H. Im, R.B. Brem.* Molecular and Cell Biology, University of California, Berkeley, CA.

Gene expression varies widely between individuals of a population, and regulatory change can underlie phenotypes of evolutionary and biomedical relevance. A key question in the field is how DNA sequence variants impact gene expression, with most mechanistic studies to date focused on the effects of genetic change on regulatory regions upstream of protein-coding sequence. By contrast, the role of 3' RNA processing in regulatory variation remains largely unknown, owing in part to the challenge of identifying functional elements in 3' untranslated regions. In this work, we conducted a genomic survey of transcript ends in lymphoblastoid cells from genetically distinct human individuals. Our analysis mapped the cis-regulatory architecture of 3' gene ends, finding that transcript end positions did not fall randomly in untranslated regions, but rather preferentially flanked the locations of 3' regulatory elements, including miRNA sites. The usage of these transcript length forms and motifs varied across human individuals, and polymorphisms in polyadenylation signals and other 3' motifs were significant predictors of expression levels of the genes in which they lay. Independent single-gene experiments confirmed the effects of polyadenylation variants on steady-state expression of their respective genes, and validated the regulatory function of 3' cis-regulatory sequence elements that mediated expression of these distinct RNA length forms. Focusing on the immune regulator IRF5, we established the effect of natural variation in 3' RNA processing on regulatory response to antigen stimulation. Our results underscore the importance of two mechanisms at play in the genetics of 3'-end variation: the usage of distinct 3'-processing signals and the effects of 3' sequence elements that determine transcript fate. Our findings suggest that the strategy of integrating observed 3'-end positions with inferred 3' regulatory motifs will prove to be a critical tool in continued efforts to interpret human genome variation.

572W

Identification of a novel CFTR mRNA isoform in nasal epithelial cells. *A. Hinzpeter[1,2], A. de Becdelièvre[1,3], E. Bieth[4], F. Brémont[5], C. Costa[1,3], C. Gameiro[3], A. Aissat[1,2], M. Goossens[1,2,3], P. Fanen[1,2,3], E. Girodon[1,3].* 1) U955, INSERM, Créteil, France; 2) Université Paris-Est, Créteil, France; 3) GH Henri Mondor, Department of Genetics, APHP, Créteil, France; 4) Purpan Hospital, Department of Genetics, Toulouse, France; 5) Pediatric CF Center, Department of Pediatric Pulmonology, Toulouse, France.

Nasal polyposis (NP) starting early in life and/or permanently infected is uncommon. In children, it leads to suspect cystic fibrosis, a recessive disorder caused by mutations in the *CFTR* gene and characterized by a large phenotypic spectrum. In atypical phenotypes, extensive analysis at the gDNA level of the *CFTR* coding regions occasionally fails to identify mutations, which may lie in deep intronic or regulatory regions and affect the normal splicing or transcriptional regulation.

We studied a 12y-old child with isolated NP, no lung or digestive disease, and a negative sweat test. He was found to be heterozygous for c.2551C>T (p.Arg851*), a nonsense mutation inherited from his father. Analysis at the mRNA level performed in the child and his parents showed qualitative and quantitative abnormalities. A partial skipping of exon 15 was observed in the patient and his father, associated with the c.2551C>T mutation. Quantitative analysis (TaqMan gene expression assay) showed a significant increase of about 30% of total CFTR mRNA in the patient and his mother, as compared to 3 healthy controls. Search for a regulatory mutation by sequencing led to identify the c.-2954G>A variant in the *CFTR* promoter region. This variant, previously reported in dbSNP 135 with a low frequency in the European population (0.34%), increases *CFTR* promoter activity as assessed by luciferase assays. This effect could be related to the total CFTR mRNA increase observed in the patient. However, this mRNA variant could be non functional. Indeed, 5'RACE analysis of the proband's mRNA showed the presence of a transcript in which the canonical exon 1 was replaced by the alternative exon 1a. Two mRNA isoforms including alternative exons have been previously described in the developing fetal lung: -1a,1a,2to27 and -1a,2to27, which were not predicted to encode a fully functional channel but are hypothesized to have regulatory functions. We describe here for the first time *in vivo* a third isoform 1a,2, which expression could be favored by the c.2954G>A variant.

CFTR gene regulation and its impact on the lung development is far from being elucidated. Mutations affecting regulation can explain atypical forms of cystic fibrosis or CFTR-RDs. The search for such mutations could be of particular interest in NP, and contribute to understand the physiopathology of NP in the cystic fibrosis context.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

573F

Cytokine induced genetic-epigenetic interactions at autoimmune thyroiditis susceptibility gene variants. *M. Stefan[1,3], W. Zang[1], M. Keddache[2], Y. Tomer[1,3]*. 1) Mount Sinai School of Medicine, New York, NY; 2) University of Cincinnati, Cincinnati, OH; 3) James J Peters VA Medical Center, Bronx, NY.

Autoimmune thyroid diseases (AITD) result from interactions between genetic and environmental factors. Several AITD susceptibility loci have been identified through genome-wide association studies (GWAS), but the causal variants remain undefined. Recent evidence suggests that epigenetic mechanisms modulate the complex interplay between genes and intra- and extra-cellular factors to trigger pathological autoimmune responses. Since cytokines are key mediators of tissue inflammation and infiltration we tested the hypothesis that inflammatory cytokines promote thyroid cell dysfunction through epigenetic modifications of AITD-genes. One prime cytokine in the etiology of AITD is interferon alpha (IFNa), which has also been shown to precipitate AITD when used as therapeutic agent in the context of infection. We have previously shown that IFNa increases mRNA expression of major AITD susceptibility genes in both cell lines and a transgenic mouse model of IFNa thyroid over-expression. We now mapped modifications of histone patterns [histone H3 mono- and tri-methylated at Lys-4, (H3K4me1 and H3K4me3)] induced by IFNa at these loci using ChIP-seq in human thyroid cells. Integration of the ChIP-seq with global gene expression data showed that significantly up-regulated pathways included genes characterized by H3K4me3 enrichment in the 5'-regions, demonstrating a correlation between H3K4me3 and pathway activation by IFNa. Most upregulated genes/pathways participate in innate immunity and host defense response. IFNa induced enrichment of H3K4me1 in intronic and 5' gene regions. We used the potential of H3K4me1 to mark regulatory regions to identify functional AITD-associated single nucleotide polymorphisms (SNPs). H3K4me1 peaks mapped at two SNPs in thyroglobulin (TG) and thyroid stimulating hormone receptor (TSHR) genes. Interestingly, the TG SNP was previously shown by us to map within a transcription factor binding site for interferon regulatory factor-1 that modulate the TG promoter in an allele-dependent manner. Using ChIP, luciferase reporter and siRNA assays we show now that the intronic TSHR SNP delineate a regulatory element that decreases promoter's activity. Our results demonstrate that an un-biased genome-wide epigenomic screening can pinpoint the disease-associated variants and their mechanism.

574W

DMD transcript imbalance regulates dystrophin levels. *P. Spitali[1], J.C. van den Bergen[2], B. Wokke[2], A.A.M Janson[3], R. van den Eijnde[3], J.J.G.M Verschuuren[2], J.T den Dunnen[1], P.A.A 't Hoen[1], A. Aartsma-Rus[1]*. 1) Human Genetics, Leiden University Medical Center, Leiden, Zuid-Holland, Netherlands; 2) Neurology, Leiden University Medical Center, Leiden, Zuid-Holland, Netherlands; 3) Prosensa Therapeutics.

Duchenne and Becker muscular dystrophies (DMD and BMD) are caused by respectively out-of-frame and in-frame mutations in the dystrophin encoding *DMD* gene. Molecular therapies targeting either the pre-mRNA, such as antisense oligonucleotides (AONs), or the mRNA, such as read-through of stop codons, are in clinical trials and show promising results. The effect of these approaches will depend on the stability and expression levels of the DMD mRNA in skeletal muscles and heart. We report that the DMD transcript is more abundant in heart compared to skeletal muscles, in mouse and man. We show that the transcript shows a significant 5'-3' imbalance in wild type mice, which is more pronounced in the *mdx* mouse model carrying a nonsense mutation. Reading frame restoration via antisense-mediated exon skipping did not correct this imbalance. Finally, we demonstrate that transcript balance, but not transcriptional rate is an important determinant for dystrophin protein levels in 22 Becker patients. We suggest that the availability of the entire transcript is a key factor to determine protein abundance and this may influence the outcome of mRNA targeting therapies.

575F

Protein inhibitor of activated STAT3 (PIAS3) Modulates the Activity of Anterior Segment Dysgenesis Transcription Factors in Trabecular Meshwork Cells. *T. Footz, L. Huang, M. Walter.* Dept. Medical Genetics, University of Alberta, Edmonton, Alberta, Canada.

Development of the ocular anterior segment is controlled by complex gene regulatory pathways that lead to severe functional and dysmorphic defects when perturbed by gene mutation. Recently, an evolutionarily-conserved network of physically-interacting transcription factors has been described which includes PAWR, PITX2, FOXC1 and FOXC2. Mutations of the latter three genes cause a variety of anterior segment phenotypes, with increased risk of developing glaucoma, a typically late-onset progressive neurodegeneration that is one of the principle causes of blindness worldwide. A yeast two-hybrid screen was performed to discover additional interacting proteins of the established regulatory network that are expressed in trabecular meshwork cells of the human anterior segment. The protein inhibitor of activated STAT3 (PIAS3) was identified as a FOXC2-interacting protein. PIAS3 is a regulator protein that affects the function of transcription factors involved in cell proliferation and apoptosis. Molecular investigations (expression analyses, reporter assays, immunofluorescence microscopy and DNA mobility shift assays) were performed to reveal the functional consequences of interactions of the anterior segment gene network members with PIAS3. In a manner similar to PAWR, overexpression of PIAS3 also serves to affect the ability of FOXC1, FOXC2 and PITX2 to transcriptionally activate target genes. Our analyses reveal that key protein interactions underlie the complex regulatory network necessary for human eye development. Disruptions to this network have profound consequences that lead to glaucoma.

576W

Statin induced alternative splicing in HepG2 cells. *C. Stormo[1], M.K. Kringen[2], O.K. Olstad[1], J.P. Berg[1,3], A.P Piehler[4]*. 1) Dept of Medical Biochemistry, Oslo University hospital, Ullevål, Norway; 2) Dept of Pharmacology, Oslo University Hospital, Ullevål, Norway; 3) Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Norway; 4) Fürst Medical Laboratory, Oslo, Norway.

Background: Genomic diversity is greatly expanded by the process of alternative splicing, where introns are removed and exons ligated to form various mature mRNA transcripts from a single gene. Recently, alternative splicing was suggested as a mechanism underlying cholesterol homeostasis and regulation of 3-Hydroxy-3-methylglutaryl-coenzyme A reductase (HMGCR), the pharmacologic target of lipid lowering drugs (statins). In this study, we sought to comprehensively investigate alternative splicing mediated by atorvastatin and to potentially highlight novel pathways related to cholesterol homeostasis and statin treatment. Methods: HepG2 liver cells were treated with or without 10 μM atorvastatin for 24 hours. cDNA was hybridized to Affymetrix GeneChip Human Exon 1.0 ST Array. Analysis of gene expression and alternative splicing was done by Partek Genomic Suite software. Only “core” RefSeq genes were considered. Potential alternative splicing events were validated by PCR and gel electrophoresis, and RT-qPCR with SYBR Green dye. Results: We found 142 genes ($p<0.01$) of the 17,151 “core” Refseq genes to have different splicing pattern in statin-treated HepG2 cells compared to non-treated cells. Eight genes were selected for validation. RT-PCR analysis, however, could only confirm one gene (FADS2) with two alternative splicing events; a FADS2 splice variant with an alternative exon 1 (FADS2-exons_1b-12) and a truncated transcript with an extended exon 10 variant (FADS2-exons_1a-10), corresponding to GenBank ID AF108658 and BC009011, respectively. The upregulation (ranging from 1.2-to 2.1 fold change) of the canonical FADS2 variant (FADS-exon1a_12) and the two alternatively spliced variants after atorvastatin treatment was confirmed in both HepG2 liver cells and a human lymphoblastoid cell line by RT-qPCR. Discussion/Conclusion: The human FADS2 gene, encoding delta-6-saturase, is clustered with family members FADS1 and FADS3. Genetic variation in the FADS gene cluster has shown to affect human polyunsaturated fatty acids (PUFA) and long-chain PUFA blood levels which are related to various disease processes, including cardiovascular disease. In this study, we have identified two alternatively spliced FADS2 variants that are responsive to atorvastatin treatment in vitro. These preliminary results may indicate a role of alternative splicing of FADS2 in a statin treatment regimen.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

577F

Functional RNA Editing Sites Identified by Comparative Transcriptome Analysis. *R. Zhang[1], G. Ramaswami[1], L.P. Keegan[2], T. Deng[1,3], M.A. O'Connell[2], J.B. Li[1].* 1) Department of Genetics, Stanford University, Stanford, California, USA; 2) MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, Western General Hospital, Edinburgh, UK; 3) Department of Structural Biology, Stanford University, Stanford, California, USA.

Adenosine to inosine (A-to-I) RNA editing is required for normal brain function and early embryonic development. However, this important phenomenon is poorly explored, partially due to the lack of identification of functionally significant RNA editing sites. Recently, RNA editing identification has been facilitated by the next generation sequencing technologies, by delineating the sequence differences between matched RNA and genome sequencing from the same individual. Here we introduce a cross-species transcriptome comparison method to accurately identify functionally conserved RNA editing events from RNA sequencing (RNA-seq) data alone. We applied this method to Drosophila and primate lineages. In Drosophila species, ~95% of the conserved RNA variants are A-to-G mismatches, indicative of A-to-I editing. We identified 1,329 exonic A-to-I editing sites in Drosophila melanogaster, 849 of which have not been previously reported. In the primate lineage, we identified 22,748 and 668 conserved exonic A-to-I sites in Alu and non-Alu regions, respectively, including 295 sites that lead to amino acid changes. Our findings demonstrate the power of comparative transcriptomics to identify functionally significant RNA editing sites.

578W

Meta-analysis of Transcriptome Data using Pathway-based Approach Reveals Biologically Relevant Asthma Genes. *T. Mersha, S. Amirisetty, G. Hershey.* pediatrics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH.

Asthma is a disease of chronic airway inflammation that affects over 300 million individuals worldwide including 20 million in the U.S. While the causes of asthma are not fully understood, genetic differences may be partly responsible for the variations in asthma susceptibility. Using microarray experiments, many independent groups have identified differentially expressed genes (DEGs) between healthy versus asthmatic patients. However, there is little overlap in the DEG lists identified among different experiments due to variation from random noise, biological and experimental differences, and differences in the extraction and handling of RNA samples. Therefore, a statistics-based meta-analysis is necessary to identify a set of genes that are consistently dysregulated among multiple independent microarray studies. A recent search of asthma-related expression data in the GEO dataset yielded 1379 experiments of which 1190 were human specific (http://www.ncbi.nlm.nih.gov). In this study, we focused our meta-analysis and network/pathway analysis on the 5 microarray human data sets, which had individual experiment sample sizes of over 100. Of the 5534 genes analyzed, 550 (12%) were up-regulated and 456 (8%) were down-regulated. Many key genes including SERPINB2, CLCA1 and P2RY14 were up-regulated and MUC5AC/MUC5B mucins and C6 were down-regulated. These genes were involved in inflammatory diseases, mucus over-production, airway obstructions, immunological diseases, and hypersensitivity responses. Meta-analytic dataset and network/pathway analysis of publicly available databases provide a starting point to identify biologically plausible potential candidate genes linked to asthma and to explore the mechanistic basis for the observed expression patterns in asthma.

579F

Whole transcriptome analysis of Velocardiofacial Syndrome (VCFS) using next-generation sequencing. *Y. Zhang[1], J. Jin[2], X. Zhu[3], M. Haney[3], Y. Kluger[2], S. Weissman[1], A. Urban[3].* 1) Genetics, Yale Univ School of Medicine, New Haven, CT; 2) 2. Yale University School of Medicine, Department of Pathology, Biological and Biomedical Sciences, 300 George Street, New Haven, CT; 3) Stanford University Department of Psychiatry and Behavioral Sciences and Department of Genetics, 1050A Arastradero Road, Stanford, CAWhitney.

Velocardiofacial Syndrome (VCFS) is a relatively common genomic disorder, typically caused by the heterozygous deletion of 3 million basepairs of genomic DNA sequence on chromosome 22q11. There are various developmental defects including neurodevelopmental disorders such as early onset schizophrenia, autism spectrum disorders (ASD) and learning disabilities; however, the combinations and severity of symptoms vary widely between patients with no clear genetic basis to such variability, and for most symptoms there are several plausible potential candidate genes amongst the at least 50 genes within the typical deletion boundaries, but none that have been identified yet as the clear and sole cause of any given symptom. Here we report on RNA-Sequencing analyses of VCFS using 2nd-generation sequencing (Illumina HiSeq 2000). For 5 lymphoblastoid cell lines from VCFS patients as well as 9 HapMap control cell lines, we carried out very deep RNA-Seq analysis (up to 100 million mapped 2×100bp paired-end constructs per cell line transcriptome). We performed cluster analysis for the 200 genes most differentially expressed between patients and controls, with the results showing a clear separation of the patient group from the control group. Among those 200 genes, 56 genes were down-regulated and 144 genes were up-regulated in VCFS samples. There are 38 genes within the 3 Mbp deletion region of VCFS on chromosome 22q11 that are expressed above a threshold of 0.5 FPKM and of those 10 genes are significantly down regulated ($P < 0.05$, ratios $< = 0.5$) in patients versus controls. The gene expression patterns also show evidence of gene dosage compensation for some genes, as well as under- or over- compensation for others. The mechanisms controlling such effects will have to be studied in detail on the molecular level in follow-up work.

580W

Genome-wide search for novel human uORFs and N-terminal protein extensions using ribosomal footprinting. *J. Hampe[1], C. Fritsch[1], A. Herrmann[1], M. Nothnagel[2], K. Szafranski[3], K. Huse[3], S. Schreiber[1], M. Platzer[3], M. Krawczak[2], M. Brosch[1].* 1) Department of Internal Medicine I, University-Hospital Schleswig Holstein, Kiel, Germany; 2) Institute for Medical Informatics and Statistics, University-Hospital Schleswig Holstein, Kiel, Germany; 3) Genome Analysis Group, Fritz Lipmann Institute for Age Research, Jena, Germany.

So far, the annotation of translation initiation sites (TISs) has been based mostly upon bioinformatics rather than experimental evidence. We adapted ribosomal footprinting to puromycin-treated cells to generate a transcriptome-wide map of TISs in a human monocytic cell line. A neural network was trained on the ribosomal footprints at previously annotated AUG translation initiation codons (TICs), and used for the ab initio prediction of TISs in 5062 transcripts with sufficient sequence coverage. Functional interpretation suggested 2994 novel upstream open reading frames (uORFs) in the 5' UTR (924 AUG, 2070 near-cognate codons), 1406 uORFs overlapping with the coding sequence (116 AUG, 1290 near-cognate) and 546 N-terminal protein extensions (6 AUG, 540 near-cognate). The TIS detection method was validated on the basis of previously published alternative TISs and uORFs. On average, TICs in newly annotated TISs were significantly more conserved among primates than control codons, both for AUGs ($p<10^{-10}$) and near-cognate codons ($p=5.9\times10^{-7}$). The derived transcriptome-wide map of TISs will help to explain how human proteome diversity is influenced by alternative translation initiation and regulation.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

581F

Variation in protein levels and its genetic determination in human lymphoblastoid cell lines. *L. Wu, S.I. Candille, Y. Choi, H. Tang, M. Snyder.* Genetics Dept, Stanford University, Stanford, CA.

Gene expression varies among individuals and among populations. Correlations between cis-acting genetic variants and RNA levels have been demonstrated in many studies of individual genes and genomes. Less well-characterized are the determinants of variation in protein levels. Quantitative correlation between levels of mRNA and corresponding protein are modest, indicating that at least some mechanisms of gene regulation might influence protein levels directly. Proteome-wide measurements of protein levels are required to understand how genes are regulated by mechanisms that do not act directly through modulating mRNA levels. Here we use isobaric tandem mass tag (TMT) quantitative mass spectrometry to determine relative protein levels in lymphoblastoid cell lines from 95 individuals genotyped in the HapMap Project: 53 of European ancestry, 33 of Nigerian ancestry, and 9 of East Asian ancestry, including 10 European trios and 9 African trios. A total of 12,089 proteins were quantified based on unique peptide ratios. We characterized protein variation between sexes and among individuals and populations, and found evidence for the heritability of protein levels. Using genetic association analyses between protein levels and SNPs (single nucleotide polymorphisms), we identified a number of cis pQTLs (protein quantitative trait loci), including variants not detected by previous transcriptome studies. We also measured protein covariation among individuals to detect networks of coexpressed proteins. Our results show that quantitative mass spectrometry-based proteomic technology can detect proteome differences among individuals, reveal genetic variants that influence protein levels and contribute to the exploration of the role of co-expressed proteins in cellular processes.

582W

The impact of loss of function variants on the transcriptome. *M.A. Rivas[1], T. Lappalainen[2], M. Sammeth[3], D.G. MacArthur[4], N. Kurbatova[5], T. Wieland[6], M. Lek[4], T. Griebel[3], M. Barann[7], M. Sultan[8], J. Carlsson[9], P. Hoen[10], T. Strom[6], S. Schreiber[7], H. Lehrach[8], S. Antonarakis[2], G. van Ommen[10], R. Sudbrak[8], R. Hasler[7], A. Brazma[5], A. Syvanen[9], P. Rosenstiel[7], T. Meitinger[6], R. Guigo[3], I. Gut[3], X. Estivilli[11], M.I. McCarthy[1], E.T. Dermitzakis[2], The Geuvadis Consortium.* 1) Nuffield Department of Clinical Medicine, University of Oxford, Oxford, United Kingdom; 2) Department of Genetic Medicine and Development, University of Geneva; 3) Centro Nacional de Analisis Genomico, Barcelona; 4) Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston; 5) European Bioinformatics Institute, Hinxton; 6) Institute of Human Genetics, Helmholtz Zentrum München, Munich; 7) Institute of Clinical Molecular Biology, University of Kiel; 8) Department of Vertebrate Genomics, Max Planck Institute for Molecular Genetics, Berlin; 9) Department of Medical Sciences, Uppsala University; 10) Center for Human and Clinical Genetics, Leiden University Medical Center; 11) Center for Genomic Regulation and UPF, Barcelona.

Whole genome and exome sequencing studies are rapidly driving exploration of the genetic architecture of human disease and related phenotypic traits. However, with these advancements come challenges in the interpretation of the functional effects of genetic variation that will be essential to our ability to map genes to disease and related phenotypes. Variants that create premature termination codons or PTC, commonly referred to as loss of function, are most likely to have phenotypic effects. Despite their predicted effect on protein structure the cellular consequences are not well characterized. We present a systematic analysis of loss of function variants and their impact on the transcriptome by combining DNA sequence and sequenced mRNA of LCLs from 465 individuals from the 1000 Genomes Project.

In our analysis we focus on genetic variants with in silico annotation predictions of loss of function. Allele-specific expression analysis indicates widespread nonsense-mediated mRNA decay (NMD). We subdivided the variants according to whether they were predicted to trigger NMD according to the “50 nt rule”. Allelic ratios for the mutations that are predicted to escape or trigger NMD differ significantly (P=2e-7), suggesting that NMD generally requires an exon junction >50nt downstream of the PTC mutation. Moreover, 40% of variants predicted to trigger NMD had no deviation in allelic ratios (>0.4). Thus, we are developing models to understand why these variants escape NMD. Furthermore, we investigated whether the loss of mRNA molecules by NMD could be compensated by higher expression of the wild-type allele to maintain normal gene expression levels; the individuals affected by NMD showed suggestively lower expression levels (measured by RPKM (P=.08)). Interestingly, 38% of individuals who had PTC and evidence of ASE had expression levels higher than the median of individuals who did not have the PTC. This is surprising and suggests that there is some mechanism to compensate the decay. Additionally, we characterized effects of splice-site variants, including exon skipping, intron retention, and subsequent NMD. Our analysis implies that even variants with most severe predicted effects need functional validation and modeling to enable better application to disease studies. This study offers the promise of unprecedented understanding of the impact of loss of function variants on the transcriptome.

583F

Mapping cis and trans regulatory effects across multiple tissues in twins: the MuTHER Study. *K. Small[1], E. Grundberg[1,2], A. Hedman[3], A. Nica[4], A. Buil[4], S. Keildson[3], J. Bell[1,3], T. Yang[2], A. Barrett[3], J. Nisbet[2], D. Meduri[2], S-Y. Shin[2], D. Glass[1], J. Min[3], L. Parts[2], L. Tsaprouni[2], F. Nestle[1], S. O'Rahilly[5,6], N. Soranzo[2], C. Lindgren[3], K. Zondervan[3], K. Ahmadi[1], E. Schadt[7], K. Stefansson[8], G. Davey Smith[9], M. McCarthy[3,10], P. Deloukas[2], E. Dermitzakis[4], T. Spector[1], MuTHER Consortium.* 1) Kings College London, London, United Kingdom; 2) Wellcome Trust Sanger Institute, Hinxton, United Kingdom; 3) Wellcome Trust Centre for Human Genetics, Oxford, United KIngdom; 4) University of Geneva Medical School, Geneva, Switzerland; 5) University of Cambridge, Cambridge, United Kingdom; 6) Cambridge NIHR Biomedical Research Centre, Cambridge, United Kingdom; 7) Mount Sinai School of Medicine, New York, USA; 8) deCODE genetics, Reykjavik, Iceland; 9) University of Bristol, Bristol, United Kingdom; 10) Oxford Centre for Diabetes, Endocrinology & Metabolism, University of Oxford, Oxford, United Kingdom.

Sequence-based variation in gene expression is a key driver of disease risk. Common variants regulating expression in cis have been mapped in many eQTL studies typically in single tissues from unrelated individuals. We have undertaken a comprehensive analysis of gene expression across multiple tissues (subcutaneous fat, whole skin and lymphoblastoid cell lines) sampled from 856 female twins from the MuTHER (Multiple Tissue Human Expression Resource) Study. This large set of mono- and dizygotic twins allows systematic dissection of genetic(cis and trans) and non-genetic effects on gene expression. We demonstrate that cis-regulatory effects account for ~30–36% of heritability of gene expression, but up to 40% of cis-heritability, or 12% of total heritability, is unaccounted for by common cis-variants (ie those identified in genotypes imputed to the HapMap2 reference panel). This indicates that 1) a substantial amount of cis-heritability is driven by rare or low frequency regulatory variants or non-additive effects that remain to be discovered and 2) the majority (~60%) of regulatory variation acts in trans. The missing cis-heritability has intriguing implications for genome-wide association (GWA) signals in cases where the effect of the lead SNP is mediated via a cis-eQTL, i.e. if a GWA variant is an eQTL and affects disease risk by modulating expression of a gene, any additional rare variant modulating expression of the same gene in the same tissue should also affect the same trait. This leads us to predict that on average an additional 40% of signal remains to be discovered at cis-eQTL GWAS loci (of which we identify 358 in this study). We further identify several replicating trans-variants which act predominantly in a tissue-restricted manner and may regulate the transcription of many genes.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

584W

Identifying age- and sex- associated gene expression profiles in >7,000 whole-blood samples. *M.J. Peters[1,2,17], R. Joehanes[3,17], T. Esko[4,17], K. Heim[5,17], H. Völzke[6,17], L. Pilling[7,17], J. Brody[8,17], Y.F. Ramos[9,17], B.E. Stranger[10,11], M.W. Christiansen[8], S. Gharib[8], R. Hanson[12], A. Hofman[2,13], J. Kettunen[14], D. Levy[3], P. Munson[3], C. O'Donnell[3], B. Psaty[8], F. Rivadeneira[1,2,13], A. Suchy-Dicey[8], A.G. Uitterlinden[1,2,13], H. Westra[15], I. Meulenbelt[2,9,17], D. Enquobahrie[8,17], T. Frayling[7,17], A. Teumer[16,17], H. Prokisch[5,17], A. Metspalu[4,17], J.B.J. Van Meurs[1,2,17], A.D. Johnson[3,17] on behalf of the CHARGE Gene Expression Working Group*. 1) Department of Internal Medicine, Erasmus Medical Centre, Rotterdam, the Netherlands; 2) Netherlands Genomics Initiative-Sponsored by the Netherlands Consortium for Healthy Aging, Rotterdam and Leiden, the Netherlands; 3) Framingham Heart Study, National Heart, Lung and Blood Institute, Framingham, USA; 4) Estonian Genome Center and Institute of Molecular and Cell Biology of University of Tartu, Estonia; 5) Institute of Human Genetics, Technische Universität München, Munich, Germany; 6) Institute for Community Medicine, University Medicine Greifswald, Germany; 7) Epidemiology and Public Health, Peninsula College of Medicine and Dentistry, University of Exeter, UK; 8) Cardiovascular Health Research Unit, Departments of Medicine and Epidemiology, University of Washington, Seattle, WA, United States; 9) Department of Molecular Epidemiology, Leiden University Medical Center, Leiden, the Netherlands; 10) Division of Genetics, Department of Medicine, Brigham and Women's Hospital, Boston, MA, USA; 11) Broad Institute of Harvard and MIT, Cambridge, USA; 12) Phoenix Epidemiology and Clinical Research Branch, NIDDK, National Institute of Health, Phoenix, AZ, USA; 13) Department of Epidemiology, Erasmus Medical Centre, Rotterdam, the Netherlands; 14) Institute for Molecular Medicine Finland FIMM, University of Helsinki, Finland; 15) Department of Genetics, University of Groningen, University Medical Center Groningen, the Netherlands; 16) Interfaculty Institute for Genetics and Functional Genomics, Ernst-Moritz-Arndt University Greifswald, Germany; 17) Contributed Equally.

Genome-Wide Expression Profiles (GWEPs) have been assayed in a growing number of cohort studies, but few attempts have been made to meta-analyse and cross-validate expression datasets. Consequently, many expression studies have been under powered. Therefore, we established a large-scale multi-cohort GWEP meta-analysis. The aim of this study was to robustly identify novel gene expression signatures associated with age and sex, two major risk factors for many diseases. We analyzed 6,993 European-ancestry PAXgene (whole-blood) samples from 6 cohort studies (RS, FHS, EGCUT, KORA, SHIP, INCHIANTI). GWEPs were quantile-normalized, log2-transformed, probe-centered and sample-z-transformed prior to analysis. In the discovery stage we meta-analysed age- and sex-associated signals for samples hybridized to an Illumina or Affymetrix array separately. All analyses were adjusted for plate ID, RNA quality, fasting- and smoking status, and cell counts (when available). The age analysis was additionally adjusted for sex. All significant signals were cross-validated between the Illumina and Affymetrix platforms. We examined the top-associated GWEPs in 3 additional studies: HVH (n=348), GARP (n=134), and NIDDK/PIMA (n=1457). We identified 396 age-associated transcripts with p<1E-5 and same direction in both platforms. *NELL2*, a protein kinase C-binding protein, was the most significant result with gene expression levels decreasing with age (Illumina p=8.2E-81, Affymetrix p=3.2E-64). *NELL2* is involved in cell growth regulation and differentiation, and there is evidence for developmental fluctuation in puberty. We identified 347 transcripts differentially expressed between males and females(p<1E-5, same direction both platforms), of which >200 show mapping to sex chromosomes. The top autosomal gender-differentiated transcript is *DACT1*, which has higher mRNA levels in females (Illumina p=2.4E-47, Affymetrix p=1.6E-75). *DACT1* is an antagonist of beta-catenin and prior work indicates it to be differentially methylated in testes. It is a biomarker for semen and *DACT1* knockout mice showed developmental defects. Both the *NELL2* and the *DACT1* signals were replicated in all 3 additional cohorts. With the GWEP meta-analysis, we gained power relative to individual cohort analyses, and were able to identify novel replicable significant age- and sex- associated loci. These loci may have implications for age-related disease biology, gender biology, and in sample forensics.

585F

Simplified sample preparation and streamlined library preparation method for deep sequencing Ribosome Protected mRNA Fragments. *R. Vaidyanathan[1], S. Kuersten[1], A. Radek[1], S. Swami[2], D. Vo[3], L. Penalva[3]*. 1) Epicentre, Madison, WI; 2) Illumina, Hawyard, CA; 3) University of Texas Health Science Centre at San Antonio, Dept. of Cellular and Structural Biology, San Antonio, TX.

Musashi1 (Msi1) is an evolutionarily conserved RNA-binding protein that has been implicated in a variety of cellular processes such as stem cell maintenance, nervous system development, and tumorigenesis. Msi1 is highly expressed in many cancers including glioblastoma, and has been implicated in neurogenesis, neurodegenerative diseases, and is emerging as a potential therapeutic target in both regenerative medicine and cancer. We are using the method of ribosome profiling to understand the regulatory role of Msi1 at the level of translation. Conventional RNA abundance measurements using microarrays or high-throughput RNA-sequencing reflect protein levels for transcripts that are not subjected to translation control. The isolation of translating ribosomes (polysomes) and the analysis of the associated mRNA fragments resulting from the foot-printed ribosomes (RNase protected monosomes) provide a better measure of translation rates to estimate the copies of the synthesized protein. Furthermore, comparison of total mRNA to the Ribosome Protected mRNA Fragments (RPF) can provide quantitative analysis of the translationally active verses repressed pools of transcripts present at a particular time or condition. A small hairpin (sh) RNA was used to knock down the Msi1 functions in the glioblastoma cell line U251. We prepared footprinted monosomes from these cells and constructed libraries from ribosome protected mRNA fragments for analysis by deep sequencing. The goal is to perform a comparative analysis of transcripts from actively translating ribosomes in shMsi1 cells versus control cells to identify direct mRNA targets of Msi1 and better understand how this RNA binding protein regulates translational control in the nervous system.

586W

Genetics and Epigenetic Regulation of Eukaryotic Transcriptome. *Y. Wakabayashi[1, 4], T. Ni[1], D. Corcoran[2], W. Yang[3], U. Ohler[2], W. Peng[3], C. O'Donnell[4], j. Zhu[1]*. 1) Genetics and Development Biology Center, NHLBI, NIH, Bethesda, MD; 2) Institute for Genome Sciences & Policy, Duke University Medical Center, Durham, NC; 3) Department of Physics, George Washington University, Washington, DC; 4) Framingham Heart Study, NHLBI, NIH, Framingham, MA.

Emerging sequencing technology has provided unprecedented throughput for monitoring transcriptome complexity. In particular, alternative regulations in multiple steps of mRNA biogenesis, including but not limited to transcriptional initiation, pre-mRNA splicing and polyadenylation, have been shown to serve as major contributors for variant transcript production. One fundamental question is how genetic and epigenetic codes are wired in the genome to control tissue and/or cell-specific gene expression. We have developed several sequencing-based technologies to investigate differential promoter usage, antisense transcription and alternative 3'-end formation of eukaryotic transcriptomes. Our early work showed that Drosophila promoters, similar to that of mammalian ones, exhibit “peak” and “broad” initiation patterns. Further analyses demonstrated that core promoter motifs as well as their location preferences, in together with local chromatin structure, play important roles in promoter choices. In addition, a compendium of genomics data has also been collected to investigate the connectivity and feedback regulation of transcription initiation/elongation and alternative polyadenylation at the systems level. Taken together, integration of genomics and epigenomics information is expected to shed light on the complexity and regulation of eukaryotic transcriptomes.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

587F

Differential Protein Expression in Human Induced Pluripotent and Embryonic Stem Cells. *S. Brown-Ford[1], M. Pellitteri-Hahn[2], A. De La Forest[3], B. Halligan[2], S. Duncan[3], M. Olivier[1, 2].* 1) Dept. of Physiology, Medical College of Wisconsin, Milwaukee, WI; 2) Biotechnology Bioengineering Center, Medical College of Wisconsin, Milwaukee, WI; 3) Cell Biology, Neurobiology & Anatomy, Medical College of Wisconsin, Milwaukee, WI.

Human induced pluripotent stem cells (hiPSC) have been proposed as unique resources in the study and treatment of disease. However, many questions remain regarding how similar iPSC are to embryonic stem cells (ESC). Early characterizations using mRNA expression profiling have reported that these two cell types are almost identical. Noticeably absent from early characterizations was information on protein expression in ESC and iPSC. These questions are now being addressed. In a recent study <1% of detected proteins were expressed with two-fold difference between a human iPSC and a human ESC line, implying that iPSC and ESC are truly very similar on a cellular level. Here, we analyzed protein expression in stem cell lines derived from different individuals to assess the inter-individual and stem cell-specific differences in protein expression levels. We characterized the proteome of 3 hiPSC and 3 hESC lines using mass spectrometry. Stem cell lines were grown under standard conditions, lysed, fractionated into soluble and insoluble fractions, and analyzed using an ion-trap mass spectrometer (LTQ Orbitrap Velos). Two biological replicates and at least two technical replicates for each cell line were included in the analysis. Mass spectral data were analyzed using SEQUEST, and results from replicates were pooled. Overall, few proteins were detected at significantly different levels. On average, 2.2% of proteins were significantly different between biological replicates of the same cell line. Cell lines derived from the same individual, but independent reprogramming events showed similar differences (3.0%). Between these lines, >88% of detected proteins were shared. iPSC from genetically different individuals showed the highest amount of differential expression (5.3%). Hierarchal clustering indicated that protein expression in ESC and iPSC were more similar to each other than to a fibroblast line. Based on clustering, iPSC and ESC lines were indistinguishable. Reprogramming factors Oct3/4 and Lin28 were not significantly differentially expressed between ESC and iPSC. 22% of differentially expressed proteins between the two were histones, which has been reported in recent analyses. This analysis, for the first time, demonstrates that iPSC and ESC show highly similar protein expression, and that variability in protein expression may not be cell type specific, but determined by inter-individual and genetic differences affecting protein expression.

588W

Understanding the molecular mechanisms underlying autism spectrum disorders using iPSC-derived neurons. *B.A. DeRosa[1, 2,3], J.M. Van Baaren[1,3], G.K. Dubey[1,3], J.M. Lee[1,3], M.L. Cuccaro[1,2,3], J.M. Vance[1,2,3], M.A. Pericak-Vance[1,2,3], D.M. Dykxhoorn[1,2,3].* 1) John P. Hussman Institute For Human Genomics; 2) Dr. John T. Macdonald Foundatioun Department of Human Genetics; 3) University of Miami Miller School of Medicine, Miami, FL 33136.

Autism spectrum disorder (ASD) is a heterogeneous, neurodevelopmental disorder characterized by incapacities in social interaction verbal and non-verbal communication and repetitive stereotypical behaviors, with recent reports estimating an average of 1 in 88 children being affected. In spite of this high prevalence, relatively little is known about the etiology of the disorder. A major constraint in ASD research has been the inaccessibility of disease-affected neural progenitors and neurons in which to study the molecular pathophysiology of ASD during neurodevelopment. A number of studies have used post-mortem brain tissue collected from individuals affected with autism. However, any findings that are identified in post-mortem ASD brain samples are likely to only represent an end point in the pathology of autism. This is because a formal diagnosis of ASD cannot be made until around 3 years of age. In order to study the processes that initiate the onset of ASD and those that have roles at various stages of neurogenesis, we have developed several induced pluripotent stem cell (iPSC) lines which were derived from the peripheral blood of individuals affected with varying degrees of autism. These ASD-specific iPSC lines are able to differentiate into neural stem cells and progenitors that give rise to electrophysiologically active cortical neurons (i.e. GABAergic and glutamatergic neurons) in a process that imitates in vivo neurodevelopment. Neurons derived from the ASD-iPSCs exhibit aberrations in morphology and proliferation compared to iPSC-derived neurons from unaffected individuals. This system provides an ideal opportunity for the use of next-generation sequencing to investigate the fundamental biological processes at play in autism through in-depth molecular profiling of ASD-affected neurons in a developmental stage and neuronal subtype-specific manner. This strategy will hopefully allow us to answer important questions about the pathophysiology of autism that could be used toward the advancement of novel therapeutics.

589F

Whole exome sequencing in familial non-HHT Brain Arteriovenous Malformations. *N. Bendjilali[1], L. Tang[4], DE. Guo[1], R. Hernandez[6], M. Segal[5], CE. McCulloch[5], S. Inoue[3], A. Koizumi[2], WL. Young[1], L. Pawlikowska[1,4], H. Kim[1,4,5].* 1) Center for Cerebrovascular Research, Department of Anesthesia and Perioperative Care, University of California, San Francisco, CA; 2) Department of Health and Environmental Sciences, Kyoto University Graduate School of Medicine, Yoshida, Kyoto, Japan; 3) School of Pharmacy, Suzuka University of Medical Science, Suzuka, Mie, Japan; 4) Institute for Human Genetics,University of California, San Francisco, CA; 5) Department of Epidemiology and Biostatistics, University of California, San Francisco, CA; 6) Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, CA.

Background: Brain arteriovenous malformations (BAVM) are rare vascular defects where arteries connect directly to veins, resulting in high-flow lesions prone to rupture. Most cases are sporadic, but BAVMs often occur patients with in Hereditary Hemorrhagic Telangiectasia (HHT), a disease caused by mutations in ALK1, ENG or SMAD4. Prior studies have suggested a genetic contribution to BAVM in familial non-HHT Japanese cases. We hypothesized that exome sequencing may identify novel, rare variants shared among these Japanese non-HHT familial BAVM cases. **Methods:** We performed exome capture and high throughput sequencing of 10 Japanese samples: 8 subjects with non-HHT BAVM (3 sibling pairs, and 2 affected from 2 different families) and 2 spouse controls. Reads were mapped to the reference human genome (UCSC hg19) with Burrows-Wheeler Aligner. Variants were called using Genome Analysis Toolkit and annotated with Annovar. For BAVM samples, we identified genes with potentially pathogenic variants (non-synonymous SNPs, insertions/deletions resulting in frameshift or amino acid change) shared within families and between at least two families. Variants present in controls (2 spousal Japanese control exomes, 73 Japanese 1000 Genomes Project samples, and 45 Japanese HapMap samples) were excluded. Only variants predicted to disrupt protein function using Polyphen-2 analysis were considered as possible candidates. **Results:** An average of 4.3Gb of mappable sequence was obtained for a target coverage of 69x per individual; 90% of all target positions achieved 10x plus coverage, and an average of 43,339 variants were called per sample. We identified 56 genes with possibly deleterious variants in at least 2 families but not in controls. No genes had variants present in all 5 families; 2 genes (CD3G and PSG4) had variants in 4 families and 9 genes had variants in 3 families. Four genes (PSG4, TTN, PCDHA8 and PCDHGA11) are in biological pathways that may be relevant to BAVM. One of the 56 genes with variants in at least 2 families lies in a previously reported linkage region on chromosome 13. **Conclusions:** Exome sequencing in 8 familial non-HHT BAVM cases identified several genes bearing candidate rare variants in cases only. More experiments are needed to validate these findings and determine the potential role of the genes identified in BAVM pathogenesis.

590W

The Investigation of Angiotensin Converting Enzyme I/D (ACE I/D) and Plasminogen Activator Inhibitor-1(PAI-1) 4G/5G Polymorphisms in Venous Thromboembolism Patients in Turkish patients. *H. Salifoglu, H. Kaya, M. Karkucak, D. Torun, Y. Tunca.* Gulhane Military Medical Academy, Ankara, Turkey.

Venous tromboembolism(VTE) refers to trombolytic events in the venous system that are most common manifested as deep vein thromboses in the upper or lower extremity and / or pulmoner embolism.As genetic factors differs among the populations; lots of studies were done about the polymorphism related to VTE. As a result of these studies the relationship between disease development and polymorphism is not clear yet. In this study we aimed to investigate the role of Angiotensin Converting Enzyme Insertion/Deletion(ACEI/D)and Plasminogen Activator Inhibitor-1 4G/5G polymorphism in the development of disease. The study population consisted of a group of 80 VTE patients and 79 healthy subjects as a control group. Genotypes were determined by using classical polymerase chain reaction(PCR) for ACE I/D polymorphism and the PCR based on allele spesific amplification for PAI-1 4G/5G polymorphism in the development of disease. As a result ; there were no significant differences for ACE I/D and PAI-1 4G/5G polymorphism among patient and control groups($p>0.05$) These findings revealed that there is no relationship between those polymorphisms and the development of VTE, but more studies with larger group of patients and controls are needed to be done.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

591F

Organization of alphoid sequence clusters on human chromosome 21. *W. Ziccardi, C. Putonti, J. Doering.* Dept Biol, Loyola Univ Chicago, Chicago, IL.

The human genome sequence does not include the heterochromatic regions, although these sequences comprise 10–15% of the genome. We are constructing a detailed physical map of the HC21 centromere and short arm as a model for the organization of these regions. There are two types of alphoid sequences in the human genome: homogeneous higher order repeat (HOR) arrays and more ancient heterogeneous monomeric clusters, which lack any higher organization. Our prior work showed that alphoid DNA on HC21 is present in two distinct regions: D21Z1 and α21-II. The D21Z1 cluster is required for centromere function and consists of a homogeneous array of 11-mer HORs. α21-II is a region on the p arm containing putatively monomeric alphoid sequences organized as at least five distinct clusters (Mp1–Mp5) spread over some 5 Mb and dispersed among many other non-alphoid sequences. Using 21p BACs already in GenBank as well as new clones we acquired, a total of 0.5 Mb of α21-II sequence was analyzed that included representative sequence from each of the five clusters. The α21-II alphoid sequences are in fact not purely monomeric since two of the clusters have an HOR organization. While one of these (Mp3) has a highly degenerate HOR array, the other (Mp4) has a highly homogeneous organization comparable to that seen in D21Z1. The α21-II alphoid clusters are larger than comparable regions on other chromosomes and they are also more distant from the functional centromere. While the overwhelming majority of non-centromeric alphoid arrays are on the p arm of HC21, other human chromosomes have symmetry in the location of these arrays with relatively equal amounts of alphoid sequence on their q and p arms, Overall, HC21 alphoid clusters with an HOR organization are evolutionarily distinct from the monomeric arrays, and D21Z1 appears to be the most recently evolved cluster. All of the monomeric clusters are of recent evolutionary origin and appear to be of a similar age to each other. This organization is quite different from the layered structure seen on other chromosomes where the age of a non-centromeric alphoid array is related to its distance from the functional centromere, with the oldest arrays being most distal. The different type of alphoid organization on HC21p may be due to relatively frequent interchromosomal exchanges between the p arms of the acrocentric chromosomes.

592W

A high-resolution genome-wide map of meiotic double-strand breaks in humans. *F. Pratto[1], K. Brick[1], P. Khil[1], F. Smagulova[2], G. Petukhova[2], R.D. Camerini-Otero[1].* 1) Genetics and Biochemistry Branch, National Institutes of Health, Bethesda, MD; 2) Uniformed Services University of Health Sciences, Department of Biochemistry and Molecular Biology, Bethesda, MD.

Meiotic recombination is initiated by the formation of programmed double-strand breaks (DSBs). The sites where these breaks are introduced are clustered in narrow regions called hotspots. Recent studies have determined that PRDM9, a meiosis specific histone H3 methyl-transferase, is the primary determinant of hotspot location. To date, genome-wide mapping of meiotic hotspots in humans was only possible via computational analysis of patterns of linkage disequilibrium in populations. Using LD patterns, more than 30,000 hotspots have been identified. LD-based analyses have several important caveats; they may not reflect present day crossovers, are prone to variability in recombination rate estimates and allow only for evaluation of sex-averaged and population-averaged recombination rates. In this work, we perform the first direct genome-wide mapping of meiotic recombination initiation hotspots in humans. To map hotspots we have used a sequencing-based technique that recently allowed us to generate genome-wide maps of meiotic DSBs in mouse (Khil et al. Genome Res (2012), Brick et al. Nature (2012), Smagulova et al. Nature (2011)). We found that the genome-wide distribution of human DSBs in male meiosis resembles the distribution of male specific, but not female-specific crossovers, with both male crossovers and DSBs more frequent in subtelomeric regions. This suggests that the frequency of crossovers is directly related to the frequency of DSBs and that regulation at the level of initiation is shaping the recombination landscape. Furthermore, the positions of DSB hotspots agree well with the positions of LD-based hotspots, although there are notable differences. In order to estimate inter-individual variability and the effect of different Prdm9 alleles on DSB hotspot locations, we mapped DSBs in four individuals; two were homozygous for the most common Prdm9 allele in the population (A/A) and two were heterozygous for the A allele and a closely related variant, the B allele (A/B). We found a substantial number of B allele defined hotspots and found that these hotspots are not as well correlated with LD hotspots as the A-defined hotspots. Finally, we found that DSB hotspots are enriched at breakpoints of copy number variants (disease associated or not) that arise via a homology mediated mechanism providing support for a role of programmed meiotic DSBs as the origin of unequal crossovers in humans.

593F

Genome-wide scan of disease-inversion association. *J. Ma, C.I. Amos.* Genetics, UT MD Anderson Cancer Center, Houston, TX.

Chromosomal inversions are known to be associated with susceptibility to disease. The recent advent of paired-end sequencing approaches has made it possible to map inversions across the human genome. However, neither this sequencing-based approach nor the traditional cytogenetic approaches can easily scale to the number of samples needed for inversion-disease association tests. We have recently developed a novel approach, based on principal components analysis (PCA), for detecting and genotyping inversions from high-density single nucleotide polymorphism (SNP) array data. Inside an inversion region, recombination is usually suppressed between the inverted and non-inverted segments, which represent two distinct lineages that have been diverging for many generations and accumulating mutations independently. Locally performing PCA inside or around the inversion region will generate three clusters on the space spanned by the first two eigenvectors corresponding to the three inversion genotypes. For widely available case-control data from SNP-based genome-wide association studies (GWAS), if the markers are dense enough, our approach can be directly used to infer inversion status of all samples and thus can be used to perform association tests. If the markers are not dense enough for meaningful PCA in inversion regions due to quality control in GWAS, we can use surrogate SNPs to infer the inversion genotypes. For most of the known inversions and our predicted inversions using PCA, we found that some of the SNPs inside an inversion region are fixed in the two lineages of different orientations and thus can serve as surrogate markers. Those surrogate SNPs, for which the allele frequencies are very close to 50% in inversion heterozygotes, can be readily determined from SNP data with dense markers, such as the HapMap data. We performed genome-wide association studies by applying our approach to real GWAS data sets for known inversions given in the literature and our predicted inversions and identified a few loci associated to melanoma or lung cancer.

594W

HOXA11 and MMP2 Gene Expressions in Women with Pelvic Organ Prolapsus. *N. Yilmaz[1], G. Ozaksit[1], Y.K. Terzi[2], S. Yilmaz[1], B. Budak[1], O. Aksakal[1], F.I. Sahin[2].* 1) Dr. ZTB Women's Health Education and Research Hospital, Ankara, Turkey; 2) Baskent University Faculty of Medicine Department of Medical Genetics, Ankara, Turkey.

Pelvic organ prolapse (POP) is a common and costly disorder that negatively impacts the quality of life in many women. Uterosacral ligaments (USLs) are supportive structures of the uterus and vagina and are often attenuated in women with POP. Mechanical reasons as well as genetic factors may play a role in POP. The HOXA genes regulate the development of the urogenital tract including uterosacral ligaments. We compared expression of HOXA11, MMP2 in USLs of women with and without POP. We compared HOXA11 and MMP2 expression in USLs of women with (n: 34) and without (n: 24) POP. Total RNA isolated from patient's (n: 34) and control's (n: 24) uterosacral ligament tissues with TriPure isolation reagent according to the manufacturer's instructions. Expression levels of HOXA11 and MMP2 were determined using semiquantitative RT-PCR in Light Cycler 480 system. Realtime ready catalog assays which are short FAM-labeled hydrolysis probe containing locked nucleic acids were used for RT-PCR reactions. Means of RNA expression of MMP2 were 1.17 ± 0.5 and 0.7± 0.3 in POP group vs control group, respectively (p= 0.003). Means of RNA expression of HOXA 11 were 2.12±1.9 and 1.79±1. 3 in POP group vs control group, respectively (p= 0.487). Although, we did not find a statistically significant difference for HOXA expression between the patient and the control groups, the difference in MMP2 expressions were statistically significant. Our results suggest that, further long term investigations with larger numbers of cases are needed to explain the genetic mechanisms underlying pelvic organ prolapsus.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

595F

Large scale whole genome and whole exome sequencing of rhesus monkeys (*Macaca mulatta*) in five primate research centers. *M. Raveendran[1], D. Rio Deiros[1], G. L. Fawcett[1,2], Z. Johnson[3], N. H. Kalin[4], R. W. Wiseman[5], B. Ferguson[6], E. Vallender[7], S. Kanthaswamy[8], D. M. Muzny[1], R. Gibbs[1,2], J. Rogers[1,2].* 1) Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX; 2) Dept. of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX; 3) Yerkes National Primate Research Center, Emory University, Atlanta, GA; 4) HealthEmotions Research Institute and Dept. of Psychiatry, University of Wisconsin, Madison, WI; 5) Wisconsin National Primate Research Center, Madison, WI; 6) Oregon National Primate Research Center, Oregon Health & Science Univ., Beaverton, OR; 7) New England Primate Research Center, Harvard Medical School, Southborough, MA; 8) California National Primate Research Center, University of California-Davis, Davis, CA.

Rhesus macaques (Macaca mulatta) are important animal models for biomedical research and evolutionary analyses. Prior studies show this species has greater genetic diversity than humans, but little is known about functionally significant variants in the research population of macaques. We report preliminary efforts to characterize rhesus single nucleotide polymorphisms (SNPs) and other genetic variants through whole genome and whole exome sequencing. We are sequencing animals from five National Primate Research Centers (Yerkes, Wisconsin, Oregon, California and New England NPRCs) using Illumina Hi-Seq methods. Our full dataset will include several hundred subjects each for whole genome (30x) and whole exome sequencing. Here we report preliminary results for 5 whole genome and 8 whole exome sequences. Whole genome sequence reads were mapped to the rhesus reference genome (rhemac2) using BWA, and potential SNPs called using SAMtools mpileup. Using low-stringency coverage requirements we identified >12 million potential SNPs per individual. Pooling across animals to obtain a preliminary population-wide SNP list, we identified >30 million unique single nucleotide variants (SNVs), or one per 100 bp. Counting only SNVs observed in 2 or more individuals yields a higher confidence list of >13 million SNVs, more than in comparable studies of humans. In parallel, we are developing whole exome capture sequencing protocols to increase efficiency in discovering functional SNPs. Rhesus sequences were captured using Roche EZ Choice reagents and a custom design with exons from 18,354 high-confidence human genes. Captured rhesus reads were mapped to rhemac2, which contains 145,849 annotated UCSC exons. 95.7% of human target regions had unambiguous alignments to rhesus genome. The rhesus reads mapped to rhemac2 covered 93% of identified rhesus homologs of the human target regions at >20x coverage. This extends previous studies of the value of human reagents to capture rhesus exons. However, only 76% of bases in the 145,849 annotated rhesus exons were covered at >20x. Initial preliminary analysis of the exome data shows approximately 300 frame shift mutations and about 240 stop codon mutations. Human capture arrays provide good coverage of most of the rhesus exome but both chip design and differences between the genomes in gene content and sequence divergence reduce the ability of human arrays to interrogate the entire rhesus exome.

596W

Whole Genome Haplotyping by Dilution, Amplification and Sequencing. *F. Kaper, S. Swamy, B. Klotzle, J. Cottrell, S. Munchel, M. Bibikova, S. Kruglyak, M. Ronaghi, M.A. Eberle, J.B. Fan.* Illumina, Inc., San Diego, CA.

Standard whole-genome genotyping technologies are unable to determine haplotypes. Several experimental methods have been described for whole-genome phasing, mostly based on separation of chromosomes or large genomic fragments. Generally these methods are expensive and labor-intensive. We have developed a method for rapid and cost-effective whole-genome haplotyping. Genomic DNA is diluted and distributed into multiple aliquots such that each aliquot receives a fraction of one haploid copy. The template in each aliquot is amplified by MDA, converted into barcoded sequencing libraries using Nextera® technology and sequenced in multiplexed pools. We first carried out a proof of concept experiment combining two male genomic DNA samples at equal ratios, resulting in a sample with diploid X-chromosomes with known haplotypes. Pools of the multiplexed sequencing libraries were subjected to targeted pull-down of a 1Mb contiguous region of the X-chromosome DMD gene. Additionally, we sequenced each undiluted male separately to confirm the haplotypes and the undiluted mixed sample to detect all heterozygous SNPs within the region. We identified 1265 heterozygous SNPs within the targeted region in the DNA mixture. Using the dilution sequencing approach, we identified individual contiguous regions within each dilution and determined the haplotypes based on the identified SNPs. These smaller haplotype segments were merged into two contiguous haplotype blocks with a mean length of ~494kb that span ~99% of the region. Genomic context rather than heterozygosity was responsible for the breakup of the region into two blocks. Comparison of the phased haplotype blocks with the haplotypes of the two individual males showed ~99% agreement. We subsequently employed the dilution-amplification strategy to haplotype the entire genome of a Yoruba male DNA sample (Corriell ID NA18506). Standard genomic sequencing to identify all heterozygous SNPs in the sample was combined with dilution-amplification based sequencing data to resolve the phase of identified heterozygous SNPs. Using this procedure we were able to phase >93% of the heterozygous SNPs from the diploid sequence data using 192 dilution-amplification reactions at ~0.4 haploid copies per aliquot. The average haplotype block size for the whole genome experiment was ~250kb with a maximum of approximately 5Mb. Therefore this strategy has successfully been applied to haplotyping of both targeted regions and the entire genome.

597F

Mechanism and Potential Diversity of T-Cell Receptor Rearrangement from Sequence Repertoires. *A. Murugan[1], T. Mora[2], A. Walczak[3], C. Callan[1].* 1) Dept. of Physics, Princeton University, Princeton, NJ; 2) Laboratoire de physique statistique, UMR8550, CNRS, Paris, France; 3) Laboratoire de physique theorique, UMR8549, CNRS, Paris, France.

T-cell receptor (TCR) diversity is generated by VDJ recombination. In this process, germline DNA is rearranged and edited to produce a random TCR coding sequence. This editing involves random choices of the rearranged V, D and J genes, deletions of nucleotides from their ends, and the insertion of random nucleotides at the two junctions (VD and DJ). The mechanism and potential diversity of this stochastic molecular process is still not fully understood. We use deep sequencing data of TCR beta chain CDR3 repertoires in nine individuals to infer a probabilistic generative model of the VDJ rearrangement process. Since the repertoire is shaped by molecular constraints as well as selection, we focus on non-productive sequences (out-of-frame or containing a stop codon) that have not been subjected to selective forces. We find high consistency of the model between different individuals suggesting its origin in a universal molecular mechanism. The generative model characterizes the probability distributions of deletions, insertions and gene choice and the correlations between them. We find that the deletions are highly dependent on gene choice, reflecting sequence dependent nuclease activity. The insertions show a sequence bias, captured well by a di-nucleotide model. We calculate the potential diversity of TCRs and find a nucleotide sequence entropy of 47 bits (~10^14 sequences). We also find quantitative agreement between the occurrence of identical TCRs in different individuals and the predictions of the model for such recurring sequences. The generative model also serves as a baseline for signatures of selection on the repertoire. The model is inferred by maximizing likelihood of the data using an Expectation Maximization algorithm.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

598W

Expression of mouse acidic mammalian chitinase in Escherichia coli and characterization of its properties. *A. Kashimura, Y. Kida, M. Ohno, K. Ishikawa, M. Sakaguchi, F. Oyama.* Applied Chemistr, Faculty of Engineering, Kogakuin University, Hachioji, Tokyo, Japan.

Chitinase hydrolyze chitin, a polymer of N-acetyl-D-glucosamine, which is present in a wide range of organisms, including fungi, insects, and parasites. Although chitin is not present in mammals, chitinase genes were found in both human and mouse genomes. Acidic mammalian chitinase (AMCase) have been shown to be closely associated with allergic inflammation and food processing in mouse. To understand its pathophysiological roles, we expressed the mouse AMCase in E. coli as a fusion protein. We found that most of the expressed fusion protein was present in intracellular soluble fraction in E. coli with chitinolytic activity. The recombinant protein shows a distinct pH optimum at pH 2.0. Under the acidic condition, the optimal temperature was 54 oC. The recombinant AMCase binds with chitin beads and cleaves colloidal chitin and releases mainly chitobioside fragments from colloidal chitin, comparable to those of the secreted form of AMCase expressed in CHO cells or mouse stomach. This recombinant protein can be used to elucidate detailed biological functions of the mouse chitinase.

599F

Functional analysis of mouse chitinase-like protein Ym1. *Y. Kida, M. Sakaguchi, Y. Sugahara, F. Oyama.* Dept Applied Chemistry, Kogakuin Univ, Hachioji, Tokyo 192-0015, Japan.

Chitin is a widespread environmental biopolymer of N-acetyl-D-glucosamine and provides structural rigidity to fungi, crustaceans and insects. Although chitin itself does not exist in mammals, chitinases and chitinase-like proteins (CLPs) are expressed in mouse and human. Both mammalian chitinases and CLPs belong to the family 18 of glycosyl hydrolases. CLPs are homologous to chitinases but lack the ability to degrade chitin. It is generally assumed that the lack of chitinase activity in CLPs is due to the mutation of crucial residues within the conserved catalytic domain during evolution from a common ancestor. Ym1, one of the CLPs, is expressed in mouse and induced in inflammation and pathogenesis, but its biological functions remain unknown. Ym1 is highly homologous to acidic mammalian chitinase (AMCase), but lacks chitin hydrolyzing activity. To understand the structure and function relationship of Ym1, we introduced amino acid substitutions (N136D and Q140E), in which both acidic amino acid residues are the highly conserved within chitinases. Although the mutated Ym1 possessed the identical sequence to the catalytic motif of AMCase, it remains inactive. This result indicates that lack of chitinase activity in Ym1 is not only attributed to the simple amino acid substitution in the conserved catalytic motif.

600F

Targeted deep sequencing by Mobile Element Scanning (ME-Scan) reveals Alu insertions throughout the genome and around the world. *D.J. Witherspoon[1], W.S. Watkins[1], Y. Zhang[1,2,3], J. Xing[1,2], M.A. Batzer[4], L.B. Jorde[1].* 1) Dept. of Human Genetics, University of Utah, Salt Lake City, UT; 2) Dept. of Genetics, Rutgers, the State University of New Jersey, Piscataway, NJ; 3) Bionomics Research & Technology Center, Rutgers, the State University of New Jersey Piscataway, NJ; 4) Dept. of Biological Sciences, Louisiana State University, Baton Rouge, LA.

Mobile elements are among evolution's greatest success stories. They are ubiquitous, inhabiting virtually all genomes examined to date. In humans, retrotransposons of the Alu and LINE-1 families have reproduced so effectively and persistently that their copies make up the majority of our DNA. These genomic parasites have had a profound impact on human genetic diversity and genomic structure. Their activity continues to create novel mutations and cause genetic diseases today. Despite their pervasiveness, it remains difficult to quantitatively study their evolution, population dynamics, and impact. Mobile element insertions can only be clearly observed at the level of DNA sequence, and even that is difficult to obtain due to their repetitive nature. To address this challenge, we have improved our method of targeted high-throughput sequencing (mobile element scanning, or ME-Scan; Witherspoon et al. BMC Genomics 2010) and applied it to detect novel and polymorphic AluYb8/9 insertions in 211 individuals from 13 diverse populations: Brahmin, Mala, Madiga and tribal Irula from India; Mbuti Pygmy, Yoruba, Luhya, Alur and Hema from Africa; Italians and northern Europeans from Europe; and Japanese and Chinese from East Asia. We also scanned one 14-member family to detect retrotransposition events between parents and offspring. We indexed and pooled up to 25 individuals per library for sequencing on Illumina HiSeq and GAIIx sequencers, then used custom mapping-based computational analyses to identify known and novel Alu insertions in the data. The per-individual sensitivity of ME-Scan insertion is ~93%, and ME-Scan results are 95% concordant with Alu insertion genotypes obtained by locus-specific PCR. As expected, African populations carry more novel and rare insertions than non-African ones. Based on internal replication results, we applied conservative criteria to reduce false positive results. We detected 1543 novel Alu insertions, including many that are present in only a single individual, and identified insertions corresponding to 3086 previously-known polymorphic Alu insertion loci. Despite the large number of novel insertions, only 27 were found interrupting exons, and only 2 were in protein-coding exons. This dearth of gene-interrupting insertions is statistically significant ($p<0.001$ in both cases, compared to random distribution of novel rare insertions). This implies that natural selection is extremely intolerant of Alu insertions in exons.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

601T

Identification of *HNRNPA1* as a novel regulator of cholesterol metabolism. *C. Yu[1], E. Theusch[1], D.A. Nickerson[2], K. Lo[1], M. Kutilova[1], R.M. Krauss[1], M.W. Medina[1]*. 1) Children's Hospital Oakland Research Institute, Oakland, CA; 2) Department of Genome Sciences, University of Washington, Seattle, WA.

HMGCR encodes the rate limiting enzyme in the cholesterol biosynthesis pathway and is inhibited by statins, a class of cholesterol lowering drugs. Expression of an alternatively spliced *HMGCR* transcript lacking exon 13, *HMGCR13(-)*, has been implicated in variation in plasma LDL-cholesterol (LDL-C) and is also the single most informative molecular marker of variation in LDL-C response to statin treatment. *In silico* prediction of splicing factors that interact with rs3846662, a single nucleotide polymorphism (SNP) within *HMGCR* shown to regulate exon 13 skipping, identified a putative interaction with heterogeneous nuclear ribonucleoprotein A1 (HNRNPA1). Sterol depletion of human hepatoma cell lines reduced *HNRNPA1* transcript levels, an effect that was reversed with sterol add-back. Overexpression of HNRNPA1 increased the ratio of *HMGCR13(-)* to total *HMGCR* transcripts by both directly increasing exon 13 skipping in a rs3846662 allele-specific manner, as well as preferentially stabilizing the *HMGCR13(-)* versus the *HMGCR13(+)* transcript. Importantly, HNRNPA1 overexpression also diminished HMGCR enzyme activity, enhanced LDL-C uptake, and increased cellular APOB. Furthermore, we found that a SNP, rs1920045, associated with statin-induced *HNRNPA1* exon 8 skipping, was also correlated with *in vivo* statin response as homozygous carriers had a 7–11% smaller reduction in total cholesterol after statin treatment in subjects from the Cholesterol and Pharmacogenetics (CAP) trial (simvastatin 40mg/day, 6 weeks) and the Pravastatin Inflammation/CRP Evaluation (PRINCE) study (pravastatin 40mg/day, 24 weeks). These results suggest that sterol regulation of *HNRNPA1* modulates *HMGCR* alternative splicing in a genetically regulated manner to impact cellular cholesterol metabolism, further implicating the importance of alternative splicing as a fundamental mechanism underlying variation in both risk factors and the treatment of cardiovascular disease.

602T

Novel association of dual anti-platelet drug response with a functional variant in *PPARG*. *A.S. Fisch, J. Liu, J.P. Lewis, L.M. Yerges-Armstrong, J.R. O'Connell, B.D. Mitchell, R.B. Horenstein, N. Ambulos, K. Ryan, Q. Gibson, J. Shelton, A.R. Shuldiner.* University of Maryland School of Medicine, Baltimore, MD.

Clopidogrel and aspirin (dual anti-platelet therapy; DAPT) are widely used medications for patients with coronary artery disease to prevent myocardial infarction, especially in patients with acute coronary syndrome and following percutaneous coronary intervention. Patients' responses to DAPT are highly variable, and heritability estimates suggest that 70–80% of the variability in on-treatment platelet reactivity may have a genetic basis. Clopidogrel is a prodrug that requires activation by cytochrome P450 enzymes. Studies have shown that the common loss-of-function *CYP2C19**2 allele explains approximately 12% of the variation in on-treatment platelet reactivity and is associated with a higher incidence of cardiovascular events in DAPT-treated subjects. In search of additional variants that influence DAPT response, we genotyped 781 informative SNPs (739 SNPs with MAF>1%) (Affymetrix DMET chip) in 670 healthy individuals from the Amish Pharmacogenomics of Anti-Platelet Intervention-1 (PAPI-1) Study who received DAPT, and had pre- and post-treatment platelet reactivity measured by ex vivo light transmission aggregometry. Association of the effect of SNPs on platelet aggregation was calculated using a variance component method under an additive model simultaneously adjusting for age, sex, and relatedness. After considering multiple comparisons, we established significance at $P<6.4\times10^{-5}$, and marginal significance at $6.4\times10^{-5}<P<6.4\times10^{-4}$. We replicated the previously reported significant association between on-treatment platelet reactivity and SNPs in *CYP2C19* (*2; rs4244285) and *ABCB1* (C3435T; rs1045642). In addition, we found a novel significant association between post-DAPT therapy collagen-stimulated aggregation and the common functional Pro12Ala variant in *PPARG*, a nuclear receptor expressed in liver and other tissues that regulates transcription of multiple genes; specifically the CG genotype was associated with a greater inhibition of post-DAPT platelet aggregation (rs1801282; lag time in N=620 CC and N=26 CG, 73 sec and 111 sec, respectively; $P=2.8\times10^{-5}$). We also report a significant association between DAPT response and variants in genes involved in drug sulfation (*SULTlA2*, *SULT1A3*), and marginal significance in other drug metabolizing genes (*GSTM2*, *CYP2E1*, *CYP3A43*) and drug transporters (*ABCC1*, *ABCC3*, *ABCC8*, *SLC7A8*). Replication and functional studies will be required to further define the role of these gene variants and pathways in DAPT response.

603T

Genetic Variations associated with gemcitabine treatment outcome in Pancreatic Cancer. *L. Li[1], B. Fridley[2], G. Jenkins[2], E. Carlson[2], K. Kalari[2], G. Petersen[2], R. McWilliams[3], L. Wang[1]*. 1) Dept of Molecular Pharmacology and Experimental Therapeutics, Mayo Clinic, Rochester, MN; 2) Dept of Health Sciences Research, Mayo Clinic, Rochester, MN; 3) Dept of Oncology, Mayo Clinic, Rochester, MN.

Background: Pancreatic cancer is a rapidly fatal disease with 5-year survival time less than 5%. Gemcitabine is first line therapy for patients with advanced pancreatic cancer for a decade. However, response rate to gemcitabine therapy varies widely. To understand and identify genetic factors that might improve efficacy and avoid toxicity of gemcitabine during pancreatic cancer therapy, we performed a genotype-phenotype association studies using DNA samples from pancreatic cancer patients treated with gemcitabine. Methods: Top 200 SNPs that were associated with gemcitabine cytotoxicity during a previous genome wide association study using 187 human lymphoblastoid cell lines were selected to determine their association with the overall survival (OS) in 397 subjects from pancreatic cancer patients treated with gemcitabine using Cox proportional hazards models. SNP interaction with gemcitabine was also analyzed. We then imputed the regions of +/− 200kb surrounding the top OS-associated SNPs using MACH 1.0 with 1000 Genome Project followed by association of these imputed SNPs with OS. Furthermore, additional 527 subjects were genotyped for top candidate SNPs using TaqMan assay, followed with association analysis when exposed to gemcitabine. Candidate genes were functionally characterized with siRNA knockdown in tumor cells. Results and Conclusions: Five top SNPs located in chromosome 1, 3, 7, 9, and 20 were identified to be associated with OS (P value less than 0.01) when exposed to gemcitabine. Imputation identified additional 103 SNPs that showed stronger associations than those 5 genotyped ones with the lowest p values = 6.47E-5. Those SNPs were mapped to 8 genes with 3 located in the introns of known genes, ARL4A, SICN, and CDH4. Replication was performed with 5 top observed SNPs plus 5 top imputed ones in 527 subjects. We found that one imputed SNP, rs4925193, in CDH4 still showed a trend in the association of OS with gemcitabine exposure (P = 0.051). Knockdown of CDH4 in pancreatic cancer cells significantly desensitized cells to gemcitabine treatment. Our genotype-phenotype association study and functional genomic study would enhance our understanding of the mechanisms involved in gemcitabine response and help us to better predict gemcitabine response in clinic.

604T

Large scale phenotyping using the 1000 Genomes Project lymphoblastoid cell line panel. *D. Spacek, M. Kaganovich, M. Snyder.* Genetics, Stanford University, Palo Alto, CA.

Understanding how differences in genotype lead to differences in phenotype is of fundamental importance to human biology. The 1000 Genomes Project has examined the genetic diversity of a large number of individuals through genome sequencing and has generated a valuable resource of EBV transformed Lymphoblastoid Cell Lines (LCLs). We have developed a high throughput approach that allows us to characterize the cellular response of LCLs to variety of pharmacologic compounds. So far we have begun to characterize the effects of methotrexate, simvastatin, and a number of other compounds on cell growth. The aim is to characterize the effect of these compounds on 1000s of cell lines. By challenging these cells with cytotoxic or cytoprotective compounds we expect to be able to discern the impact of genotypic variation on drug sensitivity. We also plan to expand this approach to the analysis of other phenotypes.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

605T

The association of *SLC6A4* with antidepressant induced mania: The long and the short of it. *J.M. Biernacka[1,2], D.L. Walker[2], J. Geske[1], S. Winham[1], S. Crow[3], M. Chauhan[4], S. McElroy[5], M. Frye[2]*. 1) Dept of Health Sciences Research, Mayo Clinic, Rochester, MN; 2) Dept of Psychiatry and Psychology, Mayo Clinic, Rochester, MN; 3) University of Minnesota, Minneapolis, MN; 4) Mayo Health Systems, Austin, MN; 5) Lindner Center of Hope, Mason, Ohio.

Identifying genetic risk factors associated with antidepressant induced mania (AIM) may enable individualized treatment strategies for bipolar depression. Several studies have investigated the potential association between AIM and a length polymorphism in the promoter region of the serotonin transporter gene (*SLC6A4*), known as 5HTTLPR. We previously performed a meta-analysis of these studies, and found no significant evidence of association of the 5HTTLPR long/short variant with risk of antidepressant induced mania (Biernacka et al., 2012, J Affect Disord 136:e21-9). In the current study, we characterized AIM in a sample of 285 subjects (96 AIM+ cases, 189 AIM- controls) from the Mayo Clinic Bipolar Disorder Biobank that were genotyped for the 5HTTLPR variant as well as the rs25531 SNP in the promoter region and the intron 2 VNTR in *SLC6A4*, and evaluated association of AIM with these variants and their haplotypes. Our analysis did not provide significant evidence of association of the 5HTTLPR short allele with AIM. Moreover, inclusion of this sample in the meta-analysis did not change the conclusion of no significant association between 5HTTLPR and AIM. However, our sample provides evidence of association between the intron 2 VNTR and AIM, with the 10-repeat allele being significantly associated with reduced risk of AIM after accounting for relevant covariates including rapid cycling and presence of attention deficit hyperactivity disorder. Furthermore, a haplotype analysis suggests an association between AIM and haplotypes composed of 5HTTLPR, rs25531, and the intron 2 VNTR, with the L-A-10 haplotype being associated with reduced risk of AIM. A prior pharmacogenomics study of unipolar depression had suggested that the haplotype composed of the *SLC6A4* promoter length polymorphism and intron 2 VNTR is associated with response to antidepressants (Mrazek et al., 2009, Am J Med Genet B Neuropsychiatr Genet 150:341–51); in that study the S-A-12 haplotype was associated with lower odds of remission. Further effort into understanding the relationship between *SLC6A4* variation and response to antidepressants should evaluate the role of the intron 2 VNTR as well as other variants in *SLC6A4*, in addition to the promoter long/short variant.

606T

Interactions between polymorphisms in eicosanoid metabolism genes and vasodilator responses in humans. *L.N. Gordon[1], C.R. Lee[2], M. Pretorius[1], R.N. Schuck[2], L.H. Burch[3], D.C. Zeldin[3], S.M. Williams[1], N.J. Brown[1]*. 1) Center for Human Genetics Research, Vanderbilt University School of Medicine, Nashville, TN; 2) University of North Carolina at Chapel Hill Chapel Hill, NC, USA; 3) National Institute of Environmental Health Sciences RTP, NC.

Endothelial-dependent vasodilation is a complex phenotype that is integral to the development of cardiovascular disease. Cytochrome P450 (CYP)-derived epoxyeicosatrienoic acid (EET) metabolites of arachidonic acid hyperpolarize vascular smooth muscle and have been shown to be potent vasodilators in preclinical models. EETs are formed by the CYP epoxygenases CYP2C8/9 and CYP2J2, and degraded by soluble epoxide hydrolase (EPHX2). Therefore, the steady state of EETs is the product of their production via CYPs balanced by their degradation via EPHX2. Consequently, we assayed 170 single nucleotide polymorphisms in 23 genes that regulate eicosanoid biosynthesis and function for associations with bradykinin-induced changes in forearm blood flow (FBF), forearm vascular resistance (FVR), and mean arterial pressure (MAP) in 67 African-American and 198 European-American subjects. Single SNP associations with these three measurements revealed no significant effects. However, given the pathway architecture involved in endothelial-dependent vasodilation, the effect of two-way SNP interactions was evaluated and numerous significant interactive effects were observed. In European-Americans, we observed significant interactive effects between polymorphisms in the ALOX5 and EPHX2 genes (rs7080474*rs7816586, β=-6.62, 95% CI -9.59 to 3.65, P=2.31×10-5), ALOX5 and CYP2J2 genes (rs7080474*rs10493270, β=12.8, 95% CI 6.82 to 18.7, P=4.15×10-5), and mitochondrial enzyme carbamoyl-phosphate synthase 1 (CPS1) and CYP2C9 genes (rs11681045*rs2253635, β=-6.85, 95% -10.1 to -3.57, P=5.83×10-5) in the FBF response to bradykinin. In African-Americans, we found a significant interactive effect between the angiotensin-converting enzyme (ACE) and CYP2C8 genes (rs4292*rs10882521, β=11.6, 95% CI 5.85 to 17.4, P=5.02×10-6) on the FVR response to bradykinin. Additionally, ten significant interactions were found in bradykinin FBF response. These interaction effects remained significant after correcting for multiple testing using FDR with q=0.10. Collectively, these findings suggest that complex interactions regulate vasodilator responses in humans, and that these effects differ by genetic ancestry.

607T

Pharmacogenetics of schizophrenia: predicting early response to antipsychotic therapy. *M. Gupta[1], S. Jain[2], R. Kukreti[1]*. 1) Genomics & Molecular Medicine, CSIR-IGIB, New Delhi, Delhi, India; 2) Molecular Genetic Laboratory, Department of Psychiatry, NIMHANS, Bangalore, Karnataka, India.

In the post-genome era, the primary focus of pharmacogenomic studies has been to develop rational means for modifying therapeutic treatments with respect to drug and dose regimes, in order to maximize the drug efficacy and tolerability. Such studies are essential, especially for complex diseases such as schizophrenia, where long-term therapies are required and drug response is heterogeneous, following a “hit-and-trial” mode of administration in absence of a mechanism based pre-designed medication regime. In view of this fact, the present report portrays an integrative approach to identify multiple factors influencing antipsychotic drug response and build a predictive model for response to antipsychotic drugs in a south Indian population. For this study, a genetic association strategy was employed to assess polymorphisms in candidate genes from the key drug response related genes in a large, well-characterized sample of 448 antipsychotic-treated schizophrenia patients of south Indian origin, recruited from the clinical services of the National Institute of Mental Health and Neuroscience (NIMHANS), Bangalore, India. The patients were assessed for a three-month period following antipsychotic drug administration, during which they were assessed for response to antipsychotics using Clinical Global Impressions (CGI) and Positive and Negative Syndrome Scale (PANSS). Additionally, other demographic and clinical parameters were evaluated which might influence the genetic variation-drug response relationship. Using a multivariate logistic regression approach, we were able to identify a pharmacogenetic model comprising of genetic and non-genetic correlates, which had a high predictive power to discriminate between good and poor responders to antipsychotic therapy (sensitivity=75.5%, specificity=71.8%; χ2=115.13, p-value = 3.37×10-23). Our results provide credible evidence for the advantage of combined information from response-related genes and non genetic factors to form the basis of prediction tests. Predictability testing is one of the main aims of pharmacogenomic studies and this report represents an important step towards realization of this aim. These results thus underscore the need to develop an overall predictive model for antipsychotic response comprising of genotype-phenotype relationships to understand and predict the phenotypic effects of genes at the clinical level.

608T

Identification of novel polymorphisms associated with efficiency of metformin treatment in type 2 diabetes patients. *J. Klovins[1], L. Tarasova[1], L. Zarina[2], V. Pirags[2,3], OPTIMED consortium.* 1) Latvian Genome Ctr, Latviam Biomedical Ctr, Riga, Latvia; 2) Pauls Stradins Clinical University Hospital, Riga, Latvia; 3) University of Latvia, Riga, Latvia.

Many patients with type 2 diabetes (T2D) are not reaching glycemic targets by using metformin or different combinations of antidiabetic drugs, including insulin. On the other hand, intensified multidrug therapy is associated with increased risk of severe side effects and individualized selection of medication is essential. The aim of our study is to identify genetic variations associated with efficiency of metformin therapy using prospective cohort of T2D patients. 144 patients that have received metformin monotherapy were selected from OPTIMED study, a prospective clinical trial involving the treatment naive T2D patients. Treatment efficiency was estimated based on change in HbA1c levels measured 3 and 9 month after the start of the therapy. 192 tagSNPs from 15 previously reported genes influencing T2D treatment efficiency and tolerability (OCT1,2,3, MATE1,2, PMAT and others) were selected and genotyped in all patients using GoldenGate Genotyping Assay with VeraCode technology (Illumina, Inc.). We identified block of SNPs located in region between SLC22A2 -SLC22A3 genes as well as SNPs in MATE2 gene that were associated with lack of HbA1c decrease after 3 month metformin therapy. Most strongly associated SNP displayed high significance (unadjusted association test (P=9.29 e-005, OR [CI 0.95]= 4.971[2.163-11.43]). Association maintained statistical significance after Bonferroni correction and was also present when efficiency of therapy at 9 month was tested. Number of epistatic interactions was identified between the polymorphisms at OCT1-3 locus and other genes. In conclusion we have identified number of novel SNPs that may predispose to decreased metformin efficiency.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

609T

Cholesteryl ester transfer protein (CETP) polymorphisms, statin use, and their impact on cholesterol levels and cardiovascular events. *M. Leusink[1,2], N.C. Onland-Moret[2], F.W. Asselbergs[3], B. Ding[4], S. Kotti[5], N.R. van Zuydam[6], A.C. Papp[7], N. Danchin[8,9], L. Donnelly[6], A.D. Morris[6], D.I. Chasman[10,11,12], P.A.F.M Doevendans[3], O.H. Klungel[1], P.M. Ridker[10,11,12], W.H. van Gilst[13], T. Simon[5,14], F. Nyberg[4,15], C.N.A. Palmer[6], W. Sadee[7], P. van der Harst[13,16], C. Verstuyft[17,18], P.I.W. de Bakker[19], A. de Boer[1], A-H. Maitland-van der Zee[1].* 1) Division of Pharmacoepidemiology and Clinical Pharmacology, Utrecht Institute for Pharmaceutical Sciences, Faculty of Science, Utrecht University, Utrecht, the Netherlands; 2) Julius Center for Health Sciences and Primary Care, University Medical Center, Utrecht, Utrecht, The Netherlands; 3) Department of Cardiology, Division Heart and Lungs, University Medical Center Utrecht, Utrecht, The Netherlands; 4) Global Epidemiology, AstraZeneca R&D, Mölndal, Sweden; 5) Assistance Publique-Hopitaux de Paris (APHP), Hopital St. Antoine, URC-EST, Paris, France; 6) Centre for Pharmacogenomics, Medical Research Institute, Ninewells Hospital and Medical School, University of Dundee, Dundee; 7) Program in Pharmacogenomics, Department of Pharmacology, College of Medicine, The Ohio State University, Columbus, OH; 8) Assistance Publique-Hopitaux de Paris (APHP), Hopital Europeen Georges Pompidou, Paris, France; 9) Universite Paris-Descartes, Paris, France; 10) Center for Cardiovascular Disease Prevention; 11) JUPITER Trial Coordinating Center; 12) Brigham and Women's Hospital and Harvard Medical School; 13) Department of Cardiology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands; 14) Department of Clinical Pharmacology, Universite Pierre et Marie Curie (Paris 6), Paris, France; 15) Occupational and environmental medicine, Department of Public Health and Community Medicine, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, 40530 Gothenburg, Sweden; 16) Department of Genetics, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands; 17) Assistance Publique-Hopitaux de Paris (APHP), Hopital Bicetre, Service de Genetique moleculaire, Pharmacogenetique et Hormonologie, Le Kremlin Bicetre, France; 18) Universite Paris-Sud, EA4123, Le Kremlin-Bicetre, F-94275, France; 19) Department of Medical Genetics, University Medical Center Utrecht, Utrecht, The Netherlands.

Background The association of *CETP* variation with the efficacy of statins in preventing cardiovascular (CV) events has been subject of debate. We aimed to evaluate whether common SNPs influence the efficacy of statins in modifying cholesterol levels and in preventing CV events. We focused specifically on three functional SNPs, tightly linked to the production of the *CETP* Δ9 splice variant (rs5883 and rs9930761, located in exon 9 and intron 8), or *CETP* expression levels (rs3764261, located in the promoter region). The promoter SNP was also shown to influence cholesterol levels, increasing HDL-cholesterol (HDLc), and decreasing LDL-cholesterol (LDLc).

Methods Linear regression analysis was used to investigate the relation between the three *CETP* SNPs and cholesterol levels during statin treatment. We meta-analyzed the outcomes of two studies (n=10,502) with cholesterol measures using a random-effects model. The interaction between the three SNPs and statin use on CV events was analyzed using a logistic regression model with a Statin*SNP interaction term. The results of five studies with MI outcome information available (n=16,570) were meta-analyzed using a random-effects model.

Results The promoter SNP rs3764261 significantly increased HDLc by 0.02 mmol/L per T allele (95% CI 0.02–0.03, p-value: 3.1E-09), and decreased LDLc by 0.03 mmol/L per T allele (95% CI 0.01–0.04, p-value: 0.0031) during statin use. Rs3764261 showed a nominally significant interaction with statin use (interaction OR 1.19 per T allele; 95% CI 1.01–1.40, p-value for interaction: 0.0418), but this was not significant after multiple testing correction. The SNPs influencing *CETP* splicing (rs5883 and rs9930761) did not modify statin efficacy on cholesterol levels during statin treatment or CV events.

Conclusion Using a large sample size and focusing on functional *CETP* variants, we showed a further increase of HDLc and decrease of LDLc during statin treatment for carriers of the rs3764261 T variant. This effect was accompanied by a reduced protection against MI by statins when compared to non-carriers.

610T

Association study of inflammation-related genes and depressive disorders and antidepressant response. *M.-W. Lin[1, 2], Y.-T. Chang[3], I.-H. Lin[1], Y.-M. Bai[3].* 1) Institute of Public Health, National Yang-Ming University, Taipei, Taiwan; 2) Department of Medical Research and Education, Taipei Veterans General Hospital, Taipei, Taiwan; 3) Department of Psychiatry, Taipei Veterans General Hospital, Taipei, Taiwan.

Depressive disorder is a common and complex disease with multifactorial etiology. The illness has an annual prevalence of 7% and also has high lifetime prevalence, perhaps as high as 16%. Current antidepressant treatments, which mainly target on monoamine pathways, fail to achieve remission in more than 30% of depressive patients with multiple treatment trials. Increasing amount of evidence suggest that inflammation has an important role in the pathophysiology of depression. Many studies have shown that elevation in plasma levels of inflammatory cytokines in depressive patients, however, some studies have failed to replicate this finding. There is evidence demonstrated that plasma levels of inflammatory mediators are influenced by genetic polymorphisms. To investigate if inflammation-related genes might play some role in susceptibility to depressive disorders and also in the treatment response of antidepressants, we recruited 151 depression patients receiving 12-week antidepressant treatment and 135 normal control subjects in this prospective study. The levels of cytokine including IL-2, IL-6, CRP, soluble p-selectin, adiponectin, MCP-1, and TNF receptor 1 were assayed by enzyme-linked immunosobent assay (ELISA) kit at week 0, 4, 12. Genotypings of six single nucleotide polymorphisms on the CRP, P selectin, TNF-alpha, and TNF receptor 1 gene, respectively, were performed using the TaqMan Assay. The genotypic frequencies of SNPs between groups were compared using Chi-square test and continuous data were by t test or ANOVA. We found significant improvement of depressive symptoms after 4 week treatment of antidepressant, but the treatment effect did not last after 12 week. We also found the levels of soluble P selectin and TNF receptor 1 were increased as the period of antidepressant treatment increases, while the levels of adiponectin was significantly decreased after 4 week antidepressant treatment in depressive patients. Our study also revealed that the patients with CC+CT genotypes of the CRP rs1205 demonstrated significantly improvement in HAMD depressive score between week 4 and week 12, as well as the T allele carriers of the CRP rs1130864 showing significant higher HAMD depressive score than the carriers of C allele at week 0. Our results indicate that the CRP gene SNPs may be associated with improvement of depressive symptoms in depression patients. Further studies using larger sample size are warranted to confirm the findings of our study.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

611T

Are CYP2C9 and VKORC1 polymorphisms associated with increased bleeding or thrombosis risk among Blacks receiving warfarin anticoagulation? *F. Mili[1], T.A. Deans[2], W.C. Hooper[1], P. Weinstein[1], C. Lally[3], H. Austin[3], N. Wenger[2].* 1) Division of Blood Disorders, Centers for Disease Control & Prevention, Atlanta, GA; 2) Emory University School of Medicine, Atlanta, GA; 3) Department of Epidemiology, Emory University Rollins School of Public Health, Atlanta, GA.

Introduction: Numerous studies describing the association between complications of warfarin anticoagulation and CYP2C9 or VKORC1 single nucleotide polymorphisms (SNPs) among Whites have been published. Few studies assessing these associations have been performed among Blacks. Objectives: To investigate whether 4 CYP2C9 and 18 non-coding VKORC1 SNPs identified by re-sequencing are associated with serious bleeding or thrombosis risk among Blacks receiving warfarin anticoagulation. We conducted a small retrospective cohort study among Blacks ≥ age 18 years on warfarin for a minimum of 3 consecutive months. Participants were enrolled between February 2005 and October 2006 from an anticoagulation clinic in the metropolitan Atlanta area (n = 230, 124 men and 106 women). Results: Observed genotype frequencies for each of the variant alleles consistent with Hardy-Weinberg equilibrium comprised CYP2C9*2_C>T (4.2%), CYP2C9*3_A>C (1.9%), VK6915T>C (11%) and VK8773C>T (37%). X2 tests indicated an increased risk of bleeding complications among Blacks with CYP2C9*2 or CYP2C9*3 variant alleles vs. referent alleles [risk ratio (RR) = 4.3 (95% confidence interval (CI) 1.03–17.7; p = 0.03)], whereas the risk of thrombosis was increased but not statistically significant [RR = 2.6 (95% CI 0.29–23.7)]. A lower mean daily maintenance dose of warfarin was required among Blacks with heterozygous alleles for CYP2C9*2 vs. referent allele [4.8 (1.6) vs. 7.0 (4.5), t test p = 0.0046], as well as among those with heterozygous alleles for CYP2C9*3 [4.0 (0.82) vs. 6.9 (4.4), t test p = 0.0007] to maintain an international normalized ratio (INR) measurement within the therapeutic range. Although not statistically significant, the variant allele for either the VK6915 or the VK8773 SNPs was associated with a decreased risk of bleeding [RR = 0.72 (95% CI 0.27-1.9)], and no risk of thrombosis among Blacks. In addition, Blacks who are heterozygous for VK8773 seemed to need a higher mean daily maintenance dose of warfarin vs. those with the referent allele [7.1 (4.2) vs. 6.8 (4.7)]. Conclusions: Blacks with CYP2C9 variant alleles have an increased risk of bleeding versus those with the referent alleles. Our pilot study suggests that Blacks with novel VK6915 or VK8773 variant alleles have a lower risk of bleeding. Future research involving multiple ethnic groups and larger sample sizes is needed to define the role of genetic testing in assessing bleeding risk with warfarin anticoagulation.

612T

Twelve Candidate Genes and Antidepressant Response in Obsessive-Compulsive Disorder. *G. Zai[1], E. Brandl[1], J. Deluce[2], J.L. Kennedy[1], M.A. Richter[2].* 1) Neurogenetics Section, CAMH, Univ Toronto, Toronto, ON, Canada; 2) Department of Psychiatry, Sunnybrook Health Science Centre, Toronto, ON, Canada.

Obsessive-compulsive disorder (OCD) is a chronic disorder with a strong genetic component. Genetic associations between OCD and several candidate genes including glutamate transporter (SLC1A1), monoamine oxidase (MAOA), glutamate receptor, ionotropic, N-methyl D-aspartate 2B (GRIN2B), serotonin 2A receptor (5HTR2A), serotonin transporter (SLC6A4), and catecholamine-O-methyl-transferase (COMT) genes have been reported albeit with inconsistent results. Pharmacogenetics represents an important alternate to investigate inter-individual genetic variation and drug response. In this preliminary study, we investigated 12 different genes including those mentioned above in addition to the disks large (drosophila) homolog-associated protein 2 (DLGAP2), myelin oligodendrocyte glycoprotein (MOG), serotonin 1B receptor (5HT1B), chromosome 9 open reading frame 68 (C9orf68), adenosine deaminase, RNA-specific, B2 (ADARB2), and oligodendrocyte lineage transcription factor 2 (OLIG2) genes. Thirty, two, six, sixteen, eleven, six, two, ten, sixteen, ten, ten, and ten single nucleotide polymorphisms (SNPs) in the DLGAP2, MOG, 5HT1B, SLC1A1, C9orf68, MAOA, ADARB2, GRIN2B, 5HTR2A, SLC6A4, OLIG2, and COMT genes respectively were genotyped in 117 individuals with OCD and retrospective response data collected on multiple serotonin reuptake inhibitor (SRI) trials. Individuals were grouped into those who improved following an adequate trial of one or more SRI(s) as compared with those who reported “minimal”, “no change”, or “worsening” in response to SRI(s) tried. Genotypes and response data were examined by exploratory analyses on a drug-by-drug and combined basis. Significant response associations were detected in DLGAP2 & paroxetine/clomipramine (P=0.008–0.042), 5HT1B & clomipramine/SRI (P=0.0003–0.045), SLC1A1 & sertraline/fluvoxamine/citalopram/SSRI/SRI (P=0.0008–0.035), C9orf68 & clomipramine/fluvoxamine/SSRI (P=0.007–0.031), MAOA & citalopram (P=0.028), GRIN2B & fluoxetine/paroxetine/fluvoxamine/citalopram/clomipramine (P=0.004–0.024), 5HTR2A & clomipramine (P=0.024), SLC6A4 & paroxetine (P=0.0006–0.010), OLIG2 & paroxetine (P=0.008), and COMT & paroxetine/sertraline/citalopram/SSRI/SRI (P=0.007–0.044). Analyses for association with response to other drug trials were negative. These results suggest that genetic variants may play an important role in SRI response to OCD. However, replication in larger and independent samples is required.

613T

A genome-wide integrative study of microRNAs in human liver. *W. Liu[1,2], F. Innocenti[2,3], E. Gamazon[2], S. Mirkov[2], J. Ramirez[2], S. Huang[2], N. Cox[2], M. Ratain[2].* 1) Department of Medicinal Chemistry and Molecular Pharmacology, Purdue University, West Lafayette, IN; 2) Department of Medicine, The University of Chicago, Chicago, IL; 3) Institute for Pharmacogenomics & Individualized Therapy, University of North Carolina, Chapel Hill, NC.

Recent studies have illuminated the diversity of roles for microRNAs in cellular, developmental, and pathophysiological processes. Their biological consequences are actively being explored, and their functions in the initiation and progression of cancer, the regulation of cellular activity from proliferation to apoptosis, and the modulation of therapeutic response are rapidly being uncovered. The study of microRNAs in human liver tissue promises to clarify the therapeutic and diagnostic value of this important regulatory mechanism of gene expression. To this end, we conducted genome-wide profiling of microRNA expression in liver and identified hundreds of microRNAs abundantly expressed, including those significantly associated (FDR<0.05) with mRNA expression. In particular, we report here the list of Very Important Pharmacogenes (VIP Genes), comprising of genes of particular relevance for pharmacogenomics, that are under substantial microRNA regulatory effect in liver relative to genomic background. We set out to elucidate the genetic basis of microRNA expression variation in liver and mapped microRNA expression to genomic loci as miRNA expression quantitative trait loci (miR-eQTLs). We identified common single nucleotide polymorphisms (minor allele frequency >15%;) that attain genome-wide significant association ($p<10^{-11}$) with microRNA expression (based on Bonferroni adjustment for the number of SNPs and the number of expressed microRNAs tested). A comparison of eQTLs from a genome-wide transcriptome study in liver and the miR-eQTLs clarifies the role of miR-eQTLs on (mRNA) gene regulation. We found that the miR-eQTLs are significantly more likely to predict mRNA levels at a range of p-value thresholds than a random set of allele frequency matched SNPs, showing the functional effect of these loci on the transcriptome. Finally, we show that annotating SNPs reproducibly associated with complex traits from the NHGRI repository of published genome-wide association studies as well as variants from a comprehensive catalog of manually curated pharmacogenetic associations with information on microRNA expression is likely to advance our understanding of disease pathogenesis, therapeutic outcome, and other complex human phenotypes.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

614T

Protein Quantitative Trait Loci Identify Novel Proteins Predictive for Chemotherapeutic Induced Phenotypes. *A.L. Stark[2], R.J. Hause, Jr[3,4,5], L.K. Gorsic[2], N.N. Antao[2], S.S. Wong[2], D.F. Gill[3], S.H. Chung[3], S.M. Delaney[2], C.D. Brown[1,3,5], K.P. White[1,3,5], R.B. Jones[3,4,5], M.E. Dolan[2,6]*. 1) Dept Human Gen, Univ Chicago, Chicago, IL; 2) Dept of Medicine, Univ Chicago, Chicago, IL; 3) Committee on Genetics, Genomics, and Systems Biology, Univ Chicago, Chicago, IL; 4) Ben May Dept for Cancer Research, Univ Chicago, Chicago, IL; 5) Institute for Genomics and Systems Biology, Univ Chicago, Chicago, IL; 6) Committee on Clinical Pharmacology and Pharmacogenomics, Univ Chicago, Chicago, IL.

Personalized medicine aims to develop predictive markers for oncology patients at greatest risk of experiencing chemotherapeutic related toxicities and/or nonresponse. Many of the biomarkers are proteins and by identifying protein quantitative trait loci (pQTLs), genetics and genotyping can predict protein levels. Our lab has developed models that employ HapMap lymphoblastoid cell lines (LCLs) to identify chemotherapeutic susceptibility genetic variants by evaluating chemotherapeutic induced phenotypes. To identify pQTLs as well as proteins that contribute to drug-induced phenotypes, we used newly developed micro-western arrays to assay protein levels for over 400 transcription factors and core signaling proteins in 68 unrelated Yoruba derived LCLs. In order to demonstrate the utility of these protein levels, we focused on two widely used chemotherapeutics (paclitaxel and cisplatin) with two cellular phenotypes assayed (caspase 3/7 apoptosis and cell growth inhibition). These two phenotypes are negatively correlated, but varied in strength (paclitaxel r2=0.35, cisplatin r2=0.04). For each drug, we then analyzed the two phenotypes GWAS for pQTLs ($p<0.0001$) and identified a common protein implicated through different pQTLs for apoptosis and cell growth inhibition phenotypes. ZNF569 protein levels were negatively correlated with paclitaxel-induced growth inhibition ($p=0.0005$) and positively correlated with paclitaxel-induced apoptosis ($p=0.0002$). The pQTLs of ZNF569 associated with paclitaxel-induced growth inhibition and apoptosis were on different chromosomes suggesting that different regulatory architecture converges on common proteins in the related paclitaxel phenotypes. Similarly for cisplatin, SMC1A protein levels were positively correlated for growth inhibition ($p=0.004$) and negatively correlated for apoptosis ($p=0.0007$). Again, the pQTLs of SMC1A associated with cisplatin growth inhibition and apoptosis were on different chromosomes. Using siRNA, we functionally validated both ZNF569 ($p<0.002$, 0.0001) and SMC1A ($p<0.0002$, 0.001) for apoptosis and growth inhibition induced by paclitaxel and cisplatin, respectively. This work allows pharmacogenomic discovery work to move from the transcriptome to the proteome and demonstrates the cellular utility of pQTLs. A better understanding of the variation in the proteome can aid in the discovery of genetic predictors of chemotherapy toxicity and response.

615T

Functional genetic variation in *RHOA* is associated with the cholesterol-lowering effects of statin treatment. *E. Theusch[1], D. Naidoo[1], F. Bauzon[1], K. Stevens[1], L. Mangravite[2], Y.L. Kuang[1], R.M. Krauss[1], M.W. Medina[1]*. 1) Children's Hospital Oakland Research Institute (CHORI), Oakland, CA; 2) Sage Bionetworks, Seattle, WA.

Statins are drugs designed to reduce plasma LDL-cholesterol (LDL-C) levels and cardiovascular disease risk by inhibiting HMG Co-A reductase (HMGCR), the rate-limiting cholesterol biosynthesis enzyme. There is substantial inter-individual variation in statin-induced LDL-C lowering, but few genetic loci modulating statin response have been identified using candidate gene and genome-wide association studies. As an alternate approach, we used genome-wide gene expression data from human livers and sham- and statin-exposed hepatoma cell lines and lymphoblastoid cell lines (LCLs) to discover liver-expressed, statin-responsive genes that had statin-induced expression changes strongly correlated with changes in *HMGCR* expression. This filtering approach identified ras homolog family member A (*RHOA*), a gene previously shown to mediate anti-inflammatory but not cholesterol responses to statin treatment, as a candidate for future study ($r^2=0.13$, $p=7.64\times10^{-16}$ with statin-induced *HMGCR* expression changes). *RHOA* knockdown in hepatoma cell lines decreased mRNA levels of known sterol-regulated genes (*HMGCR*, *SREBF2*, and *LDLR*), increased intracellular cholesterol ester, and increased media concentrations of APOB, the major apolipoprotein on LDL particles. Using genotype data and plasma cholesterol measurements from Caucasian participants of the Cholesterol and Pharmacogenetics (CAP) simvastatin clinical trial (N=580) and the PRINCE pravastatin clinical trial (N=1307), we found that of the four common haplotypes encompassing *RHOA*, haplotype H3B was associated with reduced *in vivo* statin-induced plasma LDL-C lowering ($p=0.0149$), while haplotype H2 was associated with increased LDL-C lowering ($p=0.0349$). H3B was the only haplotype containing the minor allele of rs11716445, a SNP located in a rare *RHOA* exon (exon 2.5) identified using RNA-Seq data. Furthermore, we concluded that rs11716445 was a cis-splicing quantitative trait locus (cis-sQTL) for *RHOA* exon 2.5 because the minor allele dramatically increased the inclusion of exon 2.5 in *RHOA* transcripts from CAP LCLs, and, in rs11716445 heterozygotes, over 90% of transcripts incorporating exon 2.5 contained the minor allele. Since *RHOA* lies in a large linkage disequilibrium block, it is possible that genetic variants in other genes in the region influence statin response, but these findings suggest that functional genetic variation in *RHOA* contributes to variation in LDL-C response to statin treatment.

616T

The modulation of CYP3A4 expression by the efflux pump ABCB1 in HepG2 cells. *W. Shou[1, 2], B. Wang[2], Z. Wang[2], J. Shi[2], W. Huang[1, 2]*. 1) Rui Jin Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China; 2) Department of Genetics, Shanghai-MOST Key Laboratory of Health and Disease Genomics, Chinese National Human Genome Center, Shanghai, China.

CYP3A4, the most abundant member of Cytochrome P450 family, is involved in the metabolism of nearly 50% of drugs currently prescribed. CYP3A4 expression is strongly modulated by nuclear receptors (NR), e.g. pregnane X receptor (PXR) and Hepatocyte nuclear factor 4 alpha (HNF4α), and by exogenous or endogenous chemicals through NR-mediated induction. Many CYP3A4 modulators are also the substrates of P-glycoprotein (ABCB1), which is an important transmembrane efflux transporter. Thus, we hypothesized that ABCB1 could regulate CYP3A4 expression. To prove this hypothesis, we constructed transient and stable transfected HepG2 cells with ABCB1 expressing plasmid, and adopted small interfering RNA (SiRNA) targeting ABCB1 expression in HepG2 cells, in order to make ABCB1 overexpressing or inhibited. HepG2 cells were cotransfected with expression plasmids of nuclear receptors (PXR and HNF4α), and treated with classic CYP3A4 inducer rifampicin after transfection. Then, luciferase reporter for CYP3A4 was assayed. CYP3A4 expression was induced by rifampicin in a concentration-dependent manner. CYP3A4 reporter activity significantly decreased in ABCB1-overexpressing cells, compared to non-overexpressing cells, especially in the stable transfected cells. In addition, CYP3A4 reporter activity was significantly higher in ABCB1 inhibited cells than in negative controls, which confirmed the results in ABCB1-overexpressing cells. The two-sided results revealed that ABCB1 indirectly modulated the CYP3A4 gene expression through regulating the intracellular modulators of the CYP3A4 gene expression. Our study suggested that the efflux pump ABCB1 might be an upstream regulator of the important drug-metabolizing enzyme CYP3A4. Consequently, gene-drug and gene-gene interaction need to be carefully considered during drug metabolism.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

617T

Genetic determinants of dabigatran plasma levels and their relation to bleeding. *G. Pare[1], N.E. Eriksson[2], T. Lehr[3], S. Connolly[1], J. Eikelboom[1], M.D. Ezekowitz[4], T. Axelsson[5], S. Haertter[3], J. Oldgren[2], P. Reilly[7], A. Siegbahn[5], A-C. Syvanen[5], C. Wadelius[6], M. Wadelius[5], H. Zimdahl-Gelling[3], S. Yusuf[1], L. Wallentin[2]*. 1) Population Health Research Institute, Hamilton Health Sciences, McMaster University, Hamilton, ON, Canada; 2) Uppsala Clinical Research Center and Department of Medical Sciences, Uppsala University, Uppsala, Sweden; 3) Boehringer Ingelheim Pharma GmbH ξ Co KG, Biberach, Germany; 4) Thomas Jefferson Medical College, Cardiovascular Research Foundation, Philadelphia, PA, USA; 5) Department of Medical Sciences, Uppsala University, Uppsala, Sweden; 6) Department of Immunology, Genetics and Pathology, Science for Life Laboratory, Uppsala University, Sweden; 7) Boehringer Ingelheim Pharma Inc., Ridgefield, CT, USA.

Background: Fixed-dose unmonitored treatment with dabigatran etexilate is effective and has a favorable safety profile in prevention of stroke in atrial fibrillation patients compared to warfarin. We hypothesized that genetic variants could contribute to inter-individual variability in blood concentrations of the active metabolite of dabigatran etexilate, and influence the safety and efficacy of dabigatran. [au0,2]Methods[xu: We successfully conducted a genome-wide association study in 2,944 RE-LY participants. Single nucleotide polymorphisms (SNP) correlated with dabigatran peak and trough concentrations were tested for an association with reported bleeding event (N= 587), major bleeding (N=101) and ischemic events (N=32).

Results: Each minor allele of the CES1 SNP rs2244613 was associated with lower trough concentrations (15% decrease per allele, 95%CI 10–19%; P=1.2 × 10-8) and a lower risk of any bleeding (OR=0.67, 95%CI 0.55–0.82; P=7 × 10-5) in dabigatran-treated participants, with a consistent but non-significant lower risk of major bleeding (OR=0.66, 95%CI 0.43–1.01). The interaction between treatment (warfarin versus all dabigatran) and carrier status was statistically significant (P=0.002) with carriers having less bleeding with dabigatran than warfarin (HR=0.59, 95%CI 0.46–0.76; P=5.2 × 10-5) in contrast to no difference in noncarriers (HR=0.96, 95%CI 0.81–1.14; P=0.65). There was no association with ischemic events. The ABCB1 SNP rs4148738 and CES1 SNP rs8192935 were associated with peak concentrations but neither was associated with bleeding or ischemic events.

Conclusions: Carriage of CES1 rs2244613 minor allele occurred in 32.8% of patients in RELY and was associated with lower dabigatran exposure. The presence of the polymorphism was associated with a lower risk of bleeding.

618T

Genome-wide study of resistant hypertension using existing genomic data and electronic medical records. *L. Dumitrescu[1], M.D. Ritchie[2], J.C. Denny[3,4], S.J. Bielinski[5], P. Peissig[6], J.A. Pacheco[7], M.G. Hayes[8], G.P. Jarvik[9], R. Li[10], I.J. Kullo[11], C.G. Chute[12], R.L. Chisholm[7], E.B. Larson[13], C.A. McCarty[14], D.M. Roden[4,15], M. de Andrade[12], D.C. Crawford[1,16] on behalf of the eMERGE Network.* 1) Center for Human Genetics Research, Vanderbilt University, Nashville, TN; 2) Department of Biochemistry and Molecular Biology, The Huck Institutes of the Life Sciences, Eberly College of Science, Penn State University, University Park, PA; 3) Department of Biomedical Informatics, Vanderbilt University, Nashville, TN; 4) Department of Medicine, Vanderbilt University, Nashville, TN; 5) Division of Epidemiology, Department of Health Sciences Research, Mayo Clinic, Rochester, MN; 6) Biomedical Informatics Research Center, Marshfield Clinic Research Foundation, Marshfield, WI; 7) Center for Genetic Medicine, Northwestern University, Chicago, IL; 8) Department of Medicine, Northwestern University, Chicago, IL; 9) Department of Medicine, University of Washington Medical Center, Seattle WA; 10) Office of Population Genomics, National Human Genome Research Institute, Bethesda, MD; 11) Division of Cardiovascular Diseases, Department of Medicine, Mayo Clinic, Rochester, MN; 12) Biomedical Statistics and Informatics, Department of Health Sciences Research, Mayo Clinic, Rochester, MN; 13) Group Health Research Institute, Seattle WA; 14) Center for Human Genetics, Marshfield Clinic Research Foundation, Marshfield, WI; 15) Department of Pharmacology, Vanderbilt University, Nashville, TN; 16) Departments Molecular Physiology and Biophysics, Vanderbilt University, Nashville, TN.

Resistant hypertension (RH) is defined as high blood pressure (BP) that remains above treatment goals in spite of the concurrent use of at least three antihypertensive agents of different classes. Despite the increasing prevalence of RH, genetic assessments of patients with RH are limited. Large-scale DNA biobanks linked to electronic medical records (EMRs) may enable researchers to extract patients with RH to identify novel genotype-phenotype associations using existing genomic data. The first phase of the Electronic Medical Records and Genomics (eMERGE) Network was a collaboration of five EMR-linked biobanks. For each study site, GWAS were conducted for a number of common diseases and traits, including RH. Individuals of European (E)-descent or African (A)-descent were genotyped on Illumnia 660W-Quad or 1M, respectively. Electronic selection logic using billing codes, laboratory values, text queries, and medication records was used to identify RH cases and controls, which were then validated by manual review. Cases were defined by concurrent use of four or more antihypertension medication classes or concurrent use of three medication classes with elevated BP. Controls consisted of normotensives and/or hypertensives controlled on one antihypertensive medication. Single SNP tests of association using logistic regression were performed. Models were stratified by race/ethnicity and were adjusted for sex, decade of birth, and global ancestry. Among E-descent RH cases (n=1,648) and combined normotensive and controlled hypertensive controls (n=2,365), no SNP was associated with RH at genome-wide significance (p>1E-06). When E-descent cases were compared to controlled hypertensive controls (n=718) alone, the most significant association was with rs12135020 (p=1.98E-06; OR=0.61) found 8kb upstream of *CSRP1*, which is important in the development of arterial smooth muscle cells. Among cases (n=595) and combined controls (n=527) of A-descent, intergenic rs6677984 was associated with RH at genome-wide significance (p=2.37E-14; OR=0.41). We then examined whether SNPs known to influence BP or hypertension also influenced RH. Only 3/26 and 5/42 of the previously-identified SNPs examined in E- and A-descent populations, respectively, were associated with RH at p<0.05. These data suggest that some genetic variants associated with RH are independent of those known to be associated with BP and/or hypertension and highlight the utility of EMR-linked genomic data.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

619T

Genetic and clinical determinants of warfarin dose - a RE-LY genomics substudy. *N. Eriksson[1,3], L. Wallentin[1,3], L. Berglund[1], T. Axelsson[3], S. Connolly[2], M. Ezekowitz[4], J. Oldgren[1], P. Guillaume[2], P. Reilly[5], A. Siegbahn[3], A-C. Syvanaen[3], C. Wadelius[6], S. Yusuf[2], M. Wadelius[3]*. 1) Uppsala Clinical Research Center, Uppsala, Uppsala, Sweden; 2) Population Health Research Institute, Hamilton Health Sciences and McMaster University, Hamilton, ON, Canada; 3) Department Medical Sciences, Uppsala University, Uppsala, Sweden; 4) Thomas Jefferson Medical College, Cardiovascular Research Foundation; 5) Boehringer Ingelheim Pharma Inc; 6) Department Genetics and Pathology, Uppsala University, Sweden.

Purpose: The response to warfarin treatment is influenced by several genetic factors. We conducted a genome-wide association study in patients from the RE-LY study, which compared two doses of dabigatran with warfarin. We hypothesized that genetic risk factors, other than CYP2C9, VKORC1 and CYP4F2, might affect warfarin maintenance dose and percent time in therapeutic range (TTR). We also investigated the effects of VKORC1 and CYP2C9 on TTR and risk of major bleeding. Further, we investigated the performance of the International Warfarin Pharmacogenetics Consortium (IWPC) prediction model and if any additional variables (renal function, smoking, CYP4F2) might improve performance of future prediction models. **Methods:** 3,076 participants in the RE-LY genomics sub study, of which 982 were in the warfarin arm, were genotyped using the Illumina Human610-quad chip. Following quality control, 554,725 SNPs in 956 patients were available for analysis. A p-value threshold of genome-wide significance was set at 9 × 10-8. To account for possible population stratification, all genome-wide analyses were adjusted for the first four genetic principal components. TTR was calculated according to the Rosendaal method. GWAS of continuous traits was performed using linear regression. **Results:** Novel SNPs associated with warfarin dose (adjusted by clinical factors and VKORC1/CYP2C9) at a genome-wide significance level were found in the genes DDHD1 and NEDD4. We identified a SNP in the gene ASPH that was associated with TTR (-6.8 % per minor allele). During the first three months, VKORC1 rs9923231 was associated with INR outside treatment range (2.0–3.0) and CYPC29 *3 was associated with INR above 3.0. Neither CYP2C9 *2/*3 nor VKORC1 rs9923231 were associated with risk of major bleeding. The IWPC model gave an R^2 of 50.9 % and a MAE of 6.6 mg/week. The predictive value increased slightly when CYP4F2, DDHD1, NEDD4 and Creatinine Clearance (CrCl) were added to a model including the IWPC variables (added R^2 ~1 % per SNP, 0.4 % for CrCl), whereas smoking was not associated with dose. **Conclusions:** We verified the importance of VKORC1 and CYP2C9 for warfarin dose and the predictive value of the IWPC model for dose selection. We identified two novel SNPs in DDHD1 and NEDD4 affecting warfarin maintenance dose, but the incremental information provided by them and by adding Creatinine Clearance is probably limited. We also identified a SNP in ASPH associated with total TTR.

620T

The Kaiser Permanente/UCSF Genetic Epidemiology Research Study on Adult Health and Aging: A Genome-wide association study of quetiapine-induced weight gain derived from electronic medical records and outpatient pharmacy data. *N. Gonzaludo[1], S. Sciortino[2], L. Walter[2], N. Risch[3], P.Y. Kwok[3], C. Schaefer[2]*. 1) Dept of Bioengineering & Therapeutic Sciences, UCSF, San Francisco, CA; 2) Kaiser Permanente Northern California Division of Research, Oakland, CA; 3) Institute for Human Genetics, UCSF, San Francisco, CA.

Atypical antipsychotics are widely prescribed to treat schizophrenia, bipolar disorder, and other psychiatric conditions. Weight gain is a common side effect of many atypical antipsychotics and may contribute to the onset of diabetes and lipid disorders, along with increasing the likelihood of non-compliance. While previous studies have identified genetic factors underlying this phenotype, such studies have been limited to small sample sizes and primarily used candidate gene approaches. Since weight is commonly measured during patient visits, electronic medical records provide a means of capturing this side effect in a clinical setting across a larger and broader range of patients. As part of the Kaiser Permanente Research Program on Genes, Environment, and Health, 110,266 members have undergone genome-wide genotyping using Affymetrix Axiom technology. These subjects constitute the Genetic Epidemiology Research on Adult Health and Aging (GERA) cohort. With data from this cohort, we are conducting a genome wide association study (GWAS) for weight gain induced by Seroquel (quetiapine), the most commonly prescribed atypical antipsychotic in the Kaiser Permanente Northern California system. Of the 110,266 subjects enrolled in the GERA cohort, 1,468 had been given at least one outpatient prescription of quetiapine in the past 10 years, with 677 patients on lower doses of quetiapine (100 mg or less daily). Of these patients, 441 had at least one weight and/or BMI measurement taken during a treatment period and at least two measurements taken up to one year prior. 48% of patients showed an increase in BMI, with a mean increase of 0.9 BMI. Of those that gained weight, 11% gained greater than 7% baseline BMI (5% over all patients). For these 441 patients, over 650K SNPs from the Axiom arrays were tested for an association with overall BMI gain using linear regression, as well as gain greater than 7% from baseline using logistic regression. Covariates for the association test included age, sex, prior atypical antipsychotic use, baseline BMI, and a diagnosis of Type 2 Diabetes. We compared our findings to known candidate genes for weight gain induced by atypical antipsychotics, including cytochrome P450 enzymes, drug targets, and energy homeostasis genes. Our study demonstrates the value of this resource for extracting robust phenotypic data and identifying genetic components underlying clinically observed drug response.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

621T

Genome wide association analysis of pain reduction in rheumatoid arthritis patients treated with anti-TNF medication. *M. Umicevic Mirkov[1], H. Scheffer[1], S. Krintel[2], S.H. Vermeulen[1, 3], J. Johansen[2], W. Kievit[4, 6], M.A.F. van de Laar[5, 6], P.C.L. van Riel[4, 6], B. Franke[1], M. Hetland[2], M.J.H. Coenen[1]*. 1) Department of Human Genetics, Radboud University Nijmegen Medical Centre, Nijmegen, Netherlands; 2) Department of Rheumatology, Copenhagen University Hospital, Glostrup, Denmark; 3) Department of Epidemiology, Biostatistics and HTA, Radboud University Nijmegen Medical Centre, Nijmegen, Netherlands; 4) Departments of Rheumatology, Radboud University Nijmegen Medical Centre, Nijmegen, Netherlands; 5) University Twente & Medisch Spectrum Twente, Enschede, Netherlands; 6) on behalf of the Dutch Rheumatoid Arthritis Monitoring registry.

Background Rheumatoid arthritis (RA) patients rate pain relief as the highest priority in treatment. Treatment strategies blocking tumor necrosis factor (anti-TNF) have proven very successful in pain reduction. Interestingly, recent research indicates that inflammation and pain pathways are important, partly independent, targets of anti-TNF. Therefore genetic factors predicting pain relief are important to personalize treatment of RA patients. **Objective** We aimed to identify genetic factors predicting pain reduction upon anti-TNF treatment in patients with RA using a genome-wide association approach. **Methods** We included 508 RA patients treated with anti-TNF agents from the Dutch Rheumatoid Arthritis Monitoring (DREAM) registry. Single nucleotide polymorphism (SNP) markers were genotyped using the Illumina HumanHap550-Duo or Human660W-Quad BeadChip. Association analysis using the Visual Analogue Pain Scale (VASpain) change 14 weeks after treatment initiation as outcome was performed on an imputed dataset under an additive genetic model with adjustment for base line VAS-pain. The top markers demonstrating association with VASpain change (p-value<10^{-3}) were selected for replication in 207 RA patients treated with anti-TNF from the Danish DANBIO registry. We performed a meta-analysis using METAL and pathway analysis using Ingenuity. **Results** 2,557,253 SNPs and 406 patients passed quality control. No findings reached the threshold for genome-wide significance (p-value≤1×10^{-8}) in the discovery cohort. Meta-analysis with the Danish sample led to the identification of three SNPs with a p-value <1×10^{-7}. These SNPs can be linked to functions that might involve pain processing in the brain and are located within or nearby *ATXN1*, a gene involved in the pathology of spinocerebellar ataxia type 1, a neurodegenerative disorder and *NRG3* which influences neuroblast proliferation, migration and differentiation. Pathway analysis, including all SNPs with a p-value<10^{-3} (n=2649) from the discovery cohort also point to the brain. The analysis shows that genes involved in neuritogenesis are overrepresented in our dataset (p=6.78×10^{-6}). In addition, six of the identified SNPs map to the α-adrenergic signaling pathway (p=6.16×10^{-4}), an important target for pain medication. **Conclusions** Significant findings will be replicated in a third patient population. Confirmed biomarkers can be used to personalize medication for the individual patient.

622T

A Genome-Wide Association Analysis of Corticosteroid Response in Asthmatics. *Q. Duan[1], B.E. Himes[1,3], B. Schuemann[1], K.G. Tantisira[1,2], S.T. Weiss[1,3]*. 1) Channing Laboratory, Brigham & Women's Hospital and Harvard Medical School, Boston, MA; 2) Pulmonary Division, Brigham & Women's Hospital and Harvard Medical School, Boston, MA; 3) Partners Center for Personalized Genetic Medicine, Partners Health Care, Boston, MA.

Corticosteroids (CS) are the most widely used, anti-inflammatory treatment for controlling asthma, which affects approximately 300 million individuals worldwide and over 20 million Americans. However, great inter-individual variability in response to this class of drugs has been described, whereby up to 25% of patients are non-responsive. CS response is a complex trait that is partially regulated by genetic factors. In this study, the physiological response to CS is measured as a percent change in baseline FEV_1 (forced expiratory volume in one second) following 4 to 8 weeks of treatment with an inhaled CS. A genome-wide association study (GWAS) of this quantitative trait was performed in 702 non-Hispanic white asthmatic subjects from four asthma clinical trials: the Childhood Asthma Management Program (CAMP), the Leukotriene Modifier or Corticosteroid Salmeterol study (LOCCS), the Childhood Asthma Research and Education (CARE) network and the Asthma Clinical Research Network (ACRN). We used genotype data for CAMP and LOCCS from the Illumina 550K and 610K Beadarrays and imputed SNP data for ACRN and CARE (based on the 1000 Genomes Project) that had been genotyped on the Affymetrix 6.0 array. A total of 557,367 single nucleotide polymorphisms (SNPs) were tested for association with CS response using a linear regression model in PLINK with adjustments for age, sex, and height. Combined P-values were calculated using the Liptak's weighted Z method. Two of the top GWAS SNPs (P=3.17E-07 and 1.06E-05) were located in or near the *PLXDC1* gene, whose (RNA) expression has been previously correlated with asthma. Another of the top five SNPs (2.93E-06) flanks the *GALNT7* gene, the RNA expression of which has also been associated with asthma and SNPs at this locus, which regulate gene expression, have been previously associated with airway obstruction. Replication of these genetic associations is underway in white subjects from an adult asthma trial population previously ascertained by Forest Inc. In summary, this GWAS of CS response identified novel genetic loci associated with this variable drug response. Further replication analysis in independent asthma populations and functional characterization of the replicated loci is necessary to better understand the biological mechanisms underlying differential response to CS therapy. Funding sources: NHLBI R01 HL092197 and U01 HL65899.

623T

MicroRNA-based approach expands on genome-wide association study and identifies SNPs important for response to glucocorticoids in asthma patients. *H. Im[1], E.R. Gamazon[2], D. Lenkala[2], K. Wu[2], G. Clemmer[3], S.T. Weiss[3], N.J. Cox[2], K. Tantisira[3], R.S. Huang[2]*. 1) Health Studies, University of Chicago, Chicago, Il; 2) Medicine, University of Chicago, Chicago, Il; 3) Channing Laboratory, Brigham and Women's Hospital and Harvard Medical School, Boston MA.

Our objective is to take a concerted translational effort to elucidate the genetic etiology of sensitivity to glucocorticoids (GCs), which are some of the most commonly used agents in treating inflammatory diseases. Our previous work demonstrated that by integrating gene expression into genome-wide association studies (GWASs), one can identify genetic variations that are functionally relevant to the phenotype of interest. Here we expanded this integrative genomic approach to include the use of genome-wide expression of microRNAs (miRNAs), a set of small non-coding RNAs that play an important role in diverse biological processes, and applied this approach to patients from the Childhood Asthma Management Program (CAMP), a randomized trial of asthma treatment. Utilizing a linear mixed effects model, we characterized the longitudinal phenotypic data (repeated measurement of FEV1, an indicator of lung function every 4 months for 4 years) from CAMP. The log ratio of the model-predicted FEV1 at baseline under budesonide (an inhaled GC) treatment and the observed FEV1 at baseline was used as the response phenotype of interest. Genome-wide SNP data were obtained from budesonide-treated subjects using Illumina 550K and 610K arrays. We found 12 SNPs that are associated with GC sensitivity at FDR<0.25. The most significantly associated SNP (p=2.8E-8) is also associated with the expression of 8 miRNAs, many of which have been implicated in airway function and inflammatory pathway. This study suggests that genetic variation may influence miRNA expression levels with important phenotypic effect on response to GC in asthma patients.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

624T

SLC9A9 variant as a novel biomarker of response to interferon beta: results from a pharmacogenetic study on Italian, US and French Multiple Sclerosis patients. *F. Martinelli Boneschi[1,2], F. Esposito[1,2], E. Lim[3,4], P. Brambilla[2], M. Sorosina[2], M. Romeo[1], M. Rodegher[1], N. Bohossian[5], N.A. Patsopoulos[6,7], B. Kenan[3,4], V. Martinelli[1], D. Brassat[5], G. Comi[1,2], P. De Jager[6,7]*. 1) Department of Neurology, Scientific Institute San Raffaele, Milan, Italy; 2) Laboratory of Genetics of Neurological Complex Diseases, CNS Inflammatory Unit, Institute of Experimental Neurology, Scientific Institute San Raffaele, Milan, Italy; 3) Analytic and Translational Genetics Unit, Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, MA; 4) Program in Medical and Population Genetics, Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, MA; 5) University of Toulouse, Toulouse, France; 6) Program in Translational NeuroPsychiatric Genomics, Institute for the Neurosciences, Department of Neurology, Brigham & Women's Hospital, Boston, MA; 7) Harvard Medical School, Boston, MA.

Background: Despite the broad variety of drugs approved for multiple sclerosis (MS), a large proportion of patients still experience the persistence of clinical and MRI disease activity while on treatment. The early identification of the most effective MS treatment in individuals is critical to impact the long-term outcome and to move toward a more personalized approach in MS. Objective: to identify genetic variants able to predict the response to interferon beta (IFNb). Materials and methods: A cohort of Italian relapsing-remitting (RR) MS patients treated with IFNb for at least 6 months and with a 2-year follow-up were enrolled in the study. The response to IFNb was assessed using clinical and MRI measures that correlate with disease activity. A genome-wide association study was performed to compare responders vs non-responders. The top genetic variant was replicated in three independent collections of US, Italian and French patients. Results: The rs9828519^{G} allele is associated with an increased risk of non-response to IFNb therapy in the Italian MS patients (p-value=4.43×10^8). The proportion of non-responders increases for each additional copy of the G allele that is carried, suggesting an additive effect. The same allele and the same direction of effect were confirmed in an independent collection of US MS samples (p-value=5.97×10^4). On the contrary, the genetic variant did not appear to influence the response to glatiramer acetate in the US dataset (p-value= 0.14). The same direction of effect was observed also in additional Italian and French patients, even if the results were not statistically significant within each dataset. A meta-analysis across the 4 different datasets confirmed the effect of the rs9828519 (p-value=8.39×10^7). Rs9828519 is intronic to SLC9A9, which has been recently reported to impair protein sialinization. Conclusions: Rs9828519^{G} affects the response to IFNb but not to glatiramer acetate, suggesting that this genetic variant would be a predictive marker of non-response to IFNb. Functional studies are ongoing, in order to provide further evidence of the effect of this variant on influencing the alternative splicing of the gene and to explain its effect on IFN treatment response.

625T

Polygenic Heritability Estimates of Asthma and Treatment-Related Phenotypes. *M.J. McGeachie[1,2,3], E.A. Stahl[3,4,5], B.E. Himes[1,2,3], S.A. Pendergrass[6], M.D. Ritchie[6], R.M. Plenge[3,4,5], K.G. Tantisira[1,2,3]*. 1) Channing Laboratory, Dept of Medicine, Brigham and Women's Hospital, Boston, MA; 2) Partners Healthcare Center for Personalized Genetic Medicine, Boston, MA; 3) Harvard Medical School, Boston, MA; 4) Division of Rheumatology, Immunology and Allergy, Department of Medicine, Brigham and Women's Hospital, Boston, MA; 5) Program in Medical and Population Genetics, Broad Institute, Cambridge, MA; 6) Center for Systems Genomics, Department of Biochemistry and Molecular Biology, Pennsylvania State University, University Park, PA.

BACKGROUND. We estimated the heritability of pharmacogenetic phenotypes related to asthma, including bronchodilator response, steroid response, airway hyperresponsiveness, and also asthma itself, using the polygenic linear mixed modeling approach of Yang et al. (2010). This method estimates the contribution of additive SNP effects to phenotype variance for genotypic data from existing asthma cohorts (ACRN, CARE, CAMP, LOCCS, LODO, and Sepracor). This method has previously been successfully applied to estimates of the genetic contribution to human height, schizophrenia, Crohn's disease, and type I diabetes, among others. RATIONALE. Our analyses address two important issues for elucidating the connection between genes and etiology of asthma through providing evidence that the heritability in asthma is due to simple additive effects of SNPs and through identifying substantial genetic components contributing to variance in bronchodilator response and airway hyperresponsiveness. METHODS. Bronchodilator response (BDR) was measured by FEV1 change before and after the administration of a bronchodilator (albuterol); where BDR was defined as the change in FEV1 as a percentage of the pre-bronchodilator FEV1. Bronchoconstriction was assessed by methacholine challenge and measured by the log provocative concentration of methacholine resulting in a 20% decrease in airway capacity, measured by FEV1 (PC20). GWAS data was available for each population. RESULTS. We obtained estimates of heritability for bronchodilator response of 28.5% (n = 1497, p = 0.043) and airway hyperresponsiveness of 51.5% (n = 1054, p = 0.064). This corresponds with previously published heritability estimates based on intraclass correlation coefficients of 40% and 67%, respectively. The asthma liability due to additive SNP effects was found to be 61.5% (n = 2633, p < 0.001), corresponding with asthma twin studies that have put the total heritability between 70% and 90%. We note that 13 recent SNPs robustly identified through asthma GWAS meta-analyses (the GABRIEL and EVE consortia) account for no more than 1% of asthma liability. CONCLUSIONS. These results indicate that a very substantial fraction of the heritability of these conditions are due to simple SNP main-effects, and that a polygenic approach can be a key tool in better identifying the combined effect of multiple variants on complex phenotypic outcomes.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

626T

Web-based phenotyping yields replication of genetic associations with sensitivity to warfarin. *J.L. Mountain[1], A.K. Kiefer[1], M. Mullins[1], T.K. Acquaye[1], C.B. Marsh[2], J.A. Johnson[3], H.L. McLeod[4], J.Y. Tung[1], N. Eriksson[1], K.E. Barnholt[1].* 1) 23andMe, Inc, Mountain View, CA; 2) College of Medicine, The Ohio State University, Columbus, OH; 3) Department of Pharmacotherapy and Translational Research, University of Florida, Gainesville, FL; 4) Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC.

Progress in pharmacogenomics research has been hampered by the cost and time required to assemble, assess, and genotype sufficiently large cohorts of patients. 23andMe's innovative research platform leverages involvement of a large and expanding cohort of over 150,000 genotyped individuals who have the option to respond to web-based questionnaires on an ongoing basis. The primary goal of this study is to develop, assess and implement web-based surveys for the rapid collection of drug response and toxicity data from thousands of individuals taking three classes of medication: warfarin, proton pump inhibitors (PPIs), and non-steroidal anti-inflammatory drugs (NSAIDs). In the first phase of this two-part study we assessed how well web-based questionnaires elucidate drug response data by comparing online self-reported data to responses obtained via semi-structured telephone interviews. Web-based surveys performed well in terms of obtaining medication response and side effects information. Test-retest reliability for most questions was high (>70%). Reporting for prescription medications taken currently proved more reliable than for medications taken in the past and for medications taken as needed. The second phase of this project is focused on validation of known associations between warfarin sensitivity and variants in the CYP2C9 and VKORC1 genes, between reactions to NSAIDs and variants in the CYP2C9 gene, and between reactions to PPIs and variants in the CYP2C19 gene. In our first analysis we examined warfarin sensitivity and CYP2C9 and VKORC1 genotypes in 1028 individuals who had taken or were taking warfarin. A test of association between reported warfarin sensitivity and sensitivity predicted on the basis of combinations of CYP2C9 (rs1057910, rs1799853) and VKORC1 (rs9923231) variants indicated borderline significance. A more detailed analysis revealed highly significant associations between self-reported dosage and rs1799853 and rs9923231, along with a near significant association with rs1057910. These initial results provide independent evidence that web-based surveys can yield drug response data of sufficiently high quality to enable highly scalable pharmacogenomics studies. The final phase of this project will leverage 23andMe's customized genotyping chip to test for novel genes associated with drug metabolism, efficacy, and toxicity. This study is funded in part by NIH grant 1R43HG005807-01.

627T

Follow-up study of genetic risk factors for carbamazepine-induced cutaneous adverse drug reactions in Japanese population. *T. Ozeki[1], T. Mushiroda[1], A. Takahashi[2], M. Kubo[3], Y. Shirakata[4], Z. Ikezawa[5], M. Iijima[6], T. Shiohara[7], K. Hashimoto[4], Y. Nakamura[1].* 1) Research Group for Pharmacogenomics, RIKEN Center for Genomic Medicine, Yokohama, Japan; 2) Research Group for Medical Informatics, RIKEN Center for Genomic Medicine, Yokohama, Japan; 3) Research Group for Genotyping, RIKEN Center for Genomic Medicine, Yokohama, Japan; 4) Department of Dermatology, Ehime University Graduate School of Medicine, Ehime, Japan; 5) Department of Dermatology, Yokohama City University Graduate School of Medicine, Yokohama, Japan; 6) Department of Dermatology, Showa University School of Medicine, Tokyo, Japan; 7) Department of Dermatology, Kyorin University School of Medicine, Tokyo, Japan.

An anticonvulsant, carbamazepine (CBZ), is known to show incidences of cutaneous adverse drug reactions (cADRs) including Stevens-Johnson syndrome (SJS), toxic epidermal necrolysis (TEN) and drug-induced hypersensitivity syndrome (DIHS). Previously, to identify a gene(s) susceptible to CBZ-induced cADRs, we conducted a genome-wide association study (GWAS) and following HLA typing in 61 cases and 376 patients who showed no cADRs by administration of CBZ (CBZ-tolerant controls), and found that HLA-A*31:01 was present in 60.7% (37/61) of the patients with CBZ-induced cADRs, but in only 12.5% (47/376) of the CBZ-tolerant controls (odds ratio (OR) = 10.8, 95% confidence interval (CI) = 5.9–19.6, $P = 3.64 \times 10^{-15}$). To validate the significant associations of HLA-A*31:01, we performed a replication study using an independent Japanese case-control subjects which consisted of 42 CBZ-induced cADR cases and 46 CBZ-tolerant controls. The association of HLA-A*31:01 was replicated in the second subjects ($P = 2.73 \times 10^{-6}$, OR = 10.0, 95% CI of 3.6–27.9; combined-analysis $P = 2.35 \times 10^{-23}$, OR = 11.2, 95% CI of 6.9–18.3), implying that this allele has the 62.1% sensitivity and 87.2% specificity when we apply HLA-A*31:01 as a risk predictor for CBZ-induced cADRs. We further analyzed the association of HLA-A*31:01 according to the type of cADR using the combined subjects. HLA-A*31:01 showed significant associations with DIHS ($P = 6.20 \times 10^{-15}$, OR = 14.6, 95%CI of 7.3–29.2) and SJS/TEN ($P = 2.30 \times 10^{-4}$, OR = 34.1, 95%CI of 3.9–297.2) as well as other cADRs ($P = 3.54 \times 10^{-11}$, OR = 8.2, 95%CI of 4.5–15.2), respectively. For a conditional logistic regression analysis of the GWAS data, we selected rs1633021 as the covariant with the firmest association observed in the analyzed GWAS peak of association. After the conditional analysis, there was no SNP that reached GWAS-level significant association.

628T

Re-examination of the CATIE pharmacogenetics whole-genome dataset using a run of significiant findings approach, focus on tardive dyskinesia. *C. Zai, V. Gonçalves, A. Tiwari, G. Oh, V. de Luca, D. Mueller, J. Knight, J. Kennedy.* Dept Neurogenetics, CAMH, Toronto, ON, Canada.

Introduction: Tardive dyskinesia (TD) is a serious, potentially irreversible motor side effect that arises in patients treated long-term with typical antipsychotic medication. Over 15 years of genetic research into this debilitating condition has revealed a number of candidate genes for TD, including DRD2. A number of genome-wide association studies have been conducted recently, but none gave genome-wide significant findings. Even the most significant markers did not replicate between the two GWAS on the same data from the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE). Aim: We aim to search for markers that are independent of one another within and around genes that may contribute to the risk of tardive dyskinesia. Methods: We propose to screen for runs of significant findings ($p<0.05$) in the CATIE sample using PLINK and R. After quality control and multidimensionality scaling (MDS) analysis for population stratification, we pruned the number of variants based on r-square value of 0.80. We then performed linear regression analysis on baseline total Abnormal Involuntary Movement Scale AIMS) on 385 European schizophrenia patients diagnosed with schizophrenia, including sex, age, and years of antipsychotic medication as covariates. Results: Using sliding window sizes of 10, and the criterion of a minimum of seven SNPs with $p<0.05$ within each window, we identified 76 windows of significant markers that correspond to 26 overlapping chromosomal regions for total AIMS scores. We performed 1000 permutations of the results and found only four of the permutations gave more runs that our actual AIMS data (thus $p=0.005$). A number of genes from this analysis have been implicated in ataxia (RGS6) and Bell's palsy (CSGALNACT1), suggesting that variants in these genes may contribute to part of the pathophysiology of schizophrenia and tardive dyskinesia. Discussion: This analytic approach may complement current GWAS analysis strategies in identifying novel candidate genes for complex phenotypes. We will conduct similar analysis on the individual AIMS item scores to explore whether these associations may be more specific for orofacial or limb dyskinesia.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

629T

A Bayesian Approach to Genome-Wide Genetic Association Studies with Survival Time as Outcome. *L. Chien[1], I. Chang[2], C. Hsiung[3].* 1) Institute of Statistics, National Tsing-Hua University, Hsinchu, Taiwan; 2) National Institute of Cancer Research, National Health Research Institutes, Taiwan; 3) Division of Biostatistics and Bioinformatics, Institute of Population Health Sciences, National Health Research Institutes, Taiwan.

Time to progression or overall survival is a widely used outcome variable to evaluate the treatment response in many complex disease research, including cancer. In addition to candidate gene approach, there are genome-wide association studies to look for regions in the genome that harbor genes influencing this outcome variable. Such genetic associations can be used to provide insight into relevant biological pathways or to help guide clinical care in a more personalized manner. It is well-known that clinical traits like survival time are often influenced by many variables, including epigenetic modifications, somatic mutations, genetic variants, etc.; in particular, pharmacogenomics of anticancer agents are complex because genetic variations in both host and tumor genomes are involved in determining anticancer effects of medications. It is therefore of interest to estimate the proportion of variability in the outcome variable that can be explained by the (common) SNPs in a GWAS study, in addition to looking for their associations with survival time. With these in mind, we propose a Bayesian variable selection method for survival models so that survival analysis with covariates the genotype data from a GWAS will shed light on the proportion of variance explained (PVE) by the common SNPs, the sparsity of the model and the effect size of the selected SNPs. We follow the approaches in Guan and Stephens (2011) to consider prior that uses the concept of PVE and to design MCMC algorithms for posterior inferences. Simulation studies will be provided to indicate the numerical performance of this method and compare it with single SNP analysis method. In particular, we will examine if PVE can be satisfactorily studied. Illustrations by a real dataset will be included.

630T

***IL28B* Genotypes Associated with IFN-Induced Virologic Response in Clinical Studies Evaluating Eltrombopag Treatment in HCV-infected Patients with Low Platelet Counts.** *M. Chiano[1], L. Condreay[2], L. Briley[2], D. Pulford[1], D. Fraser[2], K. Johansson[2], D. Briley[2], E. Harris[2], S. Bhatti[1], D. Theodore[3], F. Campbell[4], S. Vasey[5], M. Arning[5], C. Spraggs[1].* 1) GlaxoSmithKline, Quantitative Sciences, Stevenage, UK; 2) GlaxoSmithKline, Quantitative Sciences, Research Triangle Park, NC, USA; 3) GlaxoSmithKline, Oncology, Research Triangle Park, NC, USA; 4) GlaxoSmithKline, Stockley Park, UK; 5) GlaxoSmithKline, Upper Providence, PA, USA.

Eltrombopag (PROMACTA™) is a thrombopoietin receptor agonist for treating thrombocytopenia in patients with chronic immune (idiopathic) thrombocytopenic purpura. We conducted two clinical studies investigating the utility of eltrombopag to increase platelet counts in thrombocytopenic patients with chronic hepatitis C virus (HCV) infections to enable initiation and maintenance of antiviral therapy with pegylated interferon alfa (2a or 2b) and ribavirin (IFN/RBV). Clinical benefit in these studies was measured by the proportion of subjects who achieved Sustained Virologic Responses (SVR). *IL28B* genotypes predict spontaneous HCV clearance and virologic response to IFN/RBV treatment. We evaluated the impact of *IL28B* variants, rs8099917 and rs12979860, in a large cohort of thrombocytopenic patients with HCV (N= 906), White (n=689) and Asian (n=217), treated with open-label eltrombopag and randomized to eltrombopag plus IFN/RBV or placebo plus IFN/RBV. Association with SVR, Rapid (RVR), and Early (EVR) Virologic Response was tested using logistic regression assuming a recessive genetic model. Overall, the proportion of subjects achieving virologic response was significantly higher in subjects receiving eltrombopag and IFN/RBV than in subjects randomized to Placebo and IFN/RBV. Markers rs12979860 and rs8099917 were each significantly associated with virologic response ($p<0.001$). The effect of rs12979860 conditional on rs8099917 remained strongly associated with virologic responses ($p<0.001$), but modelling the effect of rs8099917 conditional on rs12979860 was no longer statistically significant ($p>0.05$). The joint effect of both markers did not provide a significant improvement to the model fit. Therefore, rs12979860 provides a more parsimonious genetic predictor for virologic responses in these data. The rs12979860 `CC' genotype carriers had significantly higher virologic responses compared to subjects that carried one or no copy of this allele (response [OR, p-value]: SVR [3.6, 5.1×10^{-10}], RVR[4.0, 1.62×10^{-8}] and EVR [3.5, 1.08×10^{-12}]). The effect of IL28B genotype was stronger in white than in Asian subjects and in patients with viral genotypes known to be less responsive to IFN/RBV treatment. The distribution of IL28B genotypes was balanced across treatment groups and across ethnic subgroups, so the efficacy of eltrombopag treatment in these studies cannot be attributed to an imbalance in the distribution of the *IL28B* favourable response genotype.

631T

Mapping the Incidentalome: Quantifying Incidental Findings Generated Through a Clinical Pharmacogenomics Project. *K.B. Brothers[2], M.J. Westbrook[1], M.F. Wright[1], S.L. Van Driest[3], T.L. McGregor[3,4], J.C. Denny[5], R.L. Zuvich[4], E.W. Clayton[1,3].* 1) Center for Biomedical Ethics and Society, Vanderbilt University, Nashville, TN; 2) Pediatrics, University of Louisville, Louisville, KY; 3) Pediatrics, Vanderbilt University, Nashville, TN; 4) Center for Human Genetics Research, Vanderbilt University, Nashville, TN; 5) Department of Biomedical Informatics, Vanderbilt University, Nashville, TN.

The use of multiplex genetic technologies in routine clinical care is becoming increasingly feasible. Empirical data on the numbers and types of incidental findings generated are needed so that policies and practices for their clinical use can be developed. PREDICT is an institutional program to implement prospective clinical genotyping of 34 genes to guide drug selection and dosing. Using a combination of computer and hand-based curation, we reviewed 5566 journal articles to quantify and characterize the incidental, non-pharmacogenomic genotype-phenotype associations that could be generated by this panel. We found 383 putative associations supported by at least one study demonstrating a positive association, defined as an odds ratio with a 95% confidence interval that does not cross 0. Of these, 292 are supported by at least one study showing a significantly elevated risk (odds ratio ≥2.0 or ≥0.5). Patients with certain variants in GSTM1 could carry a significant risk for over 40 phenotypes. Other patients may carry no high-risk variants for these 34 genes.

CategoryDefinitionGenotype-Phenotype Associations Replicated≥2 positive findings, 0 negative findings14No Replication Attempts1 positive finding, 0 negative findings97Mixed Findings≥1 positive findings, ≥1 negative findings272**Total383**

Some patients could receive an overwhelming number of results. Careful practices for managing incidental findings are needed.

632T

The Coriell Personalized Medicine Collaborative Pharmacogenomics Appraisal and Evidence Scoring Sytem. *N. Gharani, M.A. Keller, C.A. Stack, T.J. Schmidlen, R.S. Kasper, J.P. Jarvis, N.P. Gerry, E.S. Gordon, M.F. Christman.* Coriell Institute for Medical Research, Camden, NJ.

The incorporation of pharmacogenomics (PGx) into drug therapy is anticipated to improve drug efficacy and reduce adverse drug reactions. However, there has been a lag in the adoption of PGx in clinical practice due, in part, to a lack of rigorous systems for translating published PGx data into standardized diagnostic tests with clear therapeutic recommendations. The Coriell Personalized Medicine Collaborative (CPMC) is a prospective observational study designed to evaluate the utility of personalized genomic information in health management. In this research context, PGx results are included based on the CPMC pharmacogenomics appraisal and evidence scoring (PhAES) system. The PhAES system is drug-centric, utilizes a multi-tiered evidence scoring procedure to define key genetic variants influencing variation in drug response, involves oversight by an external advisory committee, and presents guidelines for the translation of genetic results into anticipated drug response outcomes. The evidence score, assigned to each variant in a given gene, consists of 14 categories ranging from the strongest evidence (presence of *in vivo* clinical data for the drug under review) to the weakest (e.g. identification of the variant through a gene sequencing study with no additional *in vivo* or *in vitro* functional data). Scores less than seven indicate a defined effect on drug response or clinical outcome while those greater than eight represent a lack of sufficient evidence for an effect. Once identified, a Punnett square is used to pair the key genetic variants in all possible diploid combinations; published guidelines or empirical data are then used to present drug-specific genotype-phenotype interpretations for each of the anticipated genetic results. Key to the CPMC PhAES process is the use of an external expert advisory panel known as the Pharmacogenomics Advisory Group (PAG). The PAG evaluates PhAES results for each drug-gene pair proposed for inclusion in the CPMC study and provides an independent and expert vetting process that further validates the inclusion of each variant for use in the CPMC clinical utility study. This systematic approach for evaluating and translating published PGx data is likely to have broad utility in facilitating the adoption of PGx into clinical practice.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

633T

PharmGKB: From pharmacogenomics knowledge to clinical interpretation and implementation. *L. Gong[1], M. Whirl-Carrillo[1], J.M. Hebert[1], E.M. McDonagh[1], K. Sangkuhl[1], C.F. Thorn[1], D. Klein[1], R.B. Altman[1,2], T.E. Klein[1]*. 1) Department of Genetics, Stanford University 1501 S. California Avenue Palo Alto, California 94304; 2) Department of Bioengineering, Stanford University, Stanford, CA.

The PharmGKB is a publically available online database that provides knowledge about the impact of genetic variation on drug response. We have comprehensive catalog of genes and genetic variations that are most important for drug response (many directly linked to drug efficacy and toxicity), and have presented condensed knowledge in the forms of drug pathway diagrams and Very Important Pharmacogene (VIP gene) summaries. As the field advances, our focus has shifted towards facilitating clinical interpretation and implementation of pharmacogenomics discoveries. The recent redesign of our home page and development of specialized translational tools has enabled us to provide clinical interpretations of pharmacogenetic knowledge, including genotype-guided dosing guidelines, potentially clinically actionable gene-drug associations, FDA drug labels and genetic testing information. In addition to providing knowledge and tools that are critical for bridging the gap between bench and clinic on our website, the PharmGKB is host to several large international consortia groups who are carrying out pharmacogenetics research to answer important questions. PharmGKB also contributes to clinical implementation projects to through collaborations and participates in developing methods and tools for clinical interpretation of personal genomes. We will demonstrate available resources on PharmGKB to support clinical implementation of PGx and also discuss key success factors and challenges of our approaches in bringing the pharmacogenomics knowledge to clinic. The PharmGKB is freely available through the http://www.pharmgkb.org/ website.

634T

Development of a scoring tool to prioritize clinical pharmacogenomic testing. *S.F. Manzi.* Gene Partnership, Children's Hospital Boston, Boston, MA.

A scoring tool was developed to assist in the process of prioritizing drug/gene pairs for clinical implementation of pharmacogenomic testing. The provision of clinical interpretation of genotype data in the electronic medical record (EMR) along with decision support at the point of ordering and dispensing is complex and requires resources to build and implement. Once the capability to provide pharmacogenomic testing with EMR decision support becomes publicized, an increasing number of requests for testing for a specific service or patient population are filed. Competing demands for IT support, balancing the needs of a single patient population for whom the information will be critical against the needs of many for whom the information will be important, providing a comprehensive test, and ensuring cost containment are some of the many factors that we incorporated into the tool. Additionally, factors such as commercial availability of the test, published knowledge base tools and guidelines and the projected utilization were also included. The tool was designed from review of the available literature and revised based upon input from our Clinical Pharmacogenomics Service Oversight committee members. We will demonstrate a use case for the tool via an example of prioritizing between TPMT genotyping for mercaptopurine use in Acute Lymphoblastic Leukemia (ALL) and *cyp2D6* genotyping for opioid drugs in acute pain management. The overall score for *TPMT* (thiopurine S-methyltransferase) genotyping was 27 points compared with a score of 20 points for *cyp2D6* (cytochrome P450, family 2, subfamily D, polypeptide 6) genotyping. The major dissenting areas included delayed toxicity (present for thiopurines, not present for opioids), overlapping toxicities with concomitant agents (present for thiopurines, not present for opioids), frequency of use (limited for thiopurines, extensive for opioids) and the commercial availability of a comprehensive test at the time of evaluation (present for *TPMT*, not present for *cyp2D6*). To our knowledge, no other decision tool designed to assist with prioritization of movement of drug/gene pair data from research to the clinical realm has been published. We will demonstrate the tool and scores relevant to our program.

635T

Pharmacogenomic research undertaken by PharmGKB consortia groups. *E.M. McDonagh[1], M. Whirl Carrillo[1], L. Gong[1], J.M. Hebert[1], K. Sangkuhl[1], C.F. Thorn[1], D. Klein[1], R.B Altman[2], T.E. Klein[1]*. 1) Genetics Department, PharmGKB, Stanford University, Palo Alto, CA; 2) Bioengineering Department, Stanford University, Stanford, CA.

The Pharmacogenomics Knowledge Base (PharmGKB) is a comprehensive online database of associations between genetic variants and drug responses for clinicians and researchers (www.pharmgkb.org). The PharmGKB is also host to numerous international research consortia groups that are investigating important questions in the field of pharmacogenetics. The data-led consortia hosted by the PharmGKB aim to combine clinical and genotyping data from cohorts at multiple centers around the world to gain large data sets for analysis of associations between genetic variants and drug responses. The International Warfarin Pharmacogenetics Consortium - Genome Wide Association Studies (IWPC-GWAS) are combining GWAS data from different cohorts to identify novel genetic variants that are important for warfarin dosing, following work published by the IWPC combining genetic and clinical information into a warfarin-dosing algorithm. The International Tamoxifen Pharmacogenomics Consortium (ITPC) are investigating the association between CYP2D6 gene variants and efficacy and toxicity of tamoxifen, a breast cancer therapy. The International SSRI (Selective Serotonin Reuptake Inhibitors) Pharmacogenomics Consortium (ISPC) aims to identify genetic variants that affect responsiveness to SSRI antidepressant therapy. The goal of the recently formed International Clopidogrel Pharmacogenomics Consortium (ICPC) is to identify novel variants involved in clopidogrel response as well as further dissecting the role of CYP2C19 gene variants. Information and publications by the data-led consortia groups can be found at www.pharmgkb.org. PharmGKB has collaborated with the Pharmacogenomics Research Network (PGRN) to create the Clinical Pharmacogenetics Implementation Consortium (CPIC), a knowledge-based consortium of pharmacogenomic experts. CPIC publishes genotype-based drug dosing guidelines that provide guidance to clinicians who have access to preemptive genotyping results. Guideline summaries, updates and downloadable pdf files of the CPIC publications are available on PharmGKB. The consortia hosted by the PharmGKB are investigating important questions to further pharmacogenomics knowledge and contribute to the implementation of pharmacogenomics in the clinic.

636T

Effects of *CYP2C9* and *VKORC1* polymorphisms on clinical outcomes and warfarin response - focusing on personalized medicine. *P.C.J.L. Santos[1], R.A.G. Soares[1], C.M.C. Strunz[2], J.E. Krieger[1], A.C. Pereira[1]*. 1) Laboratory of Genetics and Molecular Cardiology, Heart Institute, University of Sao Paulo Medical School, Sao Paulo, SP., Brazil; 2) Clinical Laboratory, Heart Institute (InCor), University of Sao Paulo Medical School, Sao Paulo, Brazil.

Background: Warfarin therapy can be challenging because of the wide variation on inter-individual response which have been related to *CYP2C9* and *VKORC1* polymorphisms. Nowadays, a study incorporating clinical efficacy data of genotype-guided dosing algorithm has shown that warfarin pharmacogenetic would improve quality-adjusted life years gained. Aim: In this context and as an initial step towards clinical pharmacogenetic implementation, the main aim of this study was to evaluate the effects of *CYP2C9* and *VKORC1* polymorphisms on clinical outcome and warfarin response in patients from a cardiovascular center. Methods: Nine hundred thirty patients receiving warfarin therapy from Heart Institute (InCor), Sao Paulo, Brazil were selected. They were assessed for general and clinical characteristics, warfarin prescribed dosage and INR (international normalized ratio). Genotypes for the *CYP2C9*2* (430C>T, rs1799853), *CYP2C9*3* (1075A>C, rs1057910) and *VKORC1* 3673 (1639G>A, rs9923231) polymorphisms were performed. Results: Of the 930 patients (mean age 61.8±14.1), 50.5% were female, 67.0% were White, and the main indications of treatment were atrial fibrillation and heart valves. In patients who had two consecutive target INR values, warfarin dose was associated with predicted metabolic phenotypes according to *CYP2C9*2* and *CYP2C9*3* (extensive metabolizer (EM): 30.2±0.6 mg/week and intermediate or poor metabolizers (IM+PM): 26.0±0.7 mg/week; p< 0.001), and with *VKORC1* polymorphism (GG: 33.0±0.7 mg/week, GA: 26.3±0.6 mg/week, AA: 22.1±1.3; p< 0.001). Both *CYP2C9* and *VKORC1* polymorphisms were associated with the current INR value/current dose ratio. All tests were adjusted for age, gender, body mass index, and “race/color”. Patients carrying IM+PM had higher outcomes frequencies (mainly bleeding) compared to EM (30.8% and 21.6%, p= 0.003). However, no association of the *VKORC1* polymorphism with outcomes was observed. Conclusions: Our findings from a cardiovascular patient group reaffirm the influence these markers on clinical outcomes and warfarin response. In this scenario, a next step should be creating a genotype-guided dosing program selecting a contemporary control patient group to check the true cost-effectiveness of this individualized pharmacotherapy.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

637T

Performance of a biomarker risk algorithm for a primary prevention trial of Alzheimer's disease. *M.W. Lutz[1,2], D.K. Burns[1], D.G. Crenshaw[1,2], A.M. Saunders[1,2], A.D. Roses[1,2].* 1) Zinfandel Pharmaceuticals, P.O. Box 110395, Durham, NC, 27709; 2) Department of Neurology, Duke University School of Medicine, Durham, NC, 27705.

Genetic biomarkers and predictive models are increasingly being proposed for clinical trial enrichment. These models vary greatly in terms of the number of input variables, structure (algorithmic or statistical basis) and validation criteria. For prediction of the onset of mild cognitive impairment due to Alzheimer's disease, models based on genetic variants (in APOE, TOMM40), CSF-related or blood biomarkers (Aβ1-42, total tau), and imaging-related biomarkers (e.g., hippocampal atrophy measured by MRI) have been proposed. This presentation presents preliminary data on the predictive value of a biomarker risk algorithm based on age and two genetic variants: APOE genotype and TOMM40 '523 genotype. A biomarker risk algorithm has been developed to enrich a double-blind, delay-of-onset clinical trial of the efficacy of pioglitazone, a thiazolidinedione used for the treatment of Type 2 diabetes, to delay the onset of mild cognitive impairment due to Alzheimer's disease. The algorithm will be used to stratify subjects with normal cognition, as determined by a battery of neuropsychological tests, into either a high- or a low-risk group. The algorithm has been designed to allow enrichment to be applied in a reproducible manner across the ages of 68–83 years. In an independent (not used in algorithm development) cohort of 660 subjects (mean age 77 years, range 68–90 years) obtained from the Alzheimer's Neuroimaging Initiative, use of the biomarker risk algorithm was calculated to have a positive predictive value (PPV) of 79% (75–83%, 95% CI). These results compared favorably to the use of a CSF biomarker based on the combination of Aβ1-42 and total tau, which provided a PPV of 65% (53–77%, 95% CI) in a recent meta-analysis. These CSF markers were recently qualified by the European Medicines Agency for selecting patients in the context of a clinical trial. The genetic biomarker risk algorithm will be qualified for potential use as a companion prognostic biomarker at the end of the phase 3 trial when the PPV and negative predictive value (NPV) of the biomarker risk algorithm can be calculated. Use of an enrichment strategy in the context of a clinical trial allows for significant reductions in both the numbers of study subjects required and, in the case of a delay-of-onset study, trial duration.

638T

A Study of Asthma Pharmacogenomics Using RNA-Seq. *B.E. Himes[1,2,3], R. Hu[4], P. Wagner[4], D. Chan[2], B. Klanderman[2], R.A. Panettieri Jr[5], K. Tantisira[1], S.T. Weiss[1,2], Q. Lu[4].* 1) Channing Laboratory, Brigham and Women's Hospital and Harvard Medical School, Boston, MA; 2) Partners Center for Personalized Genetic Medicine, Harvard Medical School, Boston, MA; 3) Children's Hospital Informatics Program, Boston, MA; 4) Program in Molecular and Integrative Physiological Sciences, Department of Environmental Health, Harvard School of Public Health, Boston, MA; 5) Pulmonary, Allergy and Critical Care Division, University of Pennsylvania, Philadelphia, PA.

Asthma is a chronic inflammatory airway disease with well-established heritability that affects over 300 million people around the world. The most common medications used for the treatment of asthma are β2-agonists and glucocorticosteroids, and one of the primary tissues that these drugs target in the treatment of asthma is the airway smooth muscle. RNA-Seq is a high-throughput sequencing method that provides comprehensive expression analysis, including discovery of novel genes, non-coding transcripts, and splice variants. We sought to use RNA-Seq to characterize the human airway smooth muscle (HASM) transcriptome at baseline and under three asthma treatment conditions. The Illumina TruSeq assay was used to prepare RNA-seq libraries for HASM cells from four white male donors under four conditions: 1) no treatment; 2) treatement with a β2-agonist (i.e. albuterol, 1μM for 18h); 3) treatment with a glucocorticoid (i.e. dexamethasone, 1μM for 18h); 4) simultaneous treatment with a β2-agonist and glucocorticoid. Preliminary data based on reads sequenced by an Illumina Genome Analyzer 2, aligned to the human reference hg19 using TopHat, and further analyzed using Cufflinks and CummeRbund suggest that several novel genes, including C12orf12, C16orf3, IGLJ4, KIAA1199, NUDT4, and RTN4 are involved in HASM response to asthma medications. Our results provide a transcriptomic snapshot of the effects of the most common asthma medications in HASM cells and have the potential to improve our understanding of asthma pharmacogenomics.

639T

Genetic Moderators of Treatment Response to Methylphenidate and Guanfacine in Children and Adolescents with ADHD. *E.L. Nurmi, K.S. Mallya, F. Whelan, C.P. Laughlin, J.J. McGough, S.K. Loo, R.M. Bilder, J.T. McCracken.* Dept of Psychiatry, Univ of California at Los Angeles, Los Angeles, CA.

Background: Methylphenidate and guanfacine are effective treatments of hyperactive and inattentive symptoms associated with attention-deficit hyperactivity disorder (ADHD); however, variability in individual treatment response is substantial. We sought to determine whether genetic variation in monoamine drug targets could help explain differential treatment outcomes in a double-blind, randomized, placebo-controlled trial of dexmethylphenidate (d-MPH) and guanfacine in pediatric ADHD (n=209). Methods: Complete common variation in dopaminergic and adrenergic drug targets was tested, including dopamine (DA) receptors DRD1–5, α2-adrenergic receptor 2A (ADRA2A), and catabolic enzymes MAOA and MAOB. Single functional variants in the DA transporter (SLC3A6) and catabolic enzyme catechol-o-methyltransferase (COMT) were also genotyped. Results: In children receiving d-MPH, DRD2 and ADRA2A SNPs predicted treatment response. None of the 7 homozygotes for the minor allele of a synonymous variant (His313His, rs6275) met responder criteria, compared to an 80% response rate in 47 common allele carriers (p=0.0001). The minor (low expression) allele of the DRD2 Taq1A variant (rs1800497) was associated with an allele dosage-dependent improvement in both the d-MPH group and the d-MPH and guanfacine combined group (p=0.001). Furthermore, haplotypes of these 2 DRD2 SNPs showed differential effects on treatment response (p=0.0017). The minor allele of an ADRA2A promoter SNP (rs521674) was associated with poor d-MPH response (p=0.0004). Homozygotes for the minor alleles of functional variants in DRD1 (rs686) and DRD2 (rs2075654) predicted guanfacine response (p<0.0001). Conclusions: Common genetic variation in dopaminergic and adrenergic receptors influenced treatment response to standard ADHD drug therapies in our dataset. These results survive correction for multiple testing but warrant replication in larger samples and prospective treatment studies. Understanding pharmacogenetic moderators of ADHD treatment response such as these could improve clinical treatment matching and guide the design of future therapeutics. Supported by P50 MH077248.

640T

Identifying genes that influence acute ethanol responsive behaviors in *Caenorhabditis elegans*. *J.T. Alaimo, A.G. Davies, J.C. Bettinger.* Pharmacology and Toxicology, Virginia Commonwealth University, Richmond, VA.

Alcohol abuse and alcohol addiction are serious and common illnesses that are strongly influenced by genetics. One factor influencing liability for alcohol dependence in humans is functional variation in genes encoding ethanol metabolism machinery, alcohol dehydrogenase (ADH) and aldehyde dehydrogenase (ALDH) (Pautassi et. al, 2010). We directly tested the effects of altering the function of these enzymes on ethanol responsive behaviors in *C. elegans*. We tested two ADH enzymes encoded by the genes *sodh-1* and *H24K24.3*, and found that loss of *sodh-1* but not *H24K24.3* caused animals to be acutely hypersensitive to ethanol's depression of locomotion. As expected, tissue ethanol concentration was increased when *sodh-1*, but not *H24K24.3* was lost. Importantly, we found that both strains develop robust acute functional tolerance (AFT) to ethanol, indicating that these enzymes are dispensable for this process. We found similar hypersensitivity to ethanol when we inactivated a subset of the worm ALDH enzymes, *alh-6* and *alh-13*. Interestingly, internal tissue ethanol concentrations in these two strains were higher than in wild type, suggesting that the lack of ALDH function causes a buildup of acetaldehyde, which can be converted by ADH into ethanol. Collectively, these data demonstrate that altered ethanol metabolism in worms results in a mild but detectable effect on ethanol response behaviors. A second risk factor for the development of alcohol dependence in humans is an individual's naive level of response to ethanol (Shuckit, 1994). The progression from initial ethanol responses to dependence is not well understood, but many studies have shown that acute ethanol administration induces changes in gene expression, which may contribute to later dependence behaviors. Our major goal is to identify genes and molecular pathways that are important in mediating ethanol responsive behaviors. We will identify ethanol responsive genes by performing a series of microarray experiments at different times and doses of ethanol exposure. Candidate genes from this analysis will be tested for their functional relevance in ethanol responses. We will test several ethanol-responsive behaviors including those measuring initial sensitivity and the development of AFT, two components of the initial level of response phenotype that is predictive of abuse liability in humans.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

641T

Characterization of Northeast Asian ADME Gene Variation via Exome Sequencing. *D.H. Hovelson[1], Z. Xue[2], S. Rashkin[1], X. Zhan[1], M.G. Ehm[2], A. Yeo[2], I.J. Jang[3], I. Ieiri[4], J.E. Lee[5], L.R. Cardon[2], S.L. Chissoe[2], G. Abecasis[1], M.R. Nelson[2]*. 1) Center for Statistical Genetics, University of Michigan, Ann Arbor, MI, USA; 2) Quantitative Sciences, GlaxoSmithKline, RTP, NC, USA and Stevanage, UK; 3) Seoul National University, Seoul, Korea; 4) Kyushu University, Fukuoka, Japan; 5) DNA Link, Inc., Seoul, Korea.

Drug response is influenced by pharmacokinetics, the availability of drug in the body. Investigating the factors that influence pharmacokinetics and selecting proper doses is a crucial part of drug development. Traditionally, early pharmacokinetic studies have been conducted in Western countries, with subsequent studies conducted in Northeast Asian countries to address the potential for population-specific effects and satisfy local regulatory requirements. Variation in genes that regulate the absorption, distribution, metabolism and excretion (ADME) of therapeutic agents can have a substantial impact on pharmacokinetic activity and may lead to variability in drug concentrations at the relevant tissue. Understanding ADME gene variation and its functional impact is important to developing safe and effective medicines administered at proper doses. There have been few systematic attempts to identify and characterize the extent of coding variation in all key ADME genes among major world-wide populations. To address this, we have integrated the exome sequence available from 19 populations from the 1000 Genomes Project and supplemented this with the exomes of 125 Korean and 125 Japanese subjects. These combined data would identify >97% of Northeast Asian coding variants with minor allele frequencies >0.5%. Average depth across the core (N=32) and extended (N=263) ADME genes for our Asian subjects is approximately 60X; for 1000 Genomes Project samples, this average depth is ~130X. We characterize a broad spectrum of variation across these genes, including single nucleotide variants, short insertion-deletion sequences and gene copy number differences. Initial analyses show an average of 1,947 non-singleton/non-doubleton SNVs in ADME gene coding regions per individual. Of these SNVs, on average, 990 are non-synonymous, 10 are nonsense, and 130 appear to be novel. We use predictive algorithms to assess the function of known and novel variants, and compare the patterns of variation across the world. The results of this study will assess the completeness of current ADME marker panels, guide the design of future pharmacokinetic studies, and hopefully improve the efficiency of getting new medicines to patients.

642T

Drug Metabolizing Enzyme Genes and Nicotine and Cotinine Metabolism. *A.W. Bergen[1], H.S. Javitz[1], M. Michel[1], R. Krasnow[1], D. Nishita[1], C.N. Lessov-Schlaggar[2], H. Hops[3], N.J. Markward[4], S. Hall[5], T. Baker[6], D.V. Conti[7], R.F. Tyndale[8], N.L. Benowitz[9], G.E. Swan[1]*. 1) Center for Health Sciences, SRI International, Menlo Park, CA; 2) Department of Psychiatry, Washington University, St. Louis, MO; 3) Oregon Research Institute, Eugene, OR; 4) Personalized Medicine Research and Development, Medco Health Solutions, Bethesda, MD; 5) Department of Psychiatry, University of California, San Francisco, CA; 6) Department of Medicine, Center for Tobacco Research and Intervention, University of Wisconsin School of Medicine and Public Health, Madison, WI; 7) Department of Preventative Medicine, University of Southern California, Los Angeles, CA; 8) Centre for Addiction and Mental Health, and Departments of Psychiatry, Pharmacology and Toxicology, University of Toronto, Toronto, ON; 9) Departments of Bioengineering and Therapeutic Sciences and of Medicine, University of California, San Francisco, CA.

Tobacco smoking contributes to 30% of mortality in the U.S. More effective smoking cessation treatments, including treatments informed by pharmacogenetics, are needed. The Nicotine Metabolic Ratio (NMR), the ratio of two stable metabolites of nicotine, influences the number of cigarettes smoked per day, responsiveness to treatment, and tobacco carcinogen exposure. The predominant metabolic enzyme involved in nicotine and cotinine metabolism is CYP2A6. Significant additive genetic influence on nicotine and cotinine metabolism remains after accounting for *CYP2A6* common variants. Common *CYP2A6* variation accounts for ~70% of NMR variance. Pharmacogenetic analysis of the NMR may identify genes that influence the activity of CYP2A6, which is responsible for 100% of the metabolism of cotinine to 3'hydroxycotinine. We performed a pharmacogenetic study of the NMR in 220 kindreds from two West Coast metropolitan areas who received a standardized dose of deuterium-labeled nicotine and/or cotinine and from whom metabolite levels were measured using liquid chromatography-mass spectrometry. We utilized the Affymetrix DMET™ Plus Assay, which interrogates 1936 markers at 236 drug metabolizing enzyme and transporter genes. Adjusted (age, gender, BMI, smoking, hormonal status and ten principal components of population genetic variation) mean (SD) NMR values were 0.26 (0.07), plasma, in the twin dataset (N=333), and 0.29 (0.09), saliva, in the pedigree dataset (N=215), respectively. We performed hierarchical linear modeling of common DMET SNPs and adjusted NMR, followed by adjustment for correlated tests (P_{ACT}) within genes with >1 common SNP with ≥1 SNP with nominal P<0.05. We identified SNPs at 13 genes with P_{ACT}<0.05 in ≥2 transmission models in the larger twin dataset. These 13 genes include the established metabolic genes *CYP2A6* and *UGT2B7*. SNPs in 3 of 13 genes exhibit PACT<0.05 in the smaller pedigree dataset, including *CYP2A6*. We are analyzing SNPs from the two genes not previously associated with nicotine metabolism in two randomized controlled trials of smoking cessation therapy (RCTs) with NMR data. SNPs for those two genes associated with the NMR in the two RCTs will then be subsequently genotyped in an additional six RCTs, allowing us to evaluate the influence of SNPs at novel metabolic genes on smoking behaviors in eight RCTs. Funding: Collaboration Agreement between Medco Health Solutions, Affymetrix and SRI International, U01DA020830 and 1R21DA033813.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

643T

Identification of a SNP in the UGT1A1 promoter associated with increased serum bilirubin levels. *M.K. Kringen[1], A.P. Piehler[2,3], R.M. Grimholt[2], M.S. Opdal[1], K.B.F Haug[2], P. Urdal[2]*. 1) Department of Pharmacology, Oslo University Hospital, Ullevål, Oslo, Norway; 2) Department of Medical Biochemistry, Oslo University Hospital, Ullevål, Oslo, Norway; 3) Fürst Medical Laboratory, Oslo, Norway.

Background: Genetic variants are known to contribute to decreased glucuronidation activity in the liver due to decreased promoter activity of the gene encoding UDP-glucuronyltransferase 1A1 (UGT1A1). Decreased glucuronidation activity leads to increased bilirubin levels in healthy people (Gilbert's syndrome) and may cause variable responses to drugs that are metabolized by UGT1A1. In this study, we have investigated possible associations between serum bilirubin levels and seven tag-SNPs in a ~30 Kbp region upstream of the promoter of exon 1 of UGT1A1, in relation to the common TA7-repeat polymorphism (rs8175347). Methods: Buffy coat from healthy individuals with low (<17.5 μmol/L, n=150) and high (>17.5 μmol/L, n=150) bilirubin levels was obtained from the Nordic Reference Interval Project Bio-bank and Database (NOBIDA), consisting of samples from Denmark, Finland, Iceland, Norway and Sweden. DNA was extracted from 100 μl buffy coat samples and examined for common and new polymorphisms. The UGT1A1-TA7 repeat was genotyped by pyrosequencing using a PyroMark Q24, and tag-SNPs were analyzed by TaqMan® SNP Genotyping Assays on a Viia7 Real-Time PCR System. Results: Most individuals with high bilirubin levels (>17.5 μmol/L) were homozygous for the TA7-repeat (74%), while only 3% were homozygous for the TA6-repeat. However, in heterozygous TA6/TA7 individuals, both low and high levels of bilirubin were observed. By regression analysis, adjusting for age, sex and country of origin, an additional SNP (rs7564935 G>T) in the UGT1A1 promoter was found to be associated with higher bilirubin concentrations in this genotype group (P=0.02, n=114). Individuals being homozygous for the major G-allele had a mean bilirubin concentration of 18.5 μmol/L compared to 11.2 μmol/L for TT-carriers. This SNP explained about 2–3% of the bilirubin variability in TA6/TA7 individuals. Conclusions: High bilirubin levels are mainly determined by the UGT1A1-TA7 repeat. Additionally, we have identified a SNP in the UGT1A1 promoter associated with increased bilirubin concentration in TA6/TA7 heterozygous individuals. For drugs that are metabolized by UGT1A1, this SNP may predict altered drug responses in TA6/TA7-carriers.

644T

Well-characterised cohorts and exome data: investigating antipsychotic pharmacogenomics in the South African context. *B. Drogemoller[1], D. Niehaus[2], G. Wright[1,3], B. Chiliza[2], L. Asmal[2], L. Koen[2], R. Emsley[2], L. Warnich[1]*. 1) Department of Genetics, Stellenbosch University, Stellenbosch, Western Cape, South Africa; 2) Department of Psychiatry, Stellenbosch University, Stellenbosch, Western Cape, South Africa; 3) South African Bioinformatics Institute, University of the Western Cape, Bellville, Western Cape, South Africa.

Introduction: Schizophrenia places an immense socio-economic burden on society and current treatments have substantial limitations. This burden is further heightened by poverty in low- and middle-income countries, such as South Africa. As African populations have been under-represented in genome-wide research to date, this study aims to identify unique patterns of genomic variation that may affect antipsychotic response in South African patients.

Materials and Methods: A cohort of 104 South African first episode schizophrenia patients was collected and treated for up to 3 years. This cohort is well suited to pharmacogenomic research due to the fact that patients have been carefully selected and assessed, diagnosed according to specific criteria and treated uniformly. From this cohort we have identified a subset of age, sex and ethnicity matched antipsychotic non-responders (4 unrelated individuals and 1 sib-pair) and responders (n=5) for exome sequencing. The generated exome data was subsequently analysed employing a bioinformatics pipeline, which made use of the publically-available software programs BWA, SAMtools, GATK, SeattleSeqAnnotation134 and VAAST.

Results: Analysis of the variation in the 11 schizophrenia patients revealed that >5% of the total number of SNPs detected were novel. By utilising a variety of different options in VAAST to model antipsychotic response, we identified a list of potentially damaged genes that may contribute to antipsychotic treatment response. Variation predicted to alter the resulting protein product of these genes, as well as candidate genes described in the literature/databases, was prioritised for genotyping in the entire cohort of patients.

Conclusion: This study has detected a large number of variants, a substantial portion of which remain novel, despite recent efforts to comprehensively characterise the human variome. By utilising a well-characterised cohort and identifying variation affecting patients on extreme ends of the treatment response phenotypic spectrum, we have been able to identify a unique panel of variation that may contribute to antipsychotic response. By genotyping this variation in the entire cohort, we hope to refine our results and, in so doing, improve our understanding of antipsychotic response in the South African context.

645T

Can Genetic Variation of CES1A Predict Methylphenidate Response in ADHD Patients? *A. Cassidy[1], E. Sizer[3], D. Rutschow[1], K. Hogg[3], J. Berg[2], D. Coghill[3]*. 1) Human Genetics, Medical Research Institute, Dundee, United Kingdom; 2) Human Genetics, Medical Education Institute, Dundee, United Kingdom; 3) Division of Neuroscience, Medical Research Institute, Dundee, United Kingdom.

Carboxylesterase 1A (CES1A) is the principle metabolising enzyme for a number of drugs, including methylphenidate. There is early data to suggest that variation in CES1A can affect treatment response in patients with attention deficit hyperactivity disorder (ADHD). The existence of highly homologous isoforms of CES1A on the same chromosome creates a major challenge to identifying variation in the functional isoforms present in an individual. We aimed to develop a discriminative method using long PCR and direct sequencing to analyse the individual variation of these CES1A isoforms. Sequence analysis of a local cohort combined with review of current literature, identified typical variation between the promoter region and exon 1 of the CES1A isoforms. Using this information we have adapted a recently published long PCR method to include analysis of the functional CES1A1b, 1A1c and 1A2 minor haplotypes. Our method successfully enabled specific sequence analysis of 13.5kb of CES1A1a (the most common functional isoform), from exon 1 to intron 6. A 2-step process to further define the isoforms present was also developed. SNP haplotypes identified in the promoter region and exon 1 revealed a highly complex genotype of several functional CES1A isoforms, including a novel CES1A1 isoform, present in 1/48 alleles which may all contribute to methylphenidate response. This new understanding of the genomic structure and variation within the CES1 diplotype has allowed us to design a protocol using massively parallel sequencing in a larger-scale phenotype-genotype study of ADHD patients. This is an important step in studying the effect that the CES1 genotype has on drug metabolism.

646T

Whole exome analysis of African American warfarin pharmacogenomics. *R. Daneshjou[1], K. Aquino-Michaels[2], B. Burkley[3], L.H. Cavallari[4], J.A. Johnson[3], N.A. Johnson[1], T.E. Klein[1], K.J. Karczewski[5], T. Langaee[3], S.R. Patel[4], M. Perera[2], H. Sagreiya[1], H. Tang[1], N. Tatonetti[5], R.B. Altman[1,5]*. 1) Genetics, Stanford University, Stanford, CA; 2) Department of Medicine, University of Chicago, Chicago, IL; 3) Department of Pharmacotherapy and Translational Research, University of Florida, Gainesville, FL; 4) Department of Pharmacy Practice, University of Illinois at Chicago, Chicago, IL; 5) Bioinformatics Training Program, Stanford University, Stanford, CA.

Warfarin, an anti-coagulant, is prescribed to over 30 million patients a year in the United States alone; however, doses can vary between patients as much as tenfold. Underdosing can lead to diminished therapeutic protection against stroke and clotting while overdosing can lead to serious adverse events such as bleeding or hemorrhaging. Currently, dosing is largely trial-and-error — physicians start with a fixed dose and calibrate doses based on coagulation tests until the target international normalized ratio (INR) is achieved. However, warfarin pharmacokinetics and pharmacodynamics are known to be affected by single nucleotide polymorphisms (SNPs) of genes involved in those pathways, such as VKORC1 or CYP2C9. While pharmacogenetic dosing algorithms such as the one produced by the International Warfarin Pharmacogenetics Consortium (IWPC) have been able to explain some of the variability seen in Caucasian patients, its accuracy falls in less studied ethnic populations such as African Americans. Recent studies have shown that genes other than VKORC1 and CYP2C9 may play an important role for warfarin dosing in African Americans. Additional new rare variants may explain this variation in warfarin response, which we explore using whole exome sequencing of 100 African American patients on warfarin with stable doses at low and high levels (<=35 mg/week and >=49 mg/week). In addition to traditional SNP-based analysis, we also use gene-based methods of analysis to bin SNPS for increased statistical power in detecting differences in high and low dose cohorts. Our work will allow for more accurate pharmacogenetic predictions in African American warfarin dosing.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

647T

Survey of rare Variants in pharmacogenes for patients with statin-induced myopathy Identified in a biorepository linked to electronic medical records. *M.T. Oetjens[1], Q. Feng[2], A.R. Ramirez[2], E. Bowton[3], T. Clark[5], D.M Roden[2,3,4], D.C. Crawford[1], R.A. Wilke[2]*. 1) Center for Human Human Genetics; 2) Department of Medicine, Division of Clinical Pharmacology; 3) Office of Personalized Medicine; 4) Department of Pharmacology,; 5) Genome Sciences Resource, Vanderbilt University, Nashville, TN.

HMG-CoA reductase inhibitors, statins, are a widely prescribed class of drugs for prevention and treatment of cardiovascular disease. Their clinical use can be limited by musculoskeletal problems ranging from mild pain (myalgias) to severe muscle damage (myopathy). Plasma levels of the non-specific muscle enzyme creatine kinase (CK) are often used to estimate severity. True myopathy (CK >10-fold upper limit of normal) occurs in ~1 of every 1000 patients exposed. We tested the hypothesis that rare variants contribute to the risk of developing statin induced myopathy in a large clinical practice-based biobank. BioVU, the biobank at Vanderbilt University is a biorepository of DNA samples linked to de-identified electronic medical records (EMRs). As of June 1, 2012, BioVU contained clinical data and DNA from 127,744 unique adults. Case finding algorithms flagged 199 potential myopathy case subjects who experienced CK elevation in the absence of an acute coronary syndrome or major trauma. Manual chart review confirmed 16 cases suitable for exome scanning. Whole exomes were captured using an Agilent SureSelect All Exon v2 38 Mb capture kit and sequenced on the Illumina HiSeq 2000. Short reads were aligned to UCSC human reference genome build hg19 using the BWA aligner with a soft-trim quality score setting (-q) of 5. The resulting BAM files were sorted, indexed and processed with GATK v1.3. Variant calling was performed on all BAM files simultaneously, and the resulting VCF file (containing variants of quality score >20) was filter flagged according to “GATK Best Practices v3.” Seven pharmacodynamic genes (HMGCR, COQ2, CPT2, EYS, RYR1, RYR2, RYR3), and eight pharmacokinetic genes (ABCB1, ABCG2, CYP3A4, CYP3A5, SLCO1B1, SLCO1B3, UGT1A1, UGT1A3) were surveyed for their total burden of rare variants unique to statin-induced myopathy patients compared with the 1000 Genomes reference population. The genetic analysis software VAAST was leveraged to determine if each identified variant was likely to alter protein structure/function. Five genes (RYR1, EYS, ABCG2, UGT1A1 and CYP34) contained one or more rare variants unique to our cases. Only RYR1 contained a rare variant deemed potentially detrimental in VAAST. Due to the rarity of statin-induced myopathy, case patients will need to be studied across institutions to determine if genetic variability in RYR1 impacts risk at the population level.

648T

Genetic Determinants of Fetal Hemoglobin Response to Hydroxyurea. *V.A. Sheehan[1], A. Sabo[1], J.R. Crosby[3], T.A. Howard[1], D.M. Muzny[1], B. Davis[3], A. Li[1], U. Nagaswamy[1], E. Boerwinkle[1,2], R.A. Gibbs[1], R.E. Ware[1,4], J.M. Flannagan[1]*. 1) Pediatric Hematology/Oncology, Baylor College of Medicine, Houston, TX; 2) Human Genome Sequencing Center, Baylor College of Medicine, Houston, Texas; 3) Human Genetics Center, University of Texas Health Science Center at Houston, Houston, Texas; 4) Center for Global Health, Baylor College of Medicine, Houston, TX.

Sickle cell anemia (SCA) is the most common inherited blood disorder, affecting 1 in 400 African Americans, and causing significant morbidity and mortality. Hydroxyurea is the only FDA-approved pharmacologic treatment for SCA. The effectiveness of hydroxyurea is directly related to the level of fetal hemoglobin the drug can induce, with approximately 20% HbF needed for significant disease modification. The amount of HbF produced in response to hydroxyurea is highly variable between individuals, with induced HbF levels ranging from 5 to >30% even for compliant patients on similar dosing regimens. There is strong concordance of HbF response to hydroxyurea in sibling pairs, suggesting response is heritable, and amount of HbF produced at the maximum tolerated dose is a reliable and objective phenotype. To identify genetic modifiers of hydroxyurea induction of HbF, we performed whole exome sequencing on 50 pediatric patients who represented the extreme ends of HbF response to hydroxyurea; 26 high responders (final HbF >30% and >25% change from baseline) and 24 low responders (final HbF <20%, and <15% change from baseline). There was no significant difference in levels of baseline %HbF, gender or age between the two groups. Genomic DNA was obtained from participants in multicenter clinical trials with uniform drug administration and careful monitoring, producing the most reliable phenotypes available. Paired end libraries were constructed, and all exons in each sample captured and sequenced with 94% 20x base coverage. Low responders were compared to high responders. Although no single gene affected HbF response, univariate test results identified ten nonsynonymous polymorphisms with p-values below 1×10^{-3}. Two of the ten genes identified play a role in erythropoeisis; one in maintence of glomerular filtration, of interest as hydroxyurea is renally excreted. This initial pilot investigation examined only extreme phenotypes; these associations will be validated in a larger group using HbF as a continuous variable.

649T

Translating genetic associations to new medicines. *M.G. Ehm[1], S.L. Chissoe[1], M.R. Nelson[1], D.M. Waterworth[1], L.C. McCarthy[1], H. Tipney[1], T. Pastinen[2], M. Saleh[2], P. Gros[2], P. Agarwal[1], M. Lathrop[2], L.R. Cardon[1], J.B. Richards[2], P. Sanseau[1]*. 1) Genetics, GlaxoSmithKline, Res Triangle Park, NC & Stevenage, UK; 2) McGill University, Montreal, Quebec.

Aim: To develop methodologies and processes to efficiently translate human genetic observations into effective new therapies for disease. A key goal for drug development is to minimise drug attrition due to poor therapeutic efficacy in late stage clinical trials. There is a growing body of evidence that human genetic associations can reduce this attrition by identifying proteins causally related to human disease aetiology. Further efficiencies can be gained through genomics-based drug repositioning, wherein new indications are identified for existing drugs when their drug target is associated with a disease different from the drug's original indication. We previously reported that disease genes identified using genome-wide association studies (GWAS) are 2.7-fold enriched for established drug targets, and in some disease classes identify the majority of clinically validated drug targets. We identified 92 possible repositioning opportunities using published GWAS literature. The validation of these drug targets, through allelic expression studies, animal models and functional rare variants can provide rationale to undertake human drug trials. One example is *TNFSF11*, for which we demonstrate that the allele associated with Crohn's disease is associated with increased expression of *TNFSF11*, indicating that an antagonist of this gene's activity, such as Denosumab, may be an effective therapeutic for this condition. In a second example, variants in the melatonin receptor, *MTNR1B*, are associated with increased fasting glucose levels as well as susceptibility to type 2 diabetes (T2D). Cell model validation established that loss of MTNRM.1B function is associated with increased risk of T2D, indicating that a MTNR1B agonist, such as melatonin may be helpful in treatment of T2D. With the rapidly growing knowledge of genetic susceptibility factors for rare and common disease, and an almost complete catalog of common genetic variants, opportunities to validate new targets and reposition for existing targets are increasing. We are addressing the immediate need to identify the methods to bridge the gap between genetic associations and translation into effective medicine development programs through rapid target validation.

650T

Association of CYP3A4 and CYP3A5 genotypes with treatment-related hiperlipidemy. *L.E. Wong-Ley[1], L.E. Figuera[2], R. Ortiz-Orozco[2], M. Perez-Nuño[1], P. Aguiar-Garcia[1], F.A Ortega-Gutierrez[2], A. Nuñez-Zenteno[2]*. 1) Dept Gen, SSN-UAN, Tepic, Mexico; 2) Dept Gen, IMSS, Guadalajara, Mexico.

The usage of pravastatin as a medical treatment for hyperlipidemy and its main complication, coronary aterosclerosa disease, it is effective because it modifies the cholesterol homeostasis and it plays a very important role at the beginning and progression of the ischemic cardiopathy, the main cause of dead in our country for about four decades Some differences have been reported in the response to the treatment with pravastatin related with ethnic origin and there are differences in the frequencies of the polymorphisms that are studied, that is why we considerer a probable association between this polymorphisms and the response to the treatment. The objective the present study is to know the participation of CYP3A4 and CYP3A5 genes in the variation of pravastatin dose in Mexican population. They were included in the study 73 patients that were receiving pravastatin as part of the antilipidic treatment. All patients signed an informed consent and a questionnaire was answered. DNA was obtained from lymphocytes and followed the identification by PCR and enzymatic digestion of *2 and *3 of CYP3A5, *1B of CYP3A4 variants. A t of Student test was applied to determine doses differences between genotype groups. The genotype and allelic frequencies compared from the three referred polymorphisms between sensible and resistant groups without estimating differences that can refer to association among them and the clinical response to the treatment with pravastatin. When making the analysis of the polymorphisms of CYP3A4 and CYP3A5, between the reference group and patients with pravastatin treatment, we did not find significant statistical differences between the genotype and allelic frequencies. The genotype and allelic frequencies of the CYP3A4*1B, CYP3A5*3 and *6 polymorphisms in the patients with hyperlipidemy in the reference population, are similar to the ones of the general population. As well as the distribution of the CYP3A4*1B CYP3A5*3 and *6 polymorphisms in the reference population that agreed with the predictions of the Hardy-Weinberg Law. No main variables play an important role in the dosing of pravastatin in the studied population.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

651T

Pharmacogenetic investigation of diarrhea in pazopanib-treated patients with renal cell carcinoma. *Z. Xue[1], L. Huang[1], C.N. Sternberg[3], R.A. Figlin[4], L. McCann[1], C.F. Spraggs[1], C. Carpenter[2], L.N. Pandite[2], C. Xu[1]*. 1) Quantitative Sciences, GlaxoSmithKline, Stevenage UK, Research Triangle Park, NC, and King of Prussia, PA, USA; 2) Oncology, GlaxoSmithKline, Research Triangle Park, NC, and King of Prussia, PA, USA; 3) Department of Medical Oncology, San Camillo and Forlanini Hospitals, Rome, Italy; 4) Cedars-Sinai Medical Center, Samuel Oschin Comprehensive Cancer Institute, Los Angeles, USA.

Background: Diarrhea is a commonly observed adverse effect during pazopanib treatment. Although most diarrheal events are mild to moderate, this can be burdensome and inconvenient for patients and lead to non-compliance and possibly poor treatment outcomes. This pharmacogenetic study sought to explore the association of germline genetic factors with pazopanib-related diarrhea in patients with advanced renal cell carcinoma (RCC).

Methods: In the dataset of 397 patients from 3 clinical trials (a phase II, a phase III and its extension study) who had pharmacogenetic data, 90 (22.7%) patients experienced NCI CTCv3 AE grades ≥2 diarrhea during pazopanib treatment and were defined as `cases' in this analysis. `Controls' (n=132) were patients who did not experience diarrhea and received pazopanib for at least 42 days (median time to onset of diarrhea). Germline DNA was extracted from peripheral blood. 44 functional markers in 38 candidate genes that are involved in pazopanib metabolism, or pazopanib mode of action, or mechanism of diarrhea were investigated for association with diarrhea using logistic regression. A genome wide association study (GWAS, 1M markers) was also conducted to examine associations with diarrhea.

Results: Out of the 44 candidate functional markers analyzed, IL8 (encoding Interleukin-8) -251T/A (rs4073) and ADCY6 (encoding an enzyme catalyzing the formation of cAMP) A674S (rs3730071) showed nominally significant association with pazopanib-related diarrhea ($p<0.05$). The most significant association from the GWAS (rs6498041) was identified in the gene coding for GSG1-like protein, with a nominal p value of $p=6.4\times10^{-8}$. None of the markers were associated with diarrhea at the multiple-testing corrected significance level in either the candidate gene ($p<0.0006$) or GWAS ($p<2.5\times10^{-8}$) analyses.

Conclusions: Suggestive associations between genetic markers and diarrhea were found in pazopanib treated patients with RCC using candidate gene and GWAS analyses. Results from this exploratory genetic study may provide the foundation for further confirmatory studies to evaluate the mechanism of diarrhea in patients treated with pazopanib or other tyrosine kinase inhibitors.

652T

Multi-ethnic Glutathione S-transferase (GST) Copy Number Profiling. *S. Martis[1], R. Vijzelaar[2], J. Sload[1], H. Mei[1], L. Edelmann[1], R.J. Desnick[1], S.A. Scott[1]*. 1) Department of Genetics and Genomic Sciences, Mount Sinai School of Medicine, New York, New York, 10029; 2) MRC Holland, Willem Schoutenstraat 6, Amsterdam, The Netherlands.

The polymorphic glutathione S-transferases (GSTs) are a major phase II family of conjugation enzymes involved in detoxifying carcinogens and bioactivation of antineoplastic prodrugs. Of these, the most commonly studied isoenzymes are the cytosolic μ1 (GSTM1; 1p13.3), ϕ1 (GSTT1; 22q11.23), and π1 (GSTP1; 11q13.2). For example, common deletion polymorphisms in GSTM1 and GSTT1 result in reduced enzyme activity and have been implicated in several multifactorial diseases including cancer and schizophrenia. To determine the multi-ethnic frequencies of GSTM1, GSTT1 and GSTP1 copy number variants (CNVs), multiplex ligation-dependent probe amplification (MLPA) with two probes per gene was performed on over 500 DNA samples from healthy African-American, Asian, Caucasian, Hispanic, and Ashkenazi Jewish (AJ) individuals. The identified GSTM1 deletion allele frequencies were 0.561, 0.748, 0.741, 0.556 and 0.702, and the GSTT1 deletion allele frequencies were 0.561, 0.646, 0.418, 0.451 and 0.547, respectively. Selected GSTM1 and GSTT1 deletion alleles were confirmed by quantitative TaqMan® copy number assays. To determine the breakpoint regions of both GSTM1 and GSTT1 deletion CNVs, a high-resolution oligonucleotide microarray (average probe spacing of ~119 bp for GSTM1 and ~212 bp for GSTT1) was designed that interrogated 32 core pharmacogenetic genes, including the GST genes analyzed by MLPA, and array-based comparative genomic hybridization (aCGH) was performed on 46 selected multi-ethnic samples. aCGH analyses mapped the mean size of the GSTM1 deletion to ~21 kb and localized the CNV breakpoints to two directly oriented low copy repeats (LCRs; >90% similarity) resulting in full gene deletion. Similarly, the mean size of the GSTT1 deletion was ~52 kb and these narrowed breakpoints also mapped to two directly oriented LCRs with >90% similarity, indicating that common GST deletion allele formation is likely mediated by nonallelic homologous recombination. Additionally, the identified frequencies of GSTM1 and GSTT1 deletion were high in all studied populations. However, the conflicting disease and drug response association studies with GST deletions suggest that future work should focus on additional sequence and copy number variation at the GST loci.

653T

Pharmacogenomics of cardiovascular disease treatment in subjects from the Lithuanian population. *A. Pranculis[1], I. Pepalyt[1], J. Arasimavičius[1], Z.A. Kučinskien[2], V. Kučinskas[1]*. 1) Department of Human and Medical Genetics, Faculty of Medicine, Vilnius University, Vilnius, Lithuania; 2) Department of Physiology, Biochemistry, and Laboratory Medicine, Faculty of Medicine, Vilnius University, Vilnius, Lithuania.

Cardiovascular diseases (CVDs), which are the number one cause of death in developed countries, are caused by a combination of environmental and genetic components which also affect drug efficacy and likelihood of adverse drug reactions (ADRs). The aim of the study was to evaluate the pharmacogenomic peculiarities of the Lithuanian population according to the variants of genes associated with response to medicines used for CVD treatment and use the results to determine the course of further larger scale studies. Allele and genotype frequency distributions were compared to the results reported in Estonian, Russian, Swedish populations as well as data from other Caucasian populations reported in the NCBI database. The samples of the 96 healthy unrelated subjects from the Lithuanian population were genotyped for 770 000 SNPs using Illumina genotyping array. SNP allele and genotype frequencies within the genes associated with pharmacogenetic response to medications used for CVD treatment were calculated. Allele and genotype frequency distribution analysis according to the *VKORC1* -1639G>A showed that compared to other Caucasian populations, Lithuanian patients have a significantly lower frequency of the A allele (33%) and AA genotype (7%) and may thus require higher doses of warfarin to achieve the same therapeutic effect. *CYP2C19* c.681G>A is associated with an increased clopidogrel dose requirement. In our study group A allele frequency in the Lithuanian population was 9% and there were no AA genotype patients in our study group. When we analyzed the allele and genotype frequency distributions of the SNPs in the genes associated with response to treatment with statins a converse pattern emerged. According to the *HMGCR* SNP29 (G allele frequency <2%) people from the Lithuanian population are more sensitive to treatment with statins compared to other northern European populations. Furthermore the analysis of the SNPs in the *SLCO1B1* gene shows that due to a significantly higher frequency of c.521T>C C allele (20%) people from the Lithuanian population are more likely to develop ADRs like myopathy caused by statin treatment compared to other eastern and northern European populations. Even though these results should be replicated in a larger study group, these results indicate that patients from the Lithuanian population who were diagnosed with a CVD might benefit from the treatment with higher doses of anticoagulants and decreased doses of statin class drugs.

654T

Systematic functional characterization of cytochrome P450 2E1 promoter variants in the Chinese Han population. *S. Qin, L. He.* Bio-X Institutes, Key Laboratory for the Genetics of Developmental and Neuropsychiatric Disorders(Ministry of Education), Shanghai Jiao Tong University, Shanghai 200030, China;.

CYP2E1 promoter polymorphisms can lead to significant interindividual differences in expression of CYP2E1. Using a database of CYP2E1 gene polymorphisims established in 2010, our study aimed to functionally characterize the single nucleotide polymorphisms (SNPs) of the promoter region and corresponding haplotypes in the Chinese Han population. 6 novel SNPs and 7 haplotypes with a frequency equal to or greater than 0.01 were constructed on a luciferase reporter system on the basis of site-directed mutagenesis. Dual luciferase reporter systems were used to analyze regulatory activity. The construct including novel SNP mutations exhibited insignificant change in luciferase activity, whereas, the activity produced by Haplo1(GTTGCTATAT), Haplo2 (CTTGCTATAT) and Haplo7 (GAGCTCACAT), containing a -333T>A polymorphism was significantly greater than for the wild type in HepG 2 cells ($p<0.05$), being 1.5-, 2.0- and 1.4- times greater respectively. These findings suggest the possibility of significant clinical prediction of adverse drug reaction and the facilitation of personalized medicine.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

655T

Influence of common polymorphisms in *ABCB1* gene on interaction between methadone and p-glycoprotein. *Y. Teng[1], C. Hung[1,2], M. Chiou[1].* 1) Department of Pharmacy, College of Pharmacy, China Medical University, Taichung, Taiwan, Republic of China; 2) Department of Pharmacy, China Medical University Hospital, Taichung, Taiwan, Republic of China.

Purpose: Drug addiction is one of the most important issues around the world. Lots of the abusers affected by acquired immune deficiency syndrome (AIDS) and need protease inhibitors for therapy. The most common treatment for opioid dependence is substitution therapy with methadone. The present study investigated whether interaction between methadone and p-glycoprotein would be influenced by common polymorphisms in *ABCB1* gene. Method: Cell lines harbored some combinations of the common polymorphisms on *ABCB1* gene were established to exam the function of p-glycoprotein. Direct sequencing was performed to confirm genotypes of the established cell lines. Effect of methadone on p-glycoprotein function was evaluated by rhodamine123 efflux assay and calcein-AM uptake assay in different genotypes of cells. Fluorescence microscopy assay was also conducted. Results: The expression of p-glycoprotein in different cell lines was confirmed by real-time RT-PCR. Common polymorphisms in *ABCB1* gene did not affect mRNA expression. The function of p-glycoprotein was confirmed by comparisons of intracellular accumulation of rhodamine123 between the parental cells and p-glycoprotein-carrying cells. The inhibition effect of methadone on p-glycoprotein function was demonstrated both in rhodamine123 efflux assay and calcein-AM uptake assay. Treatment of 1 μM methadone significantly inhibited rhodamine123 efflux and increased intracellular calcein fluorescence (both $p<0.001$). Cells carrying different p-glycoprotein genotypes showed distinguished effect from the cells carrying wild-type p-glycoprotein. Conclusion: Due to the inhibition effect of methadone on p-glycoprotein, the concentration of protease inhibitor may be increased. Patients with different *ABCB1* genotypes may experience different drug-drug interaction. These results may be helpful in the future clinical practice.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

656T

Genetic regulation of circulating glycine concentration in the Finnish population. *J. Kettunen[1], T. Tukiainen[1], M. Ala-Korpela[2], J. Eriksson[3], O. Raitakari[4], MR. Järvelin[5], T. Lehtimäki[6], V. Salomaa[7], M. Perola[7], A. Palotie[1,8], S. Ripatti[1,8].* 1) Institute for Molecular Medicine Finland, FIMM, Helsinki, Finland; 2) Computational Medicine, Institute of Clinical Medicine, University of Oulu, Oulu, Finland; 3) Unit of Chronic Disease Epidemiology and Prevention, National Institute for Health and Welfare, Helsinki, Finland; 4) Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland; 5) Department of Epidemiology and Biostatistics, Imperial College London, London, UK; 6) Department of Clinical Chemistry, University of Tampere and Tampere University Hospital, Tampere, Finland; 7) Department of Chronic Disease Prevention, National Institute for Health and Welfare, Helsinki, Finland; 8) Human Genetics, Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, UK.

Glycine is an amino acid which acts as an inhibitory neurotransmitter in the central nervous system but is also required as a co-agonist along with glutamate for NMDA receptor which controls synaptic plasticity and memory. The level of glycine in the spinocerebral fluid and blood is regulated by the glycine cleavage enzyme complex through degradation of glycine by decarboxylation. All four key enzymes of the cleavage system have been linked with non-ketotic hyperglycinemia [OMIM 605899], a recessive Mendelian disease where the patients have severe neurological phenotypes including seizures, profound mental retardation, episodes of delirium, chorea, and vertical gaze palsy during febrile illness or progressive spastic diplegia, optic atrophy and others symptoms with varying severity. Most patients with this disease have defects in the glycine decarboxylase (*GLDC*) gene in chromosome 9. Here, we present genome wide association study for serum glycine levels using 1000 genomes imputation panel where we replicate a known glycine association in chromosome 2 (rs1047891, $p = 5.7*10^{-167}$) and present a new genome wide significant non-synonymous variant (rs143119940, Val735Leu) in the *GLDC*-gene which is associated with the blood glycine levels ($p = 3.4*10^{-11}$, MAF = 0.03) in 9238 population based samples unselected for neurological traits. This variant showed lower enzyme activity by 19% in a previous functional study and thus may contribute to the milder forms of non-ketotic hyperglycinemia where the disease variants have not yet often been identified or other diseases of central nervous system.

657F

Functional and physical interaction of the mitochondrial electron transport chain and fatty acid oxidation and its disruption in fatty acid oxidation deficient animals. *J. Vockley[1], Y. Wang[1], E. Goetzman[1], J. Palmfeldt[2], N. Gregersen[2].* 1) Pediatrics, University of Pittsburgh School of Medicine, Pittsburgh, PA; 2) Aaruhus University School of Medicine, Aarhus, Denmark.

We have recently shown that many of the functions of fatty acid oxidation (FAO) are contained in a multifunctional fatty acid oxidation complex and that this complex is physically and functionally associated with the electron transport chain (ETC) supercomplexes. In the present study, we have used blue native gel electrophoresis (BNGE) and mass spectrometry (MS) to isolate ETC complexes from heart mitochondria form a long chain acyl-CoA dehydrogenase (LCAD) deficient mouse and characterized ETC and FAO function induced by the FAO deficiency. Compared to wild type animals, complexes I, II, III, and V activity were reduced in LCAD deficient animals. Complex II activity was affected the most with a reduction to 20% of normal. A bridging assay that measures flow of electrons from FAO to ETC was reduced by ~90% when palmitoyl-CoA was used as substrate as compared to a 50% reduction in FAO alone. Activity with octanoyl-CoA was unchanged. Accumulation of reactive oxygen species related to complex I function was markedly increased in LCAD (-/-) mice compared to wild type, indicating loss of integrity of this complex. Mass spec analysis identified at least 8 FAO proteins that co-localize with respiratory chain supercomplexes and to a lesser extent with isolated complex I in mitochondrial extracts from wild type mice. These included the gene products of *HADHA, HADHB, HADH, ACAA2, VLCAD, CPT2, CPT1B, ECHS1* and *ACSL2.* BNGE showed disruption or decreased stability of the formation of ETC supercomplexes in LCAD(-/-) mice, and mass spec analysis confirmed that all supecomplex subunits were reduced at least 50% compare to wild type. In addition, the FAO proteins were predominantly associated with isolated complex I. Protein cross linking studies revealed that the mitochondrial trifunctional protein (an NAD dependent enzyme encoded by the HADHA, HADHB genes) directly interacted with the 75 kDa subunit of ETC complex I that contains the NADH binding domain for this complex. Our results confirm our previous description of a multifunctional mitochondrial energy complex that encompasses both the ETC and FAO in a way that promotes metabolic channeling for energy production. Secondary disruption of this complex in the face of primary FAO defects is likely to play a role in the pathogenesis of this heterogeneous group of disorders.

658T

Dysregulation of central nervous system sphingolipid metabolism as a unifying mechanism for neuro-cognitive manifestations of Gaucher disease and HIV infection. *A.F. Elias[1,3], N.J. Haughey[2], K. Mackey[3], T. Pitman[4], A.W. Thomas[5], S. Wakefield[4], S.P. Yang[4].* 1) McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD., Select a Country; 2) Department of Neurology, Richard T. Johnson Division of Neuroimmunology, Johns Hopkins University School of Medicine, Baltimore, MD; 3) Ryan White HIV/AIDS Clinic at Partnership Health Center, Missoula, MT; 4) Department of Medical Genetics, Shodair Children's Hospital, Helena, MT; 5) Montana Cancer Center at St. Patrick Hospital, Missoula, MT.

Gaucher disease (GD) is caused by the inherited deficiency of lysosomal glucocerebrosidase resulting in the accumulation of sphingolipids. Three major phenotypes of GD have been described based on the absence (type 1) or presence (types 2 and 3) of central nervous system (CNS) involvement. Particularly, type 3 GD is a chronically progressive neuronopathic form with variable age of onset. The molecular basis of the neurological manifestations in GD is not well understood. Cytotoxic sphingolipids such as glucosylceramide and glucosylsphingosine accumulate in the CNS in patients with GD types 2 and 3, but not in type 1, and may be involved in neuronal cell death. Similarly, abnormal cerebrospinal fluid (CSF) sphingolipid content has been associated with memory impairment in patients infected with the human immunodeficiency virus (HIV), and the accumulation of cytotoxic ceramide has been implicated in the pathogenesis of HIV-associated neurocognitive disorders (HAND). Here, we present the unusual case of a 35-year-old patient with HIV infection and GD. Shortly after HIV diagnosis in 2007, highly active antiretroviral therapy (HAART) was initiated and resulted in consistent virological suppression. An evaluation for hepatosplenomegaly, thrombocytopenia and anemia led to the diagnosis of GD type 1 in 2011. Soon after, he started enzyme replacement therapy (ERT). Moderate hematological and visceral and mild bone involvement were present at diagnosis. However, neuro-cognitive impairment affecting several areas of function as well as subtle signs of CNS involvement including dysphagia were subsequently noted. HAND would be an unexpected complication of HIV given the short time since infection and effective treatment with HAART. Likewise, course and pattern of organ manifestations would be atypical for type 1 GD. We postulate that his neurological complications result from the convergent effects of GD and HIV on sphingolipid metabolism in the CNS leading to a more severe phenotype similar to type 3 GD. This hypothesis is presented in the context of clinical data including ophthalmological, audiological, radiological and neuro-cognitive studies as well as sphingolipid content in CSF and genotypic correlation. Therapeutic consequences including the combination of ERT with substrate reduction therapy are discussed. A unifying role for sphingolipids in the pathogenesis of neurodegeneration makes GD a potential model for other neurodegenerative disorders.

659F

Autism Spectrum Disorders Functioning Levels and Metabolic Alterations. *N. Arciniegas[1,2], N. Linares[1], M. Cabrera-Rios[4], S. Carlo[1,3].* 1) Instituto Filius, University of Puerto Rico, Rio Piedras, PR; 2) Dept Pediatrics, Hosp La Concepcion, San German, PR; 3) Molecular Medicine Dept, Hosp La Concepcion, San German, PR; 4) Bio Ie Lab, University of Puerto Rico, Mayaguez, PR.

Autism Spectrum Disorder (ASD) is a neurodevelopment disorder characterized by poor socialization skills, communication problems, and repetitive behavior. The diagnosis can be apparent during the early months of age but is usually more visible between the ages of two and three years old. It has been hypothesized that the condition is the result of a neurological disorder that affects normal brain function which alters information processing in the brain. As of now ASD is divided in 5 different diagnoses that include: Autism, PDD-NOS, Asperger syndrome, Rett syndrome and Children Disintegrative disorder. This classification will change in the DSM-V. Several etiologies had been implicated in the diagnosis of autism including metabolic, genetic and neurologic. A pilot project was designed to evaluate the possibility of presence of metabolic alterations in patients with ASD and speech delay. The study consisted of 12 patients: 3 with autism, 3 with PDD-NOS, 3 with Asperger and 3 with speech delay and sensory integration problems. Laboratory evaluation consisted of Urine Organic Acids, Plasma Amino Acids, Lactate, Pyruvate, Carnitine and Acylcarnitine profiles, Karyotype and EEG. Results showed elevation in the lactate to pyruvate ratio in one hundred per cent of patients. Usually, an L: P ratio >25 is considered increased and suggestive of a respiratory chain dysfunction at the mitochondrial level. One patient with autism had an abnormal EEG without clinical manifestations. Four patients showed elevated Free to Total carnitine ratios. All karyotypes, urine organic acids, plasma amino acids were normal and three patients had elevated glutarylcarnitine (C5DC). As previously suggested we conclude that there is mitochondrial etiology implicated in ASD and other related disorders.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

660T

Creatine transporter deficiency: novel mutations and functional studies. *O. Ardon[1,2], X. Yin[2], N. Longo[1,2], A. Shilon-Hadass[3], G. Gabis[3], C. Hoffman[3], M. Tzadok[3], G. Heimer[3], S. Sadan[3], B. Ben-Zeev[3], Y. Anikster[3]*. 1) R & D, ARUP, Salt Lake City, UT; 2) Division of Medical Genetics, Dept of Pediatrics, University of Utah, Salt Lake City, UT; 3) Edmond and Lily Safra Children's hospital and Sackler Faculty of Medicine, TAU, Sheba Medical Center, Israel.

X-Linked cerebral creatine deficiency (MIM 300036) is caused by deficiency of the creatine transporter encoded by the SLC6A8 gene. This disorder is characterized by mental retardation, delays in language and speech, autistic-like behavior, seizures in about 50% of cases, and in some cases mid-facial hypoplasia and short stature in hemizygous males. Here we report the first two patients with this condition from Israel. Both were evaluated for global developmental delays and language apraxia. Borderline microcephaly was noted in one of them. Their non-verbal communication was better than other areas and they had no autistic symptoms. Diagnosis was prompted by brain MRI/MRS evaluations which revealed normal white matter distribution, but absence of the creatine peak in both patients. Biochemical testing indicated normal plasma levels of creatine and guanidinoacetate, but an increased urine creatine/creatinine ratio of 2.9 in both patients (normal values are 0.006 to 1.2). The diagnosis was confirmed by demonstrating the absence of [14]C-creatine transport in fibroblasts. Molecular studies indicated that the first patient is hemizygous for a single nucleotide change substituting a single amino acid (c.619 C>T (exon 3) p.R207W). Expression studies in HeLa cells confirmed the causative role of the R207W substitution. The second patient had a single base insertion (c.1254+1insA) at a splicing site in the intron-exon junction of exon 8. This sequence variant is being evaluated as disease causing variant by mRNA processing studies in fibroblasts. These two patients are the first reported cases of creatine transporter deficiency in Israel, and are both in Jewish families of Sephardi origin.

661F

Mice deficient for Succinyl-CoA Synthetase exhibit functionally significant mtDNA depletion and are a model for mitochondrial disease. *T. Donti[1], M. Ge[1], K. Eldin[2], B. Graham[1]*. 1) Molecular and Human Genetics, Baylor College of Medicine, Houston, TX; 2) Department of Pediatrics, Baylor College of Medicine, Houston, TX.

Over the past few years, mitochondrial DNA (mtDNA) depletion has been implicated in several mitochondrial diseases. Mutations in subunits of Succinyl-CoA Ligase (Synthetase), a component of the citric acid cycle, have been associated with mitochondrial encephalomyopathy and mitochondrial DNA depletion. In humans, mutations in SUCLA2, encoding the ADP-specific beta subunit, and SUCLG1, encoding the alpha subunit, have been identified as causes of mitochondrial disease with mtDNA depletion. The goal of this project is to develop model systems of Sucinyl-CoA Ligase deficiency to study disease pathogenesis and biology of mtDNA maintenance. A gene trap allele of Sucla2 (Sucla2SAβgeo) has been isolated in mouse embryonic stem (ES) cells and used to generate transgenic animals. Sucla2SAβgeo homozygotes exhibit recessive lethality with most mutants dying late in gestation (e18.5). Rare stillborn homozygote pups exhibit significantly elevated levels of methylmalonic acid (MMA) and varying degrees of mtDNA depletion. Histological analysis of mutant placenta reveals increased mineralization and mutant embryos are approximately 25% smaller than wild type littermates. Mutant placenta and embryonic (e17.5) brain, heart and muscle show varying degrees of mtDNA depletion (20–60%), while there is no appreciable mtDNA depletion in mutant liver. Mouse embryonic fibroblasts (MEFs) derived from e12.5 embryos show a 50% reduction in mtDNA content after five passages for mutant compared to wild type. The mtDNA depletion in MEFs and mutant tissues is functionally significant as indicated by reduced steady state levels of COXI by western and by respiratory chain deficiencies. Ongoing and future studies include performing complementation experiments in cells to determine the structural requirements of Sucla2 for mtDNA maintenance as well as expression of affinity-tagged Sucla2 in mutant MEFs to identify interacting proteins by pulldown and mass spectrometry. This mouse model of Succinyl-CoA ligase deficiency and mtDNA depletion will provide insights into the pathogenesis of mitochondrial diseases with mtDNA depletion and into the biology of mtDNA maintenance as well as facilitate the exploration of novel therapeutic strategies.

662T

Role of Medium Chain Acyl-CoA Dehydrogenase in the Metabolism of Sodium Phenylbutyrate. *K. Kormanik[1,2], H. Kang[1,2], D. Cuebas[3], B. Scharschmidt[4], J. Vockley[1,2], A-W. Mohsen[1]*. 1) Department of Pediatrics, University of Pittsburgh School of Medicine, Pittsburgh, 15224; 2) Department of Human Genetics, University of Pittsburgh, Graduate School of Public Health, Pittsburgh, PA 15213; 3) Department of Chemistry, Missouri State University, Springfield, MO 65897; 4) Hyperion Therapeutics, South San Francisco, CA 94080.

Background: Sodium phenylbutyrate is primarily used to treat urea cycle disorders. Following conversion to phenylbutyryl-CoA ester, it is postulated to undergo one round of β-oxidation to give phenylacetyl-CoA and acetyl-CoA. The former is hydrolyzed to the drug active form phenylacetate. Molecular modeling suggests that medium chain acyl-CoA dehydrogenase (MCAD), a key enzyme in straight chain fatty acid β-oxidation, could utilize phenylbutyryl-CoA as substrate. *Methods*: Phenylbutyryl-CoA was synthesized using mixed anhydride method. Recombinant human MCAD was purified to essential homogeneity using our standard technique. To test the ability of MCAD to bind the phenylbutyryl-CoA, the ester was titrated under anaerobic conditions into the purified MCAD and changes in MCAD flavin spectrum were observed. MCAD enzyme activity using phenylbutyryl-CoA as substrate was tested using the ETF (electron transfer flavoprotein) fluorescence reduction assay. To confirm that MCAD is the enzyme primarily responsible for the β-oxidation of phenylbutyryl-CoA in vivo, extracts of normal and MCAD-deficient lymphoblast cells were tested for their ability to reduce ETF in the presence of the ester. *Results*: Phenylbutyryl-CoA induced a decrease of the MCAD flavin absorbance at 447 nm and an increase at 570 nm. These changes are similar to those induced by octanoyl-CoA, its bona fide substrate, and unique to the development of the charge transfer ternary complex. The calculated apparent binding constant (Kd app) for these substrates was 2.16 μM and 0.12 μM, respectively. ETF was reduced by MCAD in the presence of phenylbutyryl-CoA. The MCAD catalytic efficiency and the Km for phenylbutyryl-CoA were 0.2 mM-1.sec-1 and 5.3 μM compared to 4.0 mM-1.sec-1 and 2.8 μM for octanoyl-CoA. While ETF fluorescence reduction activity was detected in extracts of normal cells, it was undetectable in extracts of cells deficient in MCAD. *Conclusions*: Phenylbutyryl-CoA binding to MCAD was confirmed with the observed generation of the charge transfer complex as well as the ability of this ternary complex to transfer the reducing equivalents to ETF and release the enoyl-CoA product. The results are consistent with MCAD playing a key role in phenylbutyrate metabolism. With the drug being investigated for other uses in >30 different trials, its metabolism and mode of action must be carefully investigated in patients with partial or complete MCAD deficiency.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

663F

Common dysfunctional variants of *ABCG2* decrease extra-renal urate excretion and cause hyperuricemia. *H. Matsuo[1], T. Takada[2], A. Nakayama[1], T. Shimizu[3], H. Kasuga[2], H. Nakashima[4], T. Nakamura[5], Y. Takada[6], Y. Kawamura[1], Y. Utsumi[1], H. Ogata[1], M. Nakamura[7], Y. Sakurai[4], T. Hosoya[8], N. Shinomiya[1], H. Suzuki[2], K. Ichida[7,8]*. 1) Dept Integrative Physiol, National Defense Med College, Tokorozawa, Japan; 2) Dept Pharm, Univ Tokyo Hosp, Tokyo Japan; 3) Midorigaoka Hosp, Takatsuki, Japan; 4) Dept Prev Med Publ Health, National Defense Med College, Tokorozawa, Japan; 5) Lab Math, National Defense Med College, Tokorozawa, Japan; 6) Central Res Institute, National Defense Med College, Tokorozawa, Japan; 7) Dept Pathophysiol, Tokyo Univ Pharm Life Sci, Tokyo, Japan; 8) Dept Intern Med, Jikei Univ School Med, Tokyo, Japan.

ATP-binding cassette transporter, sub-family G, member 2 (ABCG2, also known as BCRP) is identified as a high-capacity urate exporter and its dysfunction has an association with serum uric acid (SUA) levels and gout/hyperuricemia risk. Hyperuricemia causes gout and kidney stones, and accelerates development of chronic kidney disease. Generally, hyperuricemia has been classified into urate "overproduction type," "underexcretion type," and "combined type" based on only renal urate excretion, without considering extra-renal pathway such as gut excretion. Here we show that decreased extra-renal urate excretion caused by common dysfunctional variants of *ABCG2* is a novel common mechanism of hyperuricemia. Clinical parameters for urate handling including urinary urate excretion (UUE) were examined in 644 Japanese male outpatients with hyperuricemia. The severity of their ABCG2 dysfunction was estimated by genotype combination of two common *ABCG2* variants, nonfunctional Q126X (rs72552713) and half-functional Q141K (rs2231142). Contrary to the general understanding that ABCG2 dysfunction leads to decreased renal urate excretion, UUE was significantly increased by ABCG2 dysfunction (P=3.60×10^{-10}). Mild, moderate and severe ABCG2 dysfunctions significantly raised the risk of "overproduction" hyperuricemia including overproduction type and combined type, conferring risk ratios of 1.36 (95% confidence interval [CI] 1.09–1.71; P=4.55×10^{-3}), 1.66 (95% CI 1.32–2.10; P=8.58×10^{-6}) and 2.35 (95% CI 1.86–2.97; P=3.32×10^{-7}), respectively. In *Abcg2*-knockout mice treated with uricase inhibitor, oxonate, SUA and renal urate excretion were increased, while intestinal urate excretion was decreased, compared to those of wild-type mice. Together with high ABCG2 expression in extra-renal tissues, these results suggest that common dysfunctional variants of *ABCG2* decrease extra-renal urate excretion, especially in intestines, and cause hyperuricemia. This novel mechanism would have been mistaken for urate "overproduction." Thus, "overproduction type" in the current concept of hyperuricemia should be renamed "renal overload type," which is caused by two different mechanisms, "extra-renal urate underexcretion" and genuine "urate overproduction." This new concept will lead to a more accurate diagnosis and more effective therapeutic strategy for hyperuricemia and gout.

664T

Exploring the phenotype of MMACHC deficiency (*cblC*) in zebrafish with morpholinos and zinc finger nuclease targeted mutagenesis. *J.L. Sloan[1], K. Bishop[2], T.L. Blake[2], R.J. Chandler[1], B.P. Brooks[3], R. Sood[2], C.P. Venditti[1]*. 1) Organic Acid Research Section, Genetics and Molecular Biology Branch, National Human Genome Research Institute, Bethesda, MD; 2) Zebrafish Core Facility, Genetics and Molecular Biology Branch, National Human Genome Research Institute, Bethesda, MD; 3) Ophthalmic Genetics and Visual Function Branch, National Eye Institute, Bethesda, MD.

Cobalamin C deficiency (*cblC*) features a combined impairment of the cobalamin dependent enzymes, methionine synthase and methylmalonyl-CoA mutase. It is caused by mutations in the *MMACHC* gene, which encodes an enzyme that is suspected to participate in cobalamin decyanation and intracellular trafficking. The clinical spectrum of *cblC* is wide and patients can have developmental abnormalities, such as congenital microcephaly, heart defects and intrauterine growth retardation (IUGR). To examine the underlying pathophysiology of this disorder, we have studied the phenotype in zebrafish using two approaches. FITC-tagged morpholinos targeting the cognate ATG and an exonic junction were designed and used to knockdown zebrafish *MMACHC*. Morphants began to display defects at 24 hours, including brain necrosis and diminished movement when compared to controls. The morphants had delayed hatching and were significantly smaller in size, had less blood, smaller heads and eyes, and were neurologically impaired compared to controls. By 96 hours, injected fish displayed pericardial edema. Both morpholinos produced a similar phenotype, which was dose-dependent. The same phenotype was observed in the p53$^{-/-}$ fish, suggesting it is not due to off-target effects. Further investigation using transgenic fish expressing GFP in the CNS revealed underdeveloped brains and absent Rohon-Beard neurons. Additionally, histologic examination revealed a fatty liver and abnormalities in the retina. Metabolic analysis of extracts derived from the morphants showed increased methylmalonic acid (MMA) and cystathionine, and MMA was increased in fish incubated with the precursor, propionic acid (PA). Furthermore, the mutant fish were sensitive to PA, such that they died earlier than the mutant fish not incubated with PA. To extend our morpholino studies, a zinc finger nuclease targeted to *MMACHC* was engineered and an allelic series was created: p.V31fsX78, p.H43PfsX21, p.L44HinsM. We are in the process of breeding and phenotyping these mutant lines. The model presented here is the first animal model of *cblC* and faithfully replicates some of the more severe findings observed in humans. Furthermore, it demonstrates the utility of zebrafish to easily examine aspects of metabolic diseases that will be difficult to study in other organisms, such as embryonic manifestations, and should facilitate the exploration of the underlying pathophysiological mechanisms and testing of new therapies for *cblC*.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

665F

Comparative study of gene expression profiling of skin fibroblasts of Leigh syndrome French Canadian subjects and controls. *J. Tardif, C. Laprise, LSFC Consortium.* Université du Québec à Chicoutimi, Chicoutimi, Québec, Canada.

Background: Leigh syndrome French Canadian (LSFC) is a neurodegenerative disorder caused by a mutation in the leucine-rich pentatricopeptide repeat containing (LRPPRC) gene which is implicated in the stability and the handling of mature mRNAs. This mutation leads to a cell specific decrease in cytochrome c oxydase (COX)(complex IV of the respiratory chain), namely 90% in liver and brain and 50% in muscle, fibroblasts, and kidney. The clinical manifestations are hypotonia, developmental delay, chronic well-compensated metabolic acidosis and high mortality before five years old due to frequent of episodes of severe acidosis often leading to coma. LSFC has a carrier rate of 1/23 in the Saguenay-Lac-Saint-Jean region, due to a founder effect. The goal of this study was to determine and compare the gene expression profile of LSFC and control subjects. Methods: Fibroblasts were isolated from skin biopsies of 12 LSFC subjects and 12 control subjects in the LSFC Consortium Biobank that were age and sex paired. RNA was extracted using QIAGEN RNeasy mini kits and genomic expression analysis was performed using Affymetrix Genechip Hgu133plus2 microarrays. A robust multiarray analysis (RMA) normalization, a “batch effects” correction, and a Smyth's moderated t test were used to select differentially expressed genes (p value < 0.05 and absolute fold change ≥ 1.5) (Affy, Limma, and ComBat libraries (R 2.14.1). Findings: Forty-nine differentially expressed genes were identified and classified into seven biological categories. From those genes, eight are involved in energy production and electron respiratory chain whereas five are involved in lipid and glucose metabolism. Differences in the expression of six genes (ACSS3, ND4, ND6, NDUFA4L2, ANGPTL4 and HES1) in these two categories were validated using real-time PCR. Among those genes, NADH dehydrogenase [ubiquinone] 1 alpha subcomplex, 4-like 2 (NDUFA4L2) is of particularly interest in LSFC. NDUFA4L2 is a hypoxia-induced gene, which reduces mitochondrial oxygen consumption through inhibition of complex I activity. This, in turn, limits the production of intracellular reactive oxygen species that would otherwise occur under conditions of low oxygen. Relevance: These findings suggest that the over-expression of NDUFA4L2 may be a compensatory mechanism involved in the survival of LSFC cells and advances our understanding of LSFC. Further studies should permit a better understanding of the role of NDUFA4L2 in LSFC pathology.

666T

***SUCLG1* Mutations Causing Mitochondrial Depletion Syndrome With Congenital Anomalies.** *H.C. Andersson[1,2], M. Landsverk[3], T. Yeh[2], D. Cejas[2], D. Singh[2], L.J. Wong[3].* 1) Hayward Genetics Center, Tulane Univ Med Ctr, New Orleans, LA; 2) Dept Pediatrics, Tulane Univ Med Ctr, New Orleans, LA; 3) Dept. Molecular and Human Genetics, Baylor College of Medicine, Houston, TX.

A newborn with congenital anomalies (cleft lip/palate, ectopic left kidney and severe ductal-dependent coarctation of aorta) developed marked and unremitting lactic acidosis on day of life 1 while hemodynamically stable with good blood perfusion. Urine organic acids showed massive lactic acid, small amount of methylmalonic acidemia and 3-hydroxypropionic acid. Serum pyruvate was low (0.1mg/dl) and lactate:pyruvate ratio (>40) was markedly elevated. Muscle biopsy demonstrated markedly decreased complex I and IV and normal pyruvate dehydrogenase complex (CIDEM, Case Western). Muscle mtDNA was markedly depleted (27% of control). Sequencing analysis of SUCLG1 showed homozygous c.749A>G(p.E250G) mutation previously reported to have caused infantile mitochondrial DNA (mtDNA) depletion syndrome. Causes of primary lactic acidosis due to mtDNA depletion are not typically considered to cause congenital anomalies but two prior reports have described three patients with SUCLG1 mutations and mitochondrial depletion with congenital anomalies similar to ours. Rivera et al (Mitochondrion 10:362, 2010) described one male with brain cysts, horseshoe kidney and hypospadius and one female with interrupted aortic arch who both had SUCLG1-related mtDNA depletion syndrome. Randolph et al (MGM 102:149, 2011) describe a Mexican female with craniofacial dysmorphisms and R single palmar crease with c.280-1A>G mutation in SUCLG1. We propose that the phenotype of mtDNA depletion caused by SUCLG1 mutations should be expanded to include congenital anomalies. Physicians evaluating infants with lactic acidosis and congenital anomalies should consider the mtDNA depletion syndrome.

667F

A mouse model of mitochondrial complex I deficiency exhibits progressive cardiac disease and seizures. *W.J. Craigen[1,2,3], T. Donti[1], Y.C. Lai[2], S. Ather[3], X. Wehrens[3], B.H. Graham[1].* 1) Molecular & Human Genetics, Baylor Col Med, Houston, TX; 2) Pediatrics, Baylor Col Med, Houston, TX; 3) Molecular Physiology and Biophysics, Baylor Col Med, Houston, TX.

Mitochondria are essential for many fundamental cellular processes, including energy production, ROS metabolism, Ca2+ homeostasis, and apoptosis. Dysfunctional mitochondria frequently result in multi-organ disease that can include neuronal, endocrine, cardiac and skeletal muscle pathology. Mitochondrial complex I is composed of approximately 45 subunits and is the most common abnormality in mitochondrial respiratory chain disease. NADH:ubiquinone oxidoreductase iron-sulfur protein 4 (NDUFS4) is one of the nonenzymatic, nuclear-encoded subunits of complex I that is essential for assembly. Mutations in NDUFS4 cause complex I deficiency and a Leigh-like phenotype in humans that typically results in death within 3–16 months of birth. A gene trap allele of *Ndufs4* (*Ndufs4*$^{SA\beta geo}$) has been used to generate transgenic animals. Although *Ndufs4* mRNA is reduced to 5% of wildtype levels, *Ndufs4*$^{SA\beta geo}$ homozygotes are viable and exhibit a significant partial deficiency of complex I (20–30% of control activity), as well as a less pronounced partial deficiency of Complex III. Unlike other previously reported mouse strains lacking NDUFS4, *Ndufs4*$^{SA\beta geo}$ mice do not die prematurely, but exhibit increased sensitivity to kainate-induced seizures and impaired cardiac function, with a conduction abnormality, bradycardia, and diminished contractile function. Mouse embryonic fibroblasts obtained from the homozygous mutants exhibit decreased cellular respiration that is corrected by treatment with vitamin E. This mouse model will provide insights into the pathogenesis of mitochondrial complex I diseases as well as facilitate the exploration of novel therapeutic strategies.

668T

Enzyme Function Affected by Mutations in the Sucrase-Isomaltase Gene. *J. Huang, Z. Wu, S. Uhrich, C.R. Scott.* Department of Pediatrics, University of Washington, Seattle, WA. 98195-6320.

Congenital sucrase-isomaltase deficiency (CSID) is a rare, autosomal recessive disease, caused by mutations in the sucrase-isomaltase (SI) gene. The SI gene provides instructions for producing the enzyme sucrase-isomaltase. This enzyme is found in the small intestine and is responsible for breaking down sucrose and maltose into their simple sugar components. These simple sugars are then absorbed by the small intestine. Mutations that cause this condition alter the structure, disrupt the production, impair the function of sucrase-isomaltase and prevent the enzyme from digesting sucrose and isomaltose. The SI gene consists of 48 exons and encodes 1827 amino acids; amino acids from 1 to 1007 form the isomaltase domain, whereas from 1008 to 1827 form the sucrase domain. We have examined thirty-three CSID patients worldwide by sequencing the entire SI gene coding region, and found one nucleotide deletion, one presumed splicing error and 30 missense variations. Among all missense variations, 6 of them change the genetic code into a chain termination codon and 24 of them result in amino acid substitutions. Based on current knowledge, either the nucleotide deletion, splicing error or translation termination will obviously bring major transformation of the enzyme structure and lead to complete loss of enzyme function. We focused our study on the 24 amino acid variations. Applying available softwares, we predict how these variations affect enzyme SI activity. The result shows: (1) At least seven variations F15V, A231T, Q930R, I1111V, L1239L, M1523I and S1802T are expected to have no effect or only mild effect on enzyme function; we coincidentally found 5 of the seven have been listed in SNP database (http://genome.ucsc.edu) correspondingly; (2) Of the remaining 17 variations, the following 10 variations P348L, V577G, L741P, F875S, W931R, G1073D, C1531Y, R1544C, T1606I and F1745C are likely involved in abnormal folding that severely damages enzyme SI function; the mutations V577G, G1073D and F1745C have the highest frequencies in CSID patients.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

669F

Hyperammonemia and urea cycle homeostasis in methylmalonic acidemia. *R.J. Chandler[1], H. Morizono[2], P.J. McGuire[3], C.P. Venditti[1].* 1) Genetics and Molecular Biology Branch, National Human Genome Research Institute, National Institutes Health, Bethesda, MD; 2) Research Center for Genetic Medicine, Children's National Medical Center, Washington DC; 3) Genetic Disease Research Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD.

Methylmalonic Acidemia (MMA) is an autosomal recessive disorder of metabolism caused by a deficiency in the mitochondrial enzyme methylmalonyl-CoA mutase. MMA patients can experience life-threatening metabolic instability with intermittent bouts of hyperammonemia; the mechanism(s) underlying the pathology of these episodes are not well understood. Presently, the inhibition of N-acetylglutamate synthetase (NAGS) by propionyl-CoA, a metabolite elevated in MMA, is thought to indirectly decrease carbamoyl phosphate synthetase 1 (CPS1) activity and cause hyperammonemia in the patients. To investigate possible mechanism(s) of pathology, we performed proteomic analysis of liver extracts from a mouse model of MMA using two-dimensional fluorescence difference in-gel electrophoresis (DIGE) and a quantitative analysis using iTRAQ labeling and tandem mass spectrometry. Proteomic results from the liver of MMA mice revealed decreased levels of the urea cycle (UC) enzymes Cps1, ornithine transcarbamylase (Otc), argininosuccinate synthetase 1 (Ass1), argininosuccinate lyase and arginase compared to controls. The mRNA levels of *Nags* and *Otc*, but not other UC enzymes, were significant reduced in the livers of the MMA mice. To extend these observations, livers from six metabolically stable MMA patients were studied. Immunoblots of patient liver extracts showed normal levels of CPS1, OTC and decreased levels of ASS1. However, the specific activity of CPS1 and OTC were increased in patient livers relative to controls, indicating activation at the enzyme level. Since deficiencies in the urea cycle enzymes cause hyperammonemia, we hypothesize that increased metabolic stress disrupts the normal homeostasis of the UC and contributes to the intermittent hyperammonemia observed in MMA.

670T

Propionic acidemia: Is there a decreased in tricarboxylic acid cycle enzyme functions? *K. Chapman, A. Markowitz-Shulman, G. Cunningham, J. Cabrera-Luque, M. Summar.* Genetics and Metabolism, Children's National Medical Center, Washington, DC.

Background: Propionic acidemia (PA, OMIM #606054) is an organic acidemia caused by dysfunction of propionyl CoA carboxylase leading to accumulation of propionic acid and decrease of down-stream intermediates including the tricarboxylic acid cycle (TCA) intermediate, succinyl CoA. Objectives: Thus, TCA cycle function can be decreased in PA explaining, in part, the decreased energy phenotype seen in patients. Design/Methods: Lymphoblastoid cell lines (LCLs) derived from presumably healthy controls and PA patients were grown in low glucose (about 9 mg/dL) and normal conditions (about 120 mg/dL) for 1 or 3 days and the levels of OGDH and FH were semiquantified by Western blotting. Results: Our studies demonstrate that total protein levels do not change under low glucose conditions compared to regular growth conditions for lymphoblastoid cell lines from individuals with PA. In addition the level of the TCA cycle enzyme, oxoglutarate dehydrogenase (OGDH) as measured semi-quantitatively by Western blot when cells are grown in low glucose conditions as compared to typical conditions is unchanged. Unlike OGDH, fumarate hydratase (FH) levels are increased in low glucose conditions compared to typical conditions after one day. Despite increases in total protein between days one and three, fumarate hydratase amounts are decreased day 3 compared to day 1. Conclusions: Cells from PA individuals are programmed, on the short term, to bypass the site of their decreased intermediate (succinyl CoA) by increasing protein distal to the block (FH), while leaving protein levels proximal to the block (OGDH) less affected.

671F

Cap Myopathy - A Distinct Form of Congenital Myopathy. *S.J. Piteau[1], G. Smith[1], S. Ludwin[2], J. MacKenzie[1].* 1) Pediatrics, Queen's University, Kingston, Ontario, Canada; 2) Pathology, Queen's University, Kingston, Ontario, Canada.

Cap myopathy is a congenital myopathy charaterized by cap structures located at the the periphery of the muscle fiber. Cap structures consist of disarranged thin filaments with enlarged Z discs. We describe a case of cap myopathy in a 7 year old boy and contrast it with 16 other cases identified in the literature. Our case presented at birth with hypotonia and weakness, and he subsequently developed respiratory failure in infancy requiring biPAP at night. He is ambulatory but has increasing fatigue requiring a wheelchair by mid-afternoon. His initial biopsy at 3 months of age showed fiber type disproportion with Type 1 atrophy; a repeat muscle biopsy at 6 years of age was consistent of cap myopathy. Sequencing for the TMP2 gene revealed no mutation. Mutation analysis for the TMP3, ACTA1, and TNNT1 genes is being considered. The clinical presentation and natural history of cap myopathy is variable exhibiting some phenotypic overlap with nemaline myopathy. Our patient has a more severe phenotype than the majority of reported cap myopathy cases in the literature. Our case supports the identification of cap myopathy as a distinct form of congenital myopathy and further elucidates some of the features of this disease.

672T

Mutations in the phospholipid remodeling gene SERAC1 impair mitochondrial function and intracellular cholesterol trafficking and cause dystonia and deafness. *E. Morava[1], S. Wortmann[1], F. Vaz[2], T. Gardeitchik[1,3], L.E.L.M. Vissers[4], J. Schuurs-Hoeijmakers[4], R. Rodenburg[1,3], A. Grünewald[5], C. Klein[5], T. Kozicz[6,7], P. van Hasselt[8], M. Harakalova[9], W. Kloosterman[9], I. Baricá[10], K. Naess[11], Z. Krumina[12], C. Gilissen[4], H. Bokhoven[5,6], J.A. Veltman[4], J.A.M. Smeitink[1], D.J. Lefeber[13], R.A. Wevers[3], A.P.M. de Brouwer[4].* 1) Institute of Genetic and Metabolic Disorders and Department of Pediatrics, Radboud University Nijmegen Medical Center, Nijmegen, Netherlands; 2) Department of Clinical Chemistry and Pediatrics, Laboratory Genetic Metabolic Disease, Academic Medical Center, Amsterdam, The Netherlands; 3) Laboratory of Genetic, Endocrine and Metabolic Diseases (LGEM), Department of Laboratory Medicine, RUNMC, Nijmegen, The Netherlands; 4) Department of Human Genetics, RUNMC, Nijmegen, The Netherlands; 5) Section of Clinical and Molecular Neurogenetics, Department of Neurology, University of Lübeck, Lübeck, Germany; 6) Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, The Netherlands; 7) Department of Animal Physiology, Radboud University Nijmegen, Nijmegen, The Netherlands; 8) Department of Metabolic Diseases, Wilhelmina Children's Hospital Utrecht, University Medical Center Utrecht, Utrecht, The Netherlands; 9) Department of Medical Genetics, University Medical Center Utrecht, Utrecht, The Netherlands; 10) Department of Pediatrics, University Hospital Centre Zagreb and School of Medicine, Zagreb, Croatia; 11) Department of Pediatric Neurology, Karolinska University Hospital, Stockholm, Sweden 20Medical Genetics Clinic, Children's University Hospital, Riga, Latvia; 12) Medical Genetics Clinic, Children's University Hospital, Riga, Latvia; 13) Department of Neurology, RUNMC, Nijmegen,The Netherlands.

We defined MEGDEL syndrome as 3-MethylGlutaconic Aciduria with impaired oxidative phosphorylation, Deafness, Encephalopathy, Leigh-like MRI, progressive spasticity and dystonia. Whole exome/Sanger sequencing identified 17 mutations in SERAC1, a gene of unknown function in 18 patients from 16 families. Upon the similarity of MEGDEL and Barth syndrome and the presence of a lipase domain in SERAC1 we performed phospholipid analysis. This revealed abnormal phosphatidylglycerol acyl chain composition, altered cardiolipin subspecies and low bis(monoacylglycero)phosphate levels in patients fibroblasts. As firmly established in Barth syndrome, cardiolipin alterations lead to oxidative phosphorylation dysfunction. Low bis(-monoacylglycero)phosphate levels are known to accumulate intracellular cholesterol, which was confirmed by abnormal filipin staining in patients fibroblasts. SERAC1 was further shown to be localized at the interface between endoplasmatic reticulum and mitochondria, in the mitochondria associated membrane (MAM) fraction, important for the exchange of phospholipids. Complementation of patient fibroblast with wild-type human SERAC1 by lentiviral infection led to a normalization of the phosphatidylglycerol acyl chain composition, linking SERAC1 mutations to the compromised phosphatidylglycerol remodeling. SERAC1 is crucial for both mitochondrial function and intracellular cholesterol trafficking. Our data identify SERAC1 as key player in phosphatidylglycerol remodeling as well as first enzyme in the as yet unknown bis(monoacylglycero)phosphate biosynthetic pathway.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

673F

Somatic cell diagnostic studies may miss some patients with mutations in *MMAA* and *MMAB*, genes responsible for isolated methylmalonic aciduria. *L. Dempsey Nunez[1], M.L. Illson[1], J. Kent[2], Q. Huang[2], A. Brebner[1], D. Watkins[1], B.M. Gilfix[3], C. Wittwer[2], D.S. Rosenblatt[1,3].* 1) Human Genetics, McGill University, Montreal, Qc, Canada; 2) Department of Pathology, University of Utah Medical Center, Salt Lake City, Utah; 3) Department of Medicine, McGill University, Montreal, Quebec, Canada.

The gene products of MMAA and MMAB are required for the intracellular metabolism of cobalamin (Cbl). Mutations in these genes lead to the cblA and cblB class of disorders that are characterized by isolated methylmalonic aciduria. We have been concerned that somatic cell methods of diagnosis may miss patients with mild cellular phenotypes. A high resolution melting analysis (HRMA) assay was developed to rapidly scan the coding exons and flanking intronic regions of the MMAA and MMAB genes for variants. DNA from 96 unaffected reference individuals, 42 patients with complementation confirmed cblB, 72 patients with complementation confirmed cblA, and 181 patients with elevated isolated methylmalonic acid, who could not be diagnosed using complementation analysis, were scanned by HRMA. Suspected variants were confirmed using Sanger sequencing. In the cblA cohort, HRMA correctly identified all previously known mutations as well as an additional 22 variants, 10 of which had not been previously reported. Novel variants included one duplication (C.551dupG, p.C187LfsX3), one deletion (c.387delC, p.Y129YfsX13), one splice site mutation (c.440-2A>G, splice site), 4 missense mutations (c.748G>A, p.E520K; c.820G>A, p.G274S; c.627G>T, p.R209S; c.826A>G, p.K276E), and 3 nonsense mutations (c.960G>A, p.W320X; c.1075C>T, p.E359X; c.1084C>T, p.Q362X). All novel variants, listed below, affect highly conserved residues and are predicted to be damaging. In the cblB cohort, HRMA correctly identified all of the previously reported mutations as well as 9 additional variants, one of which was novel (c.12C>A, p.C4X). Scanning of MMAA in the 181 undiagnosed samples revealed a single novel heterozygous missense change (c.821G>A, p.G274D). Scanning of MMAB in the same samples identified 8 variants in 6 samples. Five of these variants were previously characterized as mutations, while 3 were novel missense variants (c.185C>T, p.T62M; c.398C>T, p.S133F; c.394T>C, p.C132R). The identification of 2 individuals among the undiagnosed samples with compound heterozygous variants in the MMAB gene suggests the existence of an unrecognized atypical MMA phenotype for cblB. This mild phenotype was not detected through somatic cell studies, and demonstrates the need for additional diagnostic approaches such as HRMA to screen clinical samples from patients with isolated MMA.

674T

The clinical spectrum of combined malonic and methylmalonic aciduria, a defect in the intramitochondrial fatty-acid-biosynthesis pathway. *I. Manoli[1], J.L. Sloan[1], J.J. Johnston[2], L. Peller[2], J.C. Sapp[2], L.G. Biesecker[2], C.P. Venditti[1].* 1) Genetics and Molecular Biology Branch, NHGRI, NIH, Bethesda, MD; 2) Genetic Disease Research Branch, NHGRI, NIH, Bethesda, MD.

We have identified acyl-CoA synthetase family member 3 (*ACSF3*) as the gene encoding a malonyl- and methylmalonyl-CoA synthetase residing in the mitochondrial matrix. ACSF3 deficiency causes a form of combined malonic and methylmalonic aciduria/acidemia (CMAMMA). In these patients, methylmalonic acid predominates and malonyl-CoA decarboxylase activity is normal. *ACSF3* mutations occur with a minor allele frequency of 0.005 in ~5,950 control individuals, predicting that CMAMMA has a population incidence of ~1:40,000. To further delineate the phenotypes associated with this enzymopathy, we evaluated fifteen subjects, 1.5–64 years of age, with CMAMMA. Seven subjects were diagnosed in adulthood with neurological or psychiatric manifestations (seizures, memory problems, psychiatric disease and/or cognitive decline, peripheral sensory neuropathy and optic nerve atrophy), after exclusion of vitamin B12 deficiency and other more common diagnoses. Three subjects presented during childhood with symptoms suggestive of an intermediary metabolic disorder (coma, ketoacidosis, hypoglycemia, failure to thrive, elevated transaminases), while the five others had more general findings including microcephaly, dystonia, axial hypotonia, muscle weakness, seizures, mild dysmorphic features, autism and global developmental delay. Three out of ten apparently unaffected siblings were diagnosed with CMAMMA after the index case was ascertained in each family. Missense mutations and one in-frame deletion were located in highly conserved acyl-CoA synthetase motifs, while none of the patients carried two nonsense mutations. One mutation, p.R558W, was present in 13/30 alleles. In one subject only one mutation was identified. In another, no damaging mutations were detected; however fibroblasts from this subject showed increased methylmalonic acid secretion following *in vitro* stimulation. Given the predicted population incidence, it is likely that many patients with ACSF3 deficiency remain undiagnosed. Furthermore, the variable clinical phenotypes observed in the patient cohort mandates the consideration of alternative diagnoses in addition to CMAMMA. The generation of an *ACSF3* mouse model will be required to understand the pathogenic role of mutations and to define disease pathophysiology.

675F

Neurochemical profiles in children with inborn errors of metabolism determined by in vivo 1H-magnetic resonance spectroscopy. *W. Al-Hertani[1], E. Mason[1], T. Tam[1], B. Schmitt[3], S. Blaser[2], H. Branson[2], A. Schulze[1].* 1) Division of Clinical and Metabolic Genetics, Hosp Sick Children, Univ Toronto, Toronto, ON, Canada; 2) Division of Diagnostic Imaging, Hosp Sick Children, Univ Toronto, Toronto, ON, Canada; 3) Dept of Radiology, Med Univ Vienna, Vienna, Austria.

Background: Proton Magnetic Resonance Spectroscopy (1H-MRS) is a useful, non-invasive tool for in vivo assessment of biochemical profiles in the human brain. However, our knowledge about the metabolites and neurochemical profiles in metabolic disease is limited. **Objective:** Quantitative description of neurochemical profiles in children with inborn errors of metabolism (IEM). **Methods:** Retrospective quantitative data analysis from children with 42 different diagnoses of IEM, including i.e. Organic Acidurias (16 patients), Remethylation defects (7), Urea Cycle defects (7), Storage disorders (43) and OXPHOS defects (7), and comparison with age-matched normative data from 2,275 children. MRS data were obtained from basal ganglia (BG) and periventricular white matter (PVW), by the point resolved spectroscopy sequence (PRESS) method: TR = 2000 ms, TE = 35/144 ms, NEX = 128. Post-processing of data and analysis was performed using LCmodel software. **Results:** A number of abnormal neurochemical profiles were identified including low total choline in MMA, Arginase deficiency, MTHFR deficiency and SSADH deficiency, low total creatine in Citrullinemia, Arginase deficiency, and Krabbe, and high Inositol in Citrullinemia and NCL. Observed changes can be related to the pathophysiology of the disease or to the treatment. **Conclusions:** Quantitative neurochemical profiling applying normative data will add to the understanding of the pathophysiology of IEM and will be a useful tool for treatment monitoring.

676T

Analyzing Functional Phenotypes in Fibroblasts from Patients with Clinically-Diagnosed Mitochondrial Disease. *B. Graham[1], T. Donti[1], A. Bessee[2], W. Craigen[1,3], P. Bonnen[1,2].* 1) Dept Molecular & Human Genetics, Baylor Col Medicine, Houston, TX; 2) Human Genome Sequencing Center, Baylor Col Medicine, Houston, TX; 3) Dept Pediatrics, Baylor Col Medicine, Houston, TX.

Mitochondria are intracellular organelles that are essential for many cellular functions including oxidative phosphorylation and energy metabolism. The mitochondrial proteome is derived from two genomes: the nuclear genome and the mitochondrial DNA (mtDNA). Mutations in either mtDNA encoded or nuclear encoded genes cause mitochondrial disease. Fibroblast cell lines derived from patients with clinically-suspected mitochondrial disease provide a convenient resource for research studies, but the perceived low sensitivity of this cell type for displaying abnormal mitochondrial phenotypes has limited their diagnostic use. For this study, a collection of primary fibroblast cell lines from at least 16 unrelated patients with clinically-diagnosed mitochondrial disease but no molecular diagnosis was subjected to a battery of assays for mitochondrial function. Tests include assays for cellular respiration, electron transport chain (ETC) activities, relative changes in mitochondrial membrane potential (MMP), increased reactive oxygen species (ROS), changes in mtDNA copy number, and abnormal mitochondrial morphology. To date, 7 out of 16 (44%) patient cell lines exhibit abnormalities in ETC activities and/or mtDNA content compared to controls with other assays pending. A battery of functional assays in fibroblasts offers the promise of increased diagnostic sensitivity and can ultimately be used to functionally screen candidate disease genes identified from whole exome studies.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

677F

Clinical Molecular Laboratory Analysis of the Genes of Fatty Acid Oxidation. *M.T. Hardison, M. Landsverk, L. Wong.* Molecular and Human Genetics, Baylor College of Medicine, Houston, TX.

Fatty acid oxidation (FAO) is integral in maintaining normal mitochondrial homeostasis and vital to the function of the electron transport chain and subsequent production of adenosine tri-phosphate (ATP). Deficiencies in genes involved in FAO are autosomal recessive disorders that result in a decreased ability to oxidize fatty acids, thereby signifying metabolic dysfunction. Early detection by newborn screening (NBS) through the measure of acylcarnitines using tandem mass spectrometry has increased the number of patients identified with FAO deficiencies through the identification of milder phenotypes. However, acylcarnitine profiles alone cannot confirm a diagnosis; activated mitochondrial fatty acid oxidation during catabolism can lead to false positives. In addition, mutations in these genes can result in phenotypes that overlap with other disorders. Additional tests such as mutation analysis of a suspected gene are often necessary to confirm a proper diagnosis. We analyzed all individuals submitted to our laboratory for clinical molecular testing of genes involved in FAO: ACADM (MCAD), ACADVL (VLCAD), HADHA (LCHAD), HADHB, LCAD, SLC25A20 (CACT), ETFA, ETFB, ETFDH, CPT1A, CPT1B, and CPT2 from January 2007 to December 2011. More than 2135 samples were reviewed with phenotypes ranging from positive newborn screen results to rhabdomyolysis to hypoglycemia. 184 patients were identified with at least two known mutations, either apparently homozygous or in a compound heterozygous state. Additionally, 381 patients were identified to carry a single known mutation and are carriers of FAO diseases. Finally, a total of 164 patients with novel variants were identified in 8 of the 12 genes investigated. Due to the large number of unclassified variants in each of these genes, frequent review of the literature is required to interpret sequencing data of FAO patients. Since the clinical presentation of FAO disorders can be very heterogeneous, this group of genes should be considered in patients with muscle myopathy or rhabdomyolysis. For example, one of the patients in our cohort presented with recurrent rhabdomyolysis and was initially tested for McArdle disease (Glycogen Storage Disease type V) when in fact, she was compound heterozygous for two CPT2 mutations. We are currently developing a next generation sequencing panel consisting of all of the FAO genes that will be a useful tool in the analysis of patients with suspected FAO disorders.

678T

Fabry Disease: Molecular analysis of GLA in females with no family history. *D. Doheny[1], I. Nazarenko[1], J. Kim[1], R. Dobrovolny[1,2], R. Desnick[1].* 1) Dept Gen and Genomic Scienc, Mount Sinai Sch Med, New York, NY; 2) Institute of Inherited Metabolic Disorders Ke Karlovu 2, Praha 2, Czech Republic.

Fabry disease (FD) is an X-linked lysosomal storage disorder due to deficiency of the lysosomal enzyme, α-galactosidase A, resulting in progressive accumulation of globotrioasylceramide (GL-3) and related glycosphingolipids in plasma and tissue lysosomes throughout the body. Classically affected males (<1% α-Gal A activity) present in childhood with acroparesthesias, angiokeratomas, corneal/lenticular opacities, hypohidrosis, and, with advancing age, renal failure and cerebrovascular and cardiovascular complications leading to early demise, whereas later-onset variants with residual α-Gal A activity present in later life with cardiac and/or renal disease without the classic childhood manifestations. Manifestations in heterozygotes (hets) of both the classic and variant forms of FD are variable due to Lyonization and often non-specific. Diagnosis of FD in males with either phenotype is reliably made by demonstrating markedly deficient α-Gal A enzyme activity. In female hets, enzymatic diagnosis can be problematic as their α-Gal A activities are variable, 50% being normal. Thus, accurate diagnosis of hets requires GLA gene analysis, beit targeted analysis of known family GLA mutation or sequencing. Due to symptom variability in hets, females are often presented for FD testing without FD-specific symptoms or family history of FD. We present our experience in FD testing of such females. Routine FD testing, both enzymatic assay and DNA analysis, was performed in a CLIA-approved laboratory. Dosage analysis to detect large GLA deletions/duplications was performed as previously reported (Dobrovolny et al, Hum Genet 2010). Among 317 females with no family history of FD, enzyme assay was performed on 280, sequencing on 193 of which 64 (33%) were found to have a GLA mutation. Based on review of clinical information of those without a mutation (N=129), dosage analysis was performed on 36, identifying three (8%) with large GLA deletions. Extensive review of medical records and clinical information revealed differences in frequency of FD-specific symptoms between those females with and without a GLA mutation, specifically corneal findings and renal morphology, whereas incidence of other FD-related but non-specific symptoms were similar. These findings suggest that corneal findings and renal morphology are primary indicators for FD testing in females without a family history of FD; such clinical findings should be differentiated from phenocopies.

679F

Maple syrup urine disease due to a new point mutation at BCKDHB in Iranian family. *M. Hamid[1], G.R Shariati[2,3], A.H Saberi[2,3], M. Mohammadi[2].* 1) Pasteur Institute of Iran, Tehran, Iran; 2) Narges Medical Genetics & PND Laboratory, No. 18, East Mihan Ave, Kianpars, Ahvaz, Iran; 3) Department of Medical Genetic, Faculty of Medicine, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran.

Maple syrup urine disease (MSUD) is a rare autosomal recessive disorder of branched-chain amino acid (BCAA) metabolism caused by the defective function of branched-chain a-ketoacid dehydrogenase complex (BCKD). The disease-causing mutations have already been described in BCKDHA, BCKDHB and DBT. In this study, we report a couple who first presented to our clinic from south of Iran with Arab ethnic origin. Mutation analysis confirmed that this couple, who were first cousins were heterozygous for a novel mutation, p.E330K (GAG>AAG) mutation in exon 9 at BCKDHB gene. This finding, potentially facilitate prenatal diagnosis during subsequent pregnancy and carrier detection. In addition, this mutation expands the mutation spectrum of this disease.

680T

Peanut Consumption Increases Levels of Plasma Very Long Chain Fatty Acids in Humans. *C.T. Lam[1,5], D.A. Wong[1], S.D. Cederbaum[1,2,3], B. Lim[4], Y. Qu[4].* 1) Department of Pediatrics, University of California Los Angeles, Los Angeles, CA; 2) Department of Psychiatry, University of California Los Angeles, Los Angeles, CA; 3) Department of Human Genetics, University of California Los Angeles, Los Angeles, CA; 4) Department of Genetic Testing, Kaiser Permanente Southern California, Los Angeles, CA; 5) Medical Genetics Branch, NHGRI, National Institutes of Health, Bethesda, MD.

Peanut consumption has been suspected of raising plasma levels of C26:0 and C26:0/C22:0 to levels indistinguishable from that found in patients with peroxisomal disorders, such as X-linked adrenoleukodystrophy. However, this occasionally observed phenomenon has not been systematically studied or documented in humans. The objective of this study is to determine if peanut consumption in the form of peanut butter will significantly elevate plasma levels of very long chain fatty acids into the range that would raise concern for a peroxisomal disorder in humans. We enrolled six subjects in a prospective, IRB-approved protocol where plasma very long chain fatty acids were measured before and after consumption of four tablespoons of peanut butter. We found that peanut butter consumption did significantly elevate plasma C26:0 and C26:0/C22:0 to levels indistinguishable from those seen in patients with peroxisomal disorders. C26:0 level exceeded upper normal limit reaching peak values within 3 to 4 hours after peanut butter consumption in all 6 subjects. The C26:0/C22:0 ratio also exceeded the upper normal limit in all 6 subjects and reached the peak value within 3 to 4 hours in 5 out of 6 subjects. Twelve hours after peanut butter consumption, levels of C26:0 and C26:0/C22:0 ratio returned back to the normal range in all 6 subjects. When comparing peak versus baseline mean very long chain fatty acid levels and their related ratios, a statistically significant difference was observed for C26:0 ($p<0.0032$) and C26:0/C22:0 ratio ($p<0.0001$). Although the means are significantly different between baseline and peak values for C22:0 ($p<0.0116$) and C24:0 ($p<0.0107$), the levels of C22:0 and C24:0 in two subjects never exceeded the upper normal limit. Of note, in our study, the participants were allowed to resume a regular diet with the exception of avoiding foods containing peanut oil, olive oil, corn oil, mustard oil, and Lorenzo's oil, and none of the levels of C26:0 or C26:0/C22:0 significantly increased with other food consumption during the study period. Thus we conclude that the consumption of peanuts is a definite cause of elevations of very long chain fatty acids and needs to be taken into account when interpreting these laboratory values. Additionally, this study suggests that dietary causes of elevations in very long chain fatty acids maybe specific to peanuts and foods known to be high in C26:0 content instead of food in general.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

681F

Diagnostic Testing for Multiple Treatable Metabolic Neurodegenerative Disorders by Next Generation Sequencing. *S. van Koningsbruggen[1], O.R. Mook[1], G. Visser[2], M.A. Haagmans[1], R.C. Hennekam[3], G.S. Salomons[4], P. Verwaijen[5], H.R. Waterham[6], F.A. Wijburg[3], M.M. Mannens[1].* 1) Clinical Genetics, University of Amsterdam, Academic Medical Center, Amsterdam, The Netherlands; 2) Department of Metabolic Diseases, Wilhelmina Children's Hospital Utrecht, University Medical Center Utrecht, Utrecht, The Netherlands; 3) Department of Pediatrics, University of Amsterdam, Academic Medical Center, Amsterdam, The Netherlands; 4) Metabolic Unit, Department of Clinical Chemistry, VU University Medical Center, Amsterdam, The Netherlands; 5) Actelion Pharmaceuticals Nederland B.V. Woerden, The Netherlands; 6) Lab. Genetic Metabolic Diseases, University of Amsterdam, Academic Medical Center, Amsterdam, The Netherlands.

The group of Metabolic Neurodegenerative disorders is comprised of a wide variety of disorders, caused by mutations in different genes. Most of these disorders are progressively degenerative and characterized by non-specific symptoms such as mental retardation and epilepsy, which can seriously hamper the diagnosis of patients. Because a subset of the disorders is (partially) treatable, this group of patients would greatly benefit from an early and accurate diagnosis. We have set up a diagnostic assay using Next Generation Sequencing in which we can screen in parallel for multiple genes involved in 50 treatable metabolic neurodegenerative disorders. We have developed a capture array to enrich all coding exons and the flanking intronic regions of 60 involved genes. These sequences are subsequently run on the Roche 454 GS-FLX sequencer and analysed using the Roche mapper software followed by a homemade analysis pipeline. We complement this platform with Sanger sequencing of the homopolymer regions and the low coverage regions such as some of the high GC rich first exons. Using this combined approach, we are now able to obtain and analyse the complete coding sequence and flanking regions of 60 genes within a turnaround time of 2 months. Our diagnostic test for 50 disorders, that can be difficult to discriminate, will allow a much faster diagnosis, which is essential in these treatable progressive disorders. Results of the first series of patients are presently pending. This study was supported by a grant from Actelion Pharmaceuticals Nederland BV.

682T

Prevalence of *GNE* p.M712T and Hereditary Inclusion Body Myopathy (HIBM) in Sangesar Population of Northern Iran. *H. Khademian[1, 2], R. Carbajo[1, 3], E. Mehravar[2], L. Sandoval[1, 4], J. Garcia-Figueroa[1, 4], Z. Khokher[1, 3], D. No[1, 3], Y. Valles-Ayoub[1], D. Darvish[1].* 1) HIBM Research Group, Reseda, CA; 2) Sangesar Charity Organization for HIBM, Sangesar, Semnan, Iran; 3) California State University Northridge, Northridge, CA, USA; 4) Los Angeles Mission College, Sylmar, CA, USA.

GNE myopathy or Hereditary inclusion body myopathy (HIBM) is an ultra-rare severely disabling autosomal recessive adult onset muscle disease which affects roughly 1–3 individuals per million worldwide. Genetically, HIBM is caused by mutations in the glucosamine (UDP-N-acetyl)-2-epimerase/N-acetylmannosamine kinase gene (*GNE*), resulting in diminished enzyme function and reduced sialic acid biosynthesis. A founder variant *GNE* p.M712T was first described in patients of Iranian and Middle-Eastern descent living outside of Iran. Asymptomatic heterozygote or carrier frequency has been reported as high as 1 in 11 within the Persian-Jewish community residing in Los Angeles, California. To investigate the prevalence of the p.M712T variant in Iran, we studied 792 random samples in Sangesar (Mahdishahr) in Northern Iran. DNA samples were obtained by buccal swab, and genotyping was performed by melting curve analysis. The results included 31/792 (3.91%) heterozygous carriers and 5 (0.31%) homozygotes for *GNE* p.M712T. All 5 homozygous individuals, age 30–64 years, were already symptomatic at the start of the study. In addition, 7 of the wild type samples had HIBM-like symptoms; interesting results were found after sequencing the entire *GNE* gene. Our findings suggest that the prevalence of *GNE* p.M712T is higher in the Sangesar population, comprised mostly of Muslim and Bahai descendants, compared to the general world population. Additional studies are warranted to further elucidate the prevalence of the *GNE* p.M712T allele within various sub-populations of Iran.

683F

Clinical application of massively parallel sequencing for molecular diagnosis of genetically heterogeneous Glycogen Storage Diseases. *J. Wang[1], H. Cui[1], NC. Lee[2], WL. Hwu[2], YH. Chien[2], WJ. Craigen[1], LC. Wong[1], VW. Zhang[1].* 1) Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, United States; 2) Department of Pediatrics and Medical Genetics, National Taiwan University Hospital, Taipei, Taiwan.

Glycogen storage diseases (GSDs) are a group of inherited genetic disorders of glycogen synthesis or catabolism that can involve multiple organ systems, including skeletal muscle and the liver. Early diagnosis with proper treatment can greatly improve the quality of life, reduce organ damage, and extend patient's life span. Due to the genetic heterogeneity and limited availability of invasive biopsy for enzyme studies, the definitive diagnosis of GSDs is based on sequence analysis of selected potentially causative genes, which is costly and time consuming, often leading to delayed diagnosis. We developed the next generation sequencing (NGS) based approach by simultaneous capture of 16 genes currently known to cause both muscle and liver forms of GSDs: GYS2, GYS1, G6PC, SLC37A4, GAA, AGL, GBE1, PYGM, PYGL, PFKM, PHKA2, PHKB, PHKG2, PHKA1, PGAM2, and PGM1, followed by massively parallel sequencing (MPS). All nucleotides in the coding regions of 16 genes have been enriched with sufficient coverage in an unbiased manner. A total of 7 samples with known mutations were validated. All disease causing mutations were identified correctly. The mutations include single nucleotide substitution, small deletions and insertion/duplications. The results showed 100% sensitivity and specificity compared to Sanger sequencing. In addition, a total of 17 DNA samples from unrelated individuals with clinical and/or enzymatic findings suggestive of GSD but without identified mutations were analyzed. Deleterious mutations or novel variants in various GSD genes were detected in 11 patients (11/17=65%) and therefore confirmed the diagnosis. Most importantly, the uniform deep coverage of all coding sequence regions allows the detection of large deletions involving single or multiple exons. Homozygous or hemizygous deletions were identified in 3 patients by a lack of coverage in the targeted regions. Heterozygous exonic deletion can also be detected by analyzing the sequence reads using computational programs. Our data underscore the importance and clinical utility of MPS-based analysis in the molecular diagnosis of a defined disorder that may be caused by defects in multiple genes. The high throughput MPS strategy to simultaneously analyze all the genes responsible for liver and muscle forms of GSD greatly improves the diagnosis of GSD in a cost and time efficient manner.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

684T

Findings from the ZOOM genetic screening study of Niemann-Pick disease type C in adults with neurological and psychiatric signs. *P. Bauer[1], D.J. Balding[2], H.H. Klünemann[3], D. Linden[4], D. Ory[5], M. Pineda[6], J. Priller[7], F. Sedel[8], A. Muller[9], H. Chadha-Boreham[9], R. Welford[9], D.S. Strasser[9], J.E. Wraith[10], M.C. Patterson[11]*. 1) Dept Medical Genetics, University of Tuebingen, Tuebingen, Germany; 2) Institute of Genetics, University College London, UK; 3) Regensburg University Clinic for Psychiatry and Psychotherapy, Regensburg, Germany; 4) Psychological Medicine and Neurology, Cardiff University, UK; 5) Washington School of Medicine, St. Louis, Missouri, USA; 6) Hospital Sant Joan de Déu, Barcelona, Spain; 7) Neuropsychiatry, Charité-Universitätsmedizin, Berlin, Germany; 8) Pitié Salpêtrière Hospital, Paris, France; 9) Actelion Pharmaceuticals Ltd, Allschwil, Switzerland; 10) Manchester Academic Health Science Centre, St Mary's Hospital, Manchester, UK; 11) Mayo Clinic, Rochester, MN, USA.

Niemann-Pick disease type C (NP-C) is a rare, autosomal-recessive, fatal progressive neurological disease caused by mutations in either the *NPC1* gene (in 95% of cases) or the *NPC2* gene. Late-onset cases of NP-C (at ≥15 years of age) are being recognised with increasing frequency. Psychiatric manifestations are prevalent among adult patients. The ZOOM study was designed to evaluate the frequency and phenotypes of NP-C in adult patients with neurological and psychiatric symptoms. Consecutive patients aged ≥18 years with psychosis and/or early-onset progressive cognitive decline combined with neurological and/or visceral symptoms were recruited from psychiatric and neurological reference centres. Diagnostic testing for NP-C involved *NPC1* and *NPC2* exonic gene sequencing and multiplex ligation-dependent probe amplification (MLPA) analysis. When available, results of filipin staining were collected. Plasma oxysterol (cholestane-3β,5α,6β-triol) and other metabolites are also being investigated as potential biomarkers for NP-C. *NPC1* and *NPC2* gene sequencing was completed in 250/256 enrolled patients from 30 centres in the EU and USA (median [range] age 38 [18–90] years). Three patients had a confirmed diagnosis of NP-C: two based on gene sequencing alone (known causal disease mutations) and one based on gene sequencing and positive filipin staining. The latter patient had one known causal disease mutation and one unclassified variant in the *NPC1* gene. All three NP-C positive patients had early-onset cognitive decline with or without psychosis combined with characteristic neurological symptoms of NP-C. A further 12 patients displayed either single mutant NP-C alleles (8 with *NPC1* mutations and 3 with *NPC2* mutations) or a known causal disease mutation and an unclassified *NPC1* allele variant (n=1). All 12 of these patients had either psychosis (n=9), early-onset progressive cognitive decline (n=1) or both (n=2), and 9 of them also had neurological symptoms. Overall, the NP-C carrier frequency in this patient sample was higher than in the general population. Further analyses are currently ongoing to investigate whether the combination of clinical manifestations, genetic analysis and filipin staining (when available), associated with measurement of oxysterol and other metabolite levels, could help to improve the diagnosis of NP-C in this patient population.

685F

Detection of mucopolysacchariduria in dried-urine filter paper samples: an useful technique for regions where transportation of liquid samples is challenging. *R. Giugliani[1,2,3], F. Bender[1,2,3], G. Marasca[1,2], R. Guidobono[1,2], J. De Mari[1,2], M. Burin[1,2,3], G. Civallero[1,3]*. 1) Dep Genetics / Med Genet Serv, HCPA/UFRGS, Porto Alegre, RS, Brazil; 2) MPS Brazil Network, Porto Alegre, RS, Brazil; 3) INAGEMP - Brazilian National Institute of Population Medical Genetics, Porto Alegre, RS, Brazil.

The mucopolysaccharidoses (MPS) are caused by storage of glycosaminoglycans (GAGs) and other complex molecules in tissues, among other pathogenic mechanisms. New therapies are available or are in development for these pathologies, and early diagnosis seems to be important for the therapy outcomes. Almost all patients have increased levels of GAGs in urine. However, test of urine may be challenging as transportation of liquid urine samples in appropriate conditions for long distances, especially across international borders, could be difficult. Objectives: To standardize and validate a method for collection and analysis of GAGs levels in urine spotted in filter paper. Methods: Random urine samples were collected from untreated MPS patients and from controls. To part of the control samples it was added chondroitin sulphate in different concentrations. An aliquot of the urine sample was refrigerated at 4°C while another aliquot was impregnated in the filter paper. The assay of GAGs in liquid urine or in DUFP samples was performed according to the method of De Jong et al, 1992, with slight modifications. Specifically for DUFP samples, 15 paper circles of 6.0 mm diameter were diluted in 1.1 ml of distillated water and incubated overnight at 4°C. After incubation, 125 ul of this solution were used for the assay of GAGs and 50 ul for creatinine determination. Non-parametric tests were used to measure differences between creatinine and GAGs assay in urine and DUFPs, and differences between GAGs concentrations in normal and affected subjects. Results: Concentrations of creatinine and GAGs in liquid urine and on their corresponding DUFP samples from normal subjects were analyzed. No statistically differences were found in these assays ($p<0.05$). Added chondroitin sulfate was similarly recovered from liquid urine or DUFP samples. Concentrations of GAGs in urine or DUFP from normal controls and MPS patients showed similar discrimination between groups. Concentration of GAGs in MPS patients never overlapped with that of normal controls. Our results showed no differences in the concentration of GAGs between liquid and dried samples. Untreated patients with MPS and normal controls were well discriminated using any of the samples. Conclusion: the assay of GAGs in dried-urine filter paper samples is a helpful method for the identification of patients with MPS, and that its incorporation in reference laboratories should be considered.

686T

Analysis of non-reducing ends (NRE) of glycosaminoglycans for the diagnosis and monitoring of therapy in mucopolysaccharidoses type I and type II. *T. Yuzyuk[1], E. Schwarz[2], J. Brown[4], B. Crawford[4], M. Pasquali[1,2,3]*. 1) Department of Pathology, University of Utah Health Sciences Center, Salt Lake City, UT; 2) ARUP Institute for Clinical and Experimental Pathology ®, Salt Lake City, UT; 3) ARUP Laboratories Inc., Salt Lake City, UT; 4) Zacharon Pharmaceuticals Inc., San Diego, CA.

Mucopolysaccharidoses Type I and Type II (MPS I and MPS II) are lysosomal storage disorders caused by deficiency of alpha-L-iduronidase and Iduronate-2-sulfatase, respectively. The deficiency of these enzymes impairs the degradation of glycosaminoglycans (GAGs), with accumulation of dermatan and heparan sulfate. Both disorders affect multiple organs and, in some cases, the central nervous system. Enzyme replacement therapy is available for both conditions and better biomarkers are necessary for monitoring treatment. Traditional methods analyze intact GAGs and cannot distinguish between the two mucopolysaccharidoses and provide only limited specificity to monitor response to treatment. GAGs are degraded in the lysosome from their non-reducing ends and deficiency of an enzyme causes the accumulation of a modified sugar in this position specific for each disorder. The Sensi-Pro® Non-Reducing End (NRE) assay can identify and quantify (by UPLC-MS/MS) the non-reducing ends of the accumulated GAGs after digestion with enzymes specific for each GAG. Here we show that this NRE assay distinguished the GAGs accumulated in patients with MPS I and MPS II: the NREs of heparan and dermatan sulfate accumulated in these two mucopolysaccharidoses differed in mass to charge ratio and/or in retention time. Normal controls had no accumulation of these species. In addition to NREs, this method generates fragments corresponding to internal disaccharides that could be used as a measure of total GAGs. Our results show that analysis of GAGs by analysis of NREs can identify biomarkers specific for each mucopolysaccharidosis and allow the simultaneous measurements of total GAGs. These biomarkers can greatly increase our ability to diagnose and monitor response to treatment in patients with mucopolysaccharidoses.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

687F

Rare, bizarre, and complex illnesses with unusual medical events in patients should be a signal to trigger a mitochondrial work-up: clinical decision tree for mitochondrial disorder work up is provided. *A. Perszyk.* Pediatric Multispecialty Ctr, Univ Florida, Jacksonville, FL.

Mitochondria produce energy for all the cell's processes. When a sudden or acute event occur these changes might first sign of cellular malfunctioning. The more an event is rare, bizarre, unexpected one should be keen to study that patient more fully to ascertain the reason. Complex illnesses involve multiple organ systems. Any unusual medical event, sudden change or unexpected illness may have a more hidden etiology then the most obvious or apparent explanation. In patients with any such events, the clinician should begin a mitochondrial work-up. Decision tree for assistance in suspecting mitochondrial disorders is presented. These are the reasons for genetics consultation that lead to an eventual diagnosis of an energy pathway defect or inborn error of metabolism in either mitochondrial oxidative phosphorylation, Krebs cycle defect, or a fatty acid oxidation defect. Any complex medical disease can mimic mitochondrial disorders. Secondary mitochondrial disturbance and disease progression can be the basis of additional interventions that can benefit patient health and may in some instances reverse recent decline in patient health. Mitochondrial diseases must be diligently sought and work up for these must be done before concluding that the symptoms in question are not mitochondrial in origin. The hospitalist and office practitioner should be keen to do a mitochondrial metabolic work-up for previously healthy patients that have sudden episodes of mental status changes, ataxia, unexplained weakness, hearing loss or idiopathic thrombocytopenia. We have made diagnoses in patients that first presented with cardiac arrhythmia, electromechanical dissociation, enlarged kidneys, renal tubular acidosis, ascites without hepatomegaly. Recurrent fevers, recurrent aspiration pneumonia, osteopenia and recurrent fractures, depression, anorexia nervosa, and cyclic vomiting syndrome. Severe seizure disorder, central apnea, lactic acidosis, neurocognitive decline and patient referral to hospice end-of-life care. Mitochondrial disorders can be misdiagnosed as polyneuropathy, HLH, abnormal CPK, post-anesthetic complications with respiratory insufficiency or rhabdomyolysis, fasting induced hypoglycemia, recurrent rhabdomyolysis, and epileptic neurodegeneration. Interventions have improved outcomes in our patients, but the key is to know when to first suspect that a mitochondrial disorder may be present.

688T

Clinical re-sequencing of 450 genes to diagnose mitochondrial disorders: Results from the first 119 patients. *J. DaRe[1], V. Vasta[2], T. Tran[2], J. Penn[1], L. Susswein[1], J. Booker[1], S. Hahn[2,3].* 1) Transgenomic, Inc., New Haven, CT; 2) Seattle Children's Hospital Research Institute, Seattle, WA; 3) University of Washington, School of Medicine, Seattle, WA.

The diagnosis of mitochondrial disorders (MD) is very challenging as they are clinically and genetically heterogeneous. We recently implemented a clinical next-generation sequencing (NGS) test to diagnose MD targeting over 450 nuclear genes. The genes were categorized into four groups; (1) encoding entire subunits of the mitochondrial respiratory chain complex (RCC), assembly factors; (2) mitochondrial DNA transcription/translation and biogenesis factors; (3) other enzymes and carrier proteins;(4) genes causing secondary mitochondrial dysfunction or presenting with similar phenotypes. After removing variants with a relatively high minor allele frequency in our cohort or public databases, an average of 7 variants per patient was thought to be potentially pathogenic. Most of these were unpublished, single heterozygous variants of unknown significance in genes with autosomal recessive (AR) disease inheritance, resulting in possible carrier status for the corresponding disease. In a total of 64 (54%) cases, we identified variants considered to be possibly disease-causing, with 11 cases involving genes with AR inheritance, 38 cases with autosomal dominant (AD) inheritance, 4 cases with X-linked inheritance, 8 cases with positive findings in both AR, AD, and/or XL disease genes, and 3 cases in genes that have both AR and AD disease inheritance. Subsequent parental testing decreased the number of cases with plausibly causative mutations cases to 47 (39%), though incomplete penetrance or differences in expression remain possible. Of note, we found only 3 cases with variants in category 1 (RCC subunits/assembly factors). Most variants were in category 4 genes. These include OTC (urea cycle disorder), sulfite oxidase deficiency, and Fabry disease, spastic paraplegia, spinocerebellar ataxia, CLN6 (neuronal ceroid lipofuscinosis), and UBE3A (Angelman syndrome). The genes that we found the second most variants in were category 2, possibly indicating that mutations in these genes may be more prevalent in primary mitochondrial disease than mutations in RCC subunits or assembly factors. Some cases had positive findings in multiple genes, indicating that complex clinical phenotypes may be caused by oligogenic mutations. While NGS shows promise for diagnosing suspected MD patients, the challenges remain high as the underlying genetic heterogeneity may be greater than suspected and defining the significance of variants detected requires additional studies.

689F

Insulin Resistance: Development of a clinical assay for *INSR* mutation detection. *M. Procter[1], O. Ardon[1], A. Carey[1], H. Amiri[1], N. Longo[1,2], R. Mao[1,2].* 1) ARUP Institute for Clinical and Experimental Pathology, Salt Lake City, UT; 2) University of Utah Department of Pathology, Salt Lake City, UT.

Background : Insulin resistance is a pathological condition decreasing insulin's ability to lower blood sugar levels. In response to insulin resistance, the pancreas tries to compensate by increasing insulin secretion. Over time, pancreatic beta cells fail and are unable to synthesize/secrete the needed insulin. When this happens, hyperglycemia and diabetes result. Insulin resistance can be localized at a pre-receptor, receptor and post-receptor levels. The insulin receptor is a transmembrane receptor activated by insulin and encoded by a single gene (*INSR*) with alternate splicing resulting in 2 isoforms. *INSR* is located on the short arm of chromosome 19 and composed of 22 exons. Mutations in *INSR* cause rare insulin resistant syndromes, the most severe of which are Donohue Syndrome (leprechaunism) and Rabson-Mendenhall Syndrome, both of which result in growth restrictions, dysmorphic features, altered glucose homeostasis, and early demise, Donohue patients by age 1 year and Rabson-Mendenhall patients up to 17 years. Both conditions are inherited as autosomal recessive traits. Mutations in *INSR* are also implicated in familial hyperinsulinemic hypoglycemia type 5 (HHF5) as well as having a possible role in diabetes type 2. **Methods:** PCR amplification of the 22 coding exons of the *INSR* gene was performed using M13-tailed primers. Bidirectional DNA sequencing was performed with BigDye Terminator chemistry and M13 primers and the product was analyzed on the ABI 3100 genetic analyzer. Data analysis was performed using Mutation Surveyor software comparing the sequence to a reference INSR sequence (Genebank NC_000019). In order to validate a clinical assay for *INSR* sequencing, we sequenced 11 samples. Eight of the 11 samples were from a collaborator who had already sequenced them. Four of these 8 samples were from individuals known to carry mutations in the INSR gene while the other 4 of 8 were shown to be normal by sequencing in their lab. The remaining 3 of 11 samples were from unaffected individuals. **Results:** We found 100% correlation for samples which had previously been sequenced and no mutations in normal individuals. The majority of the mutations we found caused premature coding termination. **Discussion:** Donohue Syndrome and Rabson-Mendenhall Syndrome are very rare and difficult to diagnose. Current diagnosis is currently made based on clinical criteria. ARUP now offers a sequencing assay to identify INSR mutations causing these disorders.

690T

Comparative transcriptome profiles in Gaucher disease mice treated with Imiglucerase (imig) or Velaglucerase Alfa (vela). *N. Dasgupta[1], Y-H. Xu[1,2], S. Oh[1,2], Y. Sun[1,2], L. Jia[3], S. Barnes[1], M. Keddache[1,2], GA. Grabowski[1,2].* 1) The Division of Human Genetics, Cincinnati Children's Hospital Research Foundation, Cincinnati, OH 45229-3039; 2) Department of Pediatrics, University of Cincinnati College of Medicine Cincinnati, OH 45229; 3) Advanced Biomedical Computing Center SAIC-Frederick Frederick National Laboratory for Cancer Research NIH Main campus bldg.37, Bethesda, MD 20892.

Gaucher disease (GD) type 1, a lysosomal storage disease, is caused by the defective activity of acid beta-glucosidase (GCase). Intravenous infusion of purified recombinant GCases reverses many disease manifestations in affected GD patients. Two approved treatments, imig and vela, are structurally and functionally similar GCases, but differ at amino acid 495 [H (imig) and R (vela, the natural aa)]. Both drugs are mannose terminated, but vela has more mannoses in its oligosaccharide side chains. The mode of action of the two drugs was evaluated by direct comparison of transcriptome expression profiles in the visceral organs of D409V/null (9V/null) GD mice treated with either imig or vela. The genome-wide transcriptomes were performed in parallel with the Affymetrix Array and Illumina RNA seq platforms. Analyses revealed high tissue-specific correlation between the two drugs. Differences between the two were identified with secondary analyses, i.e., Anova and DESeq, to identify the differentially expressed (DE) genes. RNA-seq identified 290 (spleen), 78 (liver), and 11 (Lung) DE genes from imig-treated vs. vela-treated 9V/null mice (cutoff ±1.5 fold change, FDR= 0.05). With identical RNA samples, Affy mouse arrays showed 97 (spleen) and 1 (liver) DE genes. Additionally, 47 (spleen) DE genes overlapped with the two methods. There were 243 or 50 unique genes identified by RNA-seq or Affy array, respectively. RNA-seq detected 3 times more genes than Affy array, indicating a greater dynamic range of RNA-seq. Cell division/proliferation, hematopoietic system/development, and inflammatory response/macrophage regulation were the top functional groups (p ~e03 to e06). The gene interactions networks derived from the top functional groups in the spleen focused on: 1) down-regulation of erythroid proliferation/development genes including Gata1 and 2) up-regulation of cytokine/macrophage genes. Both networks showed ~50% of DE genes from imig or vela treatment overlapped with the DE genes in the untreated 9V/null vs. shame-treated WT mice, suggesting partial correction in the spleen. Unexpectedly, this direct comparative study uncovered different molecular pathophysiologic processes in the imig- vs. vela-treated mice even though these drugs had nearly identical therapeutic efficacy by histopathologic and glycosphingolipid analyses.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

691F

Enzyme replacement therapy started at birth improves outcome in difficult-to-treat organs in mucopolysaccharidosis I mice. *G. Baldo, F. Mayer, B. Martinelli, F. Meyer, P. de Oliveira, A. Tavares, L. Meurer, T. de Carvalho, U. Matte, R. Giugliani.* Hospital de Clinicas de Porto Alegre, Porto Alegre, Rio Grande do Sul, Brazil.

Mucopolysaccharidosys type I (MPS I) is a disorder characterized by deficiency of the lysosomal hydrolase alpha-L-iduronidase (IDUA) and storage of undegraded glycosaminoglycans (GAG) heparan sulphate and dermatan sulphate. Since we previously observed that in patients with MPS, the storage of undegraded GAG occurs from birth, in the present study we aimed to compare normal, untreated MPS I mice, and MPS I mice treated with Laronidase (1.2 mg/kg every 2 weeks) started from birth (ERT-neo) or from 2 months of age (ERT-ad). All animals were sacrificed at 6 months for analysis. GAG levels were measured and both treatments were equally effective in normalizing them in the viscera (liver, lungs, kidney, heart and testis), but had no detectable effect on the joint. Heart function was accessed by echocardiography, and parameters such as ejection fraction and fractional area change also improved with both treatments. On the other hand, mice treated from birth presented better outcomes in the difficult-to-treat aortas and heart valves, who were thickened in MPS I mice. Surprisingly, both groups had improvements in behavior tests, and normalization of GAG levels in the brain. IDUA injection resulted in detectable levels in the brain tissue 1 hour after administration, suggesting that the enzyme crossed the blood-brain-barrier. ERT-ad mice developed significantly more anti-IDUA-IgG antibodies, and mice that didn't develop antibodies had better performances in behavior tests, indicating that development of antibodies may reduce enzyme bioavailability. Our results suggest that ERT started from birth leads to better outcomes in the aorta and heart valves, as well as a reduction in antibody levels. Some poor vascularized organs, such as the joints, had partial or no benefit and ancillary therapies might be needed for patients. The results presented here support the idea that ERT started from birth leads to better treatment outcomes and should be considered whenever possible, a observation that gains relevance as newborn screening programs are being considered for the treatable LSDs.

692T

Long-term bone mineral density response to enzyme replacement therapy in a retrospective pediatric cohort of Gaucher patients. *B. Bembi[1], G. Ciana[1], L. Deroma[1], A.M. Franzil[2], A. Dardis[1].* 1) Regional Coordinator Centre for Rare Diseases, University Hospital “Santa Maria della Misericordia”, Udine, Italy; 2) Pediatric Department, Ospedale degli Infermi, Biella, Italy.

Osteopenia is described as a relevant sign of bone involvement in Gaucher disease (GD) both in pediatric and adult patients. Furthermore, abnormal bone metabolism is considered to play a role in growth and pubertal delay. To analyze the long-term effect of enzyme replacement therapy (ERT) on bone mineral density (BMD), a retrospective observational study was conducted in a cohort of 18 GD pediatric patients (13 males, 5 females; median age 9.2 years). They received biweekly infusions of 20–60 IU/kg of alglucerase/imiglucerase. Clinical, laboratory and imaging parameters were evaluated every 2 years. According to the International Society of Clinical Densitometry guidelines, a Z-score≤-2.0 was considered pathological. Nine patients (group P0) began ERT during infancy and nine (group P1) during puberty. At baseline, in three patients (16.6 %; 1P0, 2P1) Z-score was≤-2.0 (range -2.47 to -2.25). In patient P0 it normalized after 2 years, while in the 2P1 patients (splenectomised siblings) it persisted abnormal. The remaining 15 patients (83.4 %) always presented a normal value. In group P0, Z-score improved in infancy but showed a significant decrease during puberty, on the contrary it constantly improved in group P1. Furthermore, at baseline group P0 showed a higher median Z-score than group P1: 0.79 (0.38; 1.50) and -1.61 (-2.25; -1.56) respectively. The use of correct BMD standards to interpret bone loss during pediatric age suggests a limited significance of bone loss in these patients. Moreover, the persistence of residual disease activity may affect normal bone growth during puberty in GD populations.

693F

Evaluation of long-term enzyme replacement therapy for children with Fabry disease. *A.M. Hebert[1], T. Taber[1], L. Lacbawan[1], M. Banikazemi[1,2], O. Goker-Alpan[1].* 1) LSD Research and Treatment Unit, O&O Alpan LLC, Fairfax, VA; 2) Comprehensive Center for Lysosomal Storage Disorders, Columbia University Medical Center, New York, NY.

Fabry disease is an X-linked disorder caused by the deficiency of the lysososomal enzyme α-galactosidase-A (a-GalA), and it affects multiple organs including kidneys, heart, eye, and nervous system. There are two recombinant α-GalA preparations for enzyme replacement therapy (ERT), agalsidase alfa (Replagal) and beta (Fabrazyme), but only Fabrazyme gained FDA approval in 2001. Short-term studies in adults indicate limited efficacy of ERT in treating major complications, but reports on the efficacy of ERT in preventing Fabry disease related complications in pediatric population are not available. We report the course of Fabry disease in 12 children (3 male, 8 female; age range: 7–16 years), diagnosed after ERT became available. In this cohort, the median age of onset of Fabry disease symptoms was 6 years (range: 3–9 yrs). All but one was diagnosed during family screening, and three were symptom free at the time of diagnosis. The mutations included c.G146C, c.G680A, c.718-719delAA, c. G730A, c.C748T, and c.G982C. There were 4 sib-pairs, including one set of twins. The most common presenting symptom was neuronopathic pain (7/12), with GI disturbance as the second most common complaint (4/12). The majority (10/12) have received ERT for a mean duration of 28 months (range: 3–94 months). ERT has been well tolerated in all but one male who developed significant infusion-related reactions. Renal assessments have included eGFR (10/12), and urinary protein excretion (4/12). eGFR ranged between 103.0–372.2 ml/min/1.73, and mean total urinary protein was 51 mg/24 hr/m2 (range: 31.2–74.5). Additionally, two females had proteinuria/microabuminuria. Among cardiac assessments, echocardiography was normal (10/10), and sinus bradycardia was recorded in 3 of 10, one female had non-sustained SVT. On ERT, almost all, including two male patients, remained symptom free, while pain was controlled successfully with adjunct pharmacotherapy. Previous reports indicate major morbidity among children with Fabry disease; however, this cohort shows better control of symptoms. This data also suggests that ERT for treatment of Fabry disease is well tolerated and can improve symptoms in the pediatric population, although further study is essential for long-term outcomes.

694T

Combined Therapy in a Patient with Gaucher Type III. *J. Navarrete, P. Galindo, R. Delgado, A. Limon.* Dept Gen, Hosp Sur PEMEX, Mexico City, Mexico.

Female patient 6 years old with Gaucher Disease TypeIII first seen at 2 years old with hepatosplenomegaly, anemia and thrombocytopenia, bone marrow aspiration was done and hematology made the diagnosis of Gaucher disease. Enzyme activity was done in leukocytes finding an activity of beta glucosidase of 0.9Umol/L/h and DNA analysis was done finding the mutation 1448C or L444P homozygous. ERT with imiglucerase 120 mgs/kg was started every other week since she was 3 years old. AFter 6 months of treatment her blood values were normal and her bone crises disappear. Two years later there was a shortage of the enzyme and the patient was without treatment during 3 months, the patient restarted with thrombocytopenia, anemia and bone crises. Oral substrate reduction therapy was given to the patient 100 mgs every day. The blood values became normal after one month of therapy. Three months later we had the enzyme back and we restarted imiglucerase 120 mgs/kg and continued with oral therapy 100 mgs everyday. We present the clinical, hematological outcome of the patient. We think that in some cases combined therapy can work for some patients with LSD's.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

695F

Biochemical Characterization of Myelin Abnormalities in the Corpus Callosum in MPS I Dogs. *S-H. Kan[1], S.Q. Le[1], J.K. Yee[1], M. Guzman[2], N.M. Ellinwood[3], P.I. Dickson[1].* 1) Department of Pediatrics, LA Biomedical Research Institute, Torrance, CA; 2) Department of Pathology, St. Louis University, St. Louis, MO, USA; 3) Department of Animal Science, Iowa State University, Ames, IA, USA.

Mucopolysaccharidosis type I (MPS I) is a lysosomal storage disease that leads to neurological deficits. Brain magnetic resonance imaging frequently shows abnormalities in white matter suggestive of demyelination. Our neuroimaging and ultrastructural studies in the corpus callosum of the MPS I dog also suggest that abnormal myelin may contribute to white matter disease.

We isolated the genu and splenium of the corpus callosum from MPS I dogs (n=4, ~25 months) and unaffected carrier dogs (n=4, 19–30 months). We performed real-time reverse-transcriptase polymerase chain reaction (RT-PCR) for the myelin related genes, including *MBP*, *PLP1*, *MOG, MAG* and *MOBP*, etc. To study the development of myelin in younger dogs, we also evaluated 2 dogs at age 4 months, and 4 dogs at age 6 months. RT-PCR results were further confirmed in protein expression level by western blots and immunohistochemistry. Oligodendrocytes were studied by immunohistochemistry using marker O4. Myelin was purified from 4 adult MPS I-affected and 4 carrier dogs by sucrose gradient centrifugation, and the fatty acid composition was analyzed by GC/MS.

The expression of myelin-related genes PLP1 and MBP was slightly reduced in the affected animals, and the diminished expression of these genes appeared as early as 4 months of age. Western blot from purified myelin fractions confirmed reduction of MBP in affected MPS I dogs. Additional studies are underway.

Biochemical studies in the MPS I dog model show myelin abnormalities in the corpus callosum. These findings suggest that abnormal myelin may play a role in white matter disease in patients with MPS I.

696T

Enzyme replacement therapy in a Gaucher type 3 patient: report of a paradigmatic case. *F. Vairo[1,2], A.D. Dornelles[1], S.D. Mittelstadt[1], M. Wilke[1], D. Doneda[1], C.B.O Netto[1], I.V.D. Schwartz[1,3].* 1) Medical Genetics Service, Hospital de Clínicas de Porto Alegre, Brazil; 2) Post Graduate Program in Genetics and Molecular Biology, Universidade Federal do Rio Grande do Sul, Brazil; 3) Genetics Department, Universidade Federal do Rio Grande do Sul, Brazil.

Introduction: There are three recombinant enzymes available for the treatment of Gaucher disease (GD): imiglucerase, alfavelaglucerase and taliglucerase alfa. Case report: A male GD type 3 patient, 14 years old, genotype p.L444P/L444, diagnosed at 2 years old. He had been treated with imiglucerase for 9 years since the diagnosis. In 2008, however, he presented a severe adverse reaction to imiglucerase, characterized by cough, laryngeal stridor and periorbital edema. The infusions were suspended for 3 months when imiglucerase was restarted with pre-medication and a slower infusion rate. After 5 months, he presented a new adverse reaction with vomiting, tachypnea, cough and periorbital edema. Intradermal testing confirmed IgE-mediated reaction but serological tests were negative. After 2 years and 10 months with no specific treatment and a significant worsening of the clinical picture, taliglucerase alfa was prescribed, with pre-medication and a slower infusion rate. At the first infusion he presented moderate adverse reaction (skin rash, sweating, headache and vomiting), and the infusions were suspended. After two months, alfavelaglucerase was initiated uneventfully. He maintains day-hospital infusions without premedication and shows improvement of clinical and laboratory parameters. Conclusion: This is the first report of the use of alfavelaglucerase in patients with GD type 3. The use of recombinant enzymes is safe for the majority of GD patients but severe reactions may occur. Premedication and slower infusion rate reduce the incidence of adverse reactions but may not solve the problem. This case report further demonstrates the different safety profile among all the recombinant enzymes available for the treatment of GD.

697F

Alteration in liver TGFβ expression and signaling in MPS I canines and the effect of enzyme replacement therapy. *M. Vera, S.Q. Le, S.H. Kan, P.I. Dickson.* Department of Pediatrics, Division of Medical Genetics, Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center, Torrance, CA.

Mucopolysaccharidosis I (MPS I) is a lysosomal storage disease caused by deficiency of α-L-iduronidase enzyme activity leading to impaired degradation of extracellular matrix glycosaminoglycans (GAG) such as heparan sulfate. Enzyme replacement therapy (ERT) with recombinant human α-L-iduronidase (rhIDU) successfully reduces lysosomal GAG storage in MPS I canines and human patients; however, in the canine model anti-rhIDU IgG antibodies develop and reduce the therapeutic efficacy of ERT. Interactions between the extracellular matrix (ECM) and cytokines such as transforming growth factor beta (TGFβ) are likely to be important in modulating the signaling events required to establish immune tolerance to ERT and prevent anti-rhIDU antibody formation. Our research hypothesis is that the ECM alterations in MPS I affect the distribution and bioavailability of TGFβ in the tissues of affected canines and that this altered distribution normalizes with the effect of ERT on the ECM. TGFβ is secreted in the form of a latent complex that is sequestered in the extracellular matrix by indirect noncovalent association with heparan sulfate. In this work we compared the distribution and signaling activity of TGFβ in the liver and spleen of untreated MPS I canines and MPS I canines treated with intravenous rhIDU ERT. We performed immunohistochemistry with primary antibodies against TGFβ, and the lymphocyte markers CD3 (T cells) and CD79a (B cells) and found that untreated MPS I canine livers had increased staining with all markers compared to normal canines, and that this effect was reduced with ERT. We performed Western blot analysis of liver tissue with primary antibodies against TGFβ and the signaling intermediate SMAD2 and found that precursor TGFβ expression was reduced in untreated compared to treated MPS I canines, while SMAD2 phosphorylation was increased in treated compared to untreated animals. These preliminary data suggest that a non-specific inflammatory response occurs in the liver of untreated MPS I canines and that ERT may reduce this inflammation.

698T

Korean experience with Citrin Deficiency. *B.H. Lee[1,2], Y.M. Kim[1], G.H. Kim[2], K.M. Kim[1], H.W. Yoo[1,2].* 1) Pediatrics, Asan Med Ctr, Seoul, South Korea; 2) Medical Genetics Center, Asan Med Ctr, Seoul, South Korea.

Two phenotypes can occur by citrin deficiency, neonatal intrahepatic cholestasis by citrin deficiency (NICCD; OMIM #605814) and adult-onset type II citrullinemia (CTLN2; OMIM #603471), and some patients with NICCD develop CTLN2 in their later lives. More cases continue to be identified in Korea as well. This study evaluated the clinical and molecular genetic characteristics of 24 Korean patients with citrin deficiency. Thirteen patients were male and 11 patients were females. Eighteen patients manifestated as NICCD, whereas six patients as CTLN2. Prolonged neonatal jaundice was noted in 83.3% of the patients with NICCD. Two patients with NICCD were identified by neonatal screening. Hyperammonemic encephalopathy was accompanied in 5 patients with CTLN2. Atypical manifestations such as severe coagulopathy (1 pt), dyslipidemia (2 pts), and peripheral neuropathy (1 pt) were also noted. At initial evaluation, most patients showed abnormal serum liver enzyme levels (89.4 %), hyperammonemia (77.8 %), and hypoproteinemia (77.8 %). All the patients had citrullinemia (350.9 ± 313.9 umol/L; nl, 1–46 umol/L), normo/hyper-argininemia (116.6 ± 68.0 umol/L; nl, 10–140 umol/L), and increased threonine to serine ratios (3.7 ± 2.9; nl, <1.1 umol/L). Methioninemia and galactosemia were observed in 11 and 4 cases, respectively. SLC25A13 mutations were identified in all patients. Five common mutations, IVS16ins3kb (115 alleles, 32.6 %), 851del4 (12, 26.1 %), IVS11+1G>A (8, 17.4%), IVS13+1G>A (2, 4.3 %), and S225X (3, 6.5 %), comprises 87 % of all the mutations. Prolonged jaundice in 6 pts with NICCD was spontaneously resolved by 3 to 7 months of age. Fatty liver or dyslipidemia were noted in 11 patients (84.6%). Recurrent episodes of hyperammonemic encephalopathy, liver cirrhosis and death were noted in 4, 2, and 1 pts with CTLN2, respectively. Liver transplantation was done in 2 pt with CTLN2. None with NICCD progressed to CTLN2 until 3.1 (0.2–8.6) years of age. In conclusion, typical presenting phenotypes and biochemical findings can be used as the hallmarks for the diagnosis of citrin deficiency. Genetic screening for 5 common mutations in SLC25A13 can be recommended for the Korean patients with citrin deficiency for the rapid diagnosis. Considering the poor clinical course of CTLN2, long-term follow-up evaluation for the patients with NICCD is crucial, and more efforts are needed to diagnose the patients with CTLN2 at early stage.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

699F

Molecular diagnosis in mitochondrial respiratory chain deficiency using exome sequencing. *R. Kopajtich[1], T.B. Haack[1,2], B. Haberberger[1,2], K. Danhauser[1,2], J.A. Mayr[4], W. Sperl[4], U. Ahting[5], G.F. Hoffmann[6], M. Tesarova[7], E. Frisch[3], M. Schülke[3], I. Wittig[8], P. Freisinger[9], D. Ghezzi[10], M. Zeviani[10], A. Rötig[11], E. Graf[1], T.M. Strom[1,2], T. Meitinger[1,2], H. Prokisch[1,2].* 1) Helmholtz Zentrum München, Munich, Germany; 2) Technical University Munich, Munich, Germany; 3) Charité University Medical Center, Berlin, Germany; 4) Paracelsus Medical University Salzburg, Salzburg, Austria; 5) Städtisches Klinikum München, Munich, Germany; 6) University Hospital Heidelberg, Heidelberg, Germany; 7) Charles University in Prague and General University Hospital in Prague, Prague, Czech Republic; 8) Goethe-Universität Frankfurt, Frankfurt, Germany; 9) Community Hospital Reutlingen, Reutlingen, Germany; 10) Neurological Institute Besta, Milano, Italy; 11) INSERM, Paris, France.

Advances in sequencing technologies offer the prospect to reduce the uncertainty and length of the diagnostic journey toward a molecular diagnosis for mitochondrial disorders. Sequencing the entire coding sequence of the human genome (exome sequencing) enables the investigation of about 18,000 genes in parallel and is especially promising in genetic and clinically heterogeneous disorders. We applied exome sequencing in combination with stepwise filtering of gene variants combined with functional complementation experiments. For this study 41 unrelated individuals were chosen who display juvenile-onset mitochondrial disorders without pathogenic variants in the mtDNA. Exome sequencing data were filtered using three different criteria: (i) the presence of mutations in known disease causing genes; (ii) mutations in candidate genes that match the biochemical and clinical phenotype; (iii) mutations in genes that code for mitochondrial proteins. In 14/41 cases we identified mutations in known disease-causing genes (BOLA3, NDUFS8, NDUFS3, MTFMT, TRMU, C8orf38, C20orf7, FOXRED1, NUBPL, GFM1) five of which have only recently been shown to be associated with mitochondrial disorders. In 9/41 cases, the exome sequencing approach led to the identification of mutations in novel genes (ACAD9, MTO1, AGK, NDUFB3, EARS2). The pathogenic character of the identified variants could be confirmed by functional complementation experiments in five of them. In 12/41 cases no obvious disease-causing mutations could be identified. Currently we investigate prioritized rare variants in genes which function in the mitochondrial transcription and translation process. In parallel to the functional studies we analyzed novel candidate genes in a collection of 600 DNA samples from patients with mitochondrial respiratory chain deficiencies. We conclude that in a research setting NGS is a promising method for the diagnostic workup of patients suffering from metabolic diseases. We aim to implement exome sequencing as a standard diagnostic procedure. Since for most patients with mitochondrial disorder there is no clear phenotype-genotype correlation, exome sequencing is prioritized over multiple rounds of single gene sequencing. In addition, unbiased genome-wide analysis allows the identification of new disease genes thereby extending our understanding of the underlying pathomechanisms.

700T

Occurrence of lysosomal storage disorders in children with hepatomegaly or hepatosplenomegaly: An institutional experience. *M. Mistri[1], C. Ankeleshwaria[1], A. Bavdekar[2], M. Kamate[3], C. Datar[4], F. Sheth[1], J. Sheth[1].* 1) Biochemical and Molecular Genetics, FRIGE's Institute of Human Genetics, Ahmedabad, Satellite, India; 2) K.E.M. Hospital, 489, Rasta peth, Sardar Mudliar road, Pune- 411011, India; 3) K.L.E.S. Dr. Prabhakar Kore Hospital & Medical Research Centre, Nehru Nagar, Belgaum, Karnataka,India; 4) Sahyadri Medical Genetics and Tissue Engineering Facility, 170/01 Barve Memorial Complex, J.M. Road, Pune- 411005.

Nine hundred and sixty eight subjects were referred from January 2001 to April 2012 for various LSDs. There major presenting clinical sign and symptoms were enlarged liver and/or spleen, regression of milestone, neurodegeneration, cherry red spot, coarse facial features, skeletal abnormalities, hypotonia, seizures, burning sensations and corneal clouding. From these, 236 patients with the signs of hepatomegaly and/or hepatosplenomegaly were investigated for LSDs such as glycolipids and lipids storage disorders, mucopolysaccharidosis, Pompe disease, Sialic acid disorder, galactosialidosis and I -cell disease. Of these, 110 (46.6%) were found to be affected with different LSDs and remaining 126 (53.4%) were found to be normal for afore mentioned LSDs. Among the affected, Gaucher was confirmed in 48 (20.3%) cases followed by NPD-A/B in 17 (7.2%), mucopolysaccharidosis in 18 (7.6 %), GM-1 gangliosidosis in 14 (5.9 %), Sandhoff in 5 (2.1 %), I-cell disease in 4 (1.7%), Galactosialidosis in 2(0.84%) and Pompe in 1(0.42 %). Patients clinically suspected for Gaucher and NPD-A/B were screened for plasma chitotriosidase(ChT), which act as screening marker for these disease and found markedly elevated (2,000–85,496 nmol/ml/hr). Of 48 patients of Gaucher, 39 (81.25%) had elevated chitotriosidase (4,274–85,496 nmol/ml/hr), 1(2.0%) with undetectable chitotriosidase, 2(4.0%) with normal chitotriosidase and in 6 (12.5 %) chitotriosidase was not carried out. From 17 patients of NPD-A/B, 12(70.5%) had elevated chitotriosidase (2,675–19,236nmol/ml/hr), 1(5.8%) had undetectable chitotriosidase, 2(11.8%) had normal chitotriosidase and in 2 (11.8%) chitotriosidase was not carried out. This study demonstrates that nearly 47% of children with hepatosplenomegaly and developmental delay have a high probability of LSDs and should be screened for LSDs and confirmative enzyme study needs to be carried out for future pregnancies and for therapeutic options.

701F

Exome sequencing reveals that mutations in the genes encoding aminoacyl tRNA synthetases (ARS) cause a variety of clinical syndromes. *M.T. Geraghty[1], M.J.M. Nowaczyk[2], P. Humphreys[1], J. Schwartzentruber[3], C.L. Beaulieu[1], A. Smith[4], L. Huang[4], J. Majewski[3], D.E. Bulman[1,4], K.M. Boycott[1] FORGE Consortiom.* 1) Children's Hosp Eastern Ontario, University of Ottawa, Ottawa, Ontario; 2) McMaster University Medical Center, Hamilton, Ontario; 3) McGill University and Genome Quebec Innovation Center, Montreal, Quebec; 4) Ottawa Hospital Research Institute, University of Ottawa, Ottawa, Ontario, Canada.

ARS are ubiquitously expressed essential enzymes required for cytosolic and mitochondrial protein translation. To date there have been few reports of deficiencies of these enzymes causing human disease. Using exome sequencing we identified mutations in genes encoding members of the ARS family in several patients with a variety of clinical presentations. Patient 1, a 12 year old female presented with exercise intolerance, left ventricular noncompaction without cardiac dysfunction and elevated lactate and alanine. Numerous investigations including muscle histology and respiratory chain analysis were normal. Compound heterozygote mutations in the gene encoding Glycine ARS (*GARS*) were found using exome sequencing. Mutations in this gene have previously been reported as causing the autosomal dominant neuropathies CMT2D and dSMA-V. Patients 2 and 3 are male and female siblings who presented in infancy with microcephaly, cortical blindness, developmental delays and a seizure disorder. Brain MRI revealed cerebral white matter and callosal hypoplasia in both. Chromosome microarray and metabolic studies were normal. Exome sequencing identified the same two compound heterozygous mutations in both children in the gene encoding the Lysine ARS (*KARS*). Autosomal recessive mutations in this gene have been reported in a patient with intermediate CMT, developmental delay, dysmorphic features and vestibular Schwannoma. Patients 4 and 5 are a brother sister pair who presented with microcephaly and mild developmental delays associated with spasticity and arreflexia. Additionally the older brother had hepatomegaly with macronodular cirrhosis and pulmonary cysts. Exome sequencing revealed the same two mutations in both siblings in *YARS* (Tyrosine ARS). Mutations in this gene have been described in dominant intermediate CMT. These cases expand the phenotype associated with deficiencies of ARS and illustrate that defects in tRNA metabolism represent an emerging class of human disease.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

702T

Novel mutations in the PC gene in patients with type B pyruvate carboxylase deficiency. *E. Oestergaard[1], M. Duno[1], LB. Møller[2], HS. Kalkanoglu-Sivri[3], A. Dursun[3], D. Aliefendioglu[4], H. Leth[5], M. Dahl[6], E. Christensen[1], F. Wibrand[1]*. 1) Clinical Genetics, Copenhagen University Hospital, Copenhagen, Denmark; 2) Center of Applied Human Molecular Genetics, Kennedy Center, Glostrup, Denmark; 3) Department of Pediatrics, Section of Nutrition & Metabolism, Hacettepe University Faculty of Medicine, Ankara, Turkey; 4) Department of Pediatrics, School of Medicine, Kirikkale University, Kirikkale, Turkey; 5) Department of Pediatrics, Roskilde University Hospital, Roskilde, Denmark; 6) Department of Pediatrics, Odense University Hospital, Odense, Denmark.

We have investigated seven patients with the type B form of pyruvate carboxylase (PC) deficiency. Mutation analysis revealed eight mutations, all novel. In a patient with exon skipping on cDNA analysis, we identified a homozygous mutation located in a potential branch point sequence, the first possible branch point mutation in PC. Two patients were homozygous for missense mutations (with normal protein amounts on western blot analysis), and two patients were homozygous for nonsense mutations. In addition, a duplication of one base pair was found in a patient who also harboured a splice site mutation. Another splice site mutation led to the activation of a cryptic splice site, shown by cDNA analysis. All patients reported until now with at least one missense mutation have had the milder type A form of PC deficiency. We thus report for the first time two patients with homozygous missense mutations with the severe type B deficiency, clinically indistinguishable from other patients with type B form of PC deficiency. The mutations found here are novel; it is noteworthy that four Turkish patients did not have any mutations in common, despite the rarity of PC deficiency. There is thus no evidence for recurrent mutations in the Turkish or other populations.

703F

Bioinformatics and Genotypic evaluation of the IDUA gene and protein in MPS I patients diagnosed in Colombia, Ecuador and Perú. *H. Velasco[1], J. Gonzalez[2], T. Pineda[3]*. 1) Genetic Institute, National University, Bogota Cundinamar, Colombia, MD, MSc; 2) Associate Professor, School of Science, Department of Computational Biology, Pontificia Universidad Javeriana, Bogota Colombia, MSc, PhD; 3) Genetic Institute, National University, Bogota Cundinamar, Colombia, MD.

Mucopolysaccharidosis Type I (MPSI) is an autosomal recessive disorder classified into the “rare diseases” with an incidence of 1:100.000–300.000. Mutations in the IDUA gene lead to deficiency in the enzyme α - L -Iduronidase, which is involved in the breakdown of dermatan and heparan sulphates. These compounds will accumulate in the lysosome and will affect many tissues, including neuronal tissue. Since the isolation and characterization of the IDUA gene by Scott et al in 1992, it has been possible to use molecular tools for MPSI diagnosis. Colombia, Ecuador and Peru lacked a molecular methodology for diagnosis and genetic evaluation of this disorder and there was a total ignorance about the mutational profile of the MPSI in our patients. The aim of this study was to analyze genotypic and with bioinformatics tools the mutations found in the 14 MPSI patients from Colombia, Ecuador and Peru. Of the 15 patients analyzed (13 with severe phenotype and one intermediate), the most common mutation found was W402X with a frequency of 42.8%. 6 patients presented 7 novel mutations, which is a high novel mutational rate in our population (33%). In addition to the experimental findings, we proceeded in the first instance to the bioinformatics validation of the novel mutations found in our populations and then to the creation and validation of a bioinformatics model of the IDUA protein for the analysis of 3 novel missense mutations (Y625C, P385L, R621L) where it was found that all of them are potentially pathogenic and some alter the RMSD and ASA values, therefore they are able to disturb the tridimensional structure of the protein and the IDUA accessibility to the substrates. This is the first study in Colombian, Ecuadorian and Peruvian about the mutational behavior of IDUA gene in MPSI patients where we propose the “COLOMBIA” bioinformatics model for subsequent studies of mutations in IDUA. These results will allow a better understanding about the behavior of the gene and the enzyme providing better approaches in the management of the MPSI.

704T

A founder mutation causing a severe methylenetetrahydrofolate reductase deficiency in Bukharian Jews. *S. Ben-Shachar[1], T. Zvi[1], A. Rolfs[2,3], A. B. Klobus[3], Y. Yaron[1,4], A. Bar-Shira[1], A. Orr-Urtreger[1,4]*. 1) Genetic Institute, Tel Aviv Sourasky Med Ctr, Tel-Aviv, Israel; 2) Albrecht-Kossel-Institute for Neuroregeneration, Medical Faculty, University of Rostock, Rostock, Germany; 3) Centogene GmbH, The Rare Disease Company, Rostock, Germany; 4) Sackler Faculty of Medicine, Tel-Aviv University, Tel Aviv, Israel.

Methylenetetrahydrofolate reductase (MTHFR) deficiency (MIM #236250) is an autosomal recessive disorder caused by mutations in *MTHFR* gene with a well characterized biochemical and clinical presentation. We have diagnosed MTHFR deficiency in two children, each from an unrelated non-consanguineous Jewish family of Bukharian origin. In both children *MTHFR* gene sequencing detected a homozygous c.474A>T (p.G158G) mutation in exon 3. As this mutation creates a potential donor splice site (GT), we tested its splicing effect and detected an abnormal transcript which includes the entire intron 3 of the gene and inserts an early termination codon. We have subsequently tested 196 healthy unrelated Jewish individuals of Bukharian origin and detected 5 carriers for this mutation (carrier frequency: 1:39). We conclude that c.474A>T is a pathogenic founder mutation, prevalent among Jews of Bukharian ancestry. Given the disease severity and the frequency of this allele, we recommend that screening for this mutation is offered to individuals of Bukharian origin, in order to facilitate prenatal or preimplantation diagnosis.

705F

Somatic cell and molecular studies on an atypical patient with combined homocystinuria and methylmalonic aciduria (*cblC*). *A. Brebner[1], C. Ficicioglu[2], D. Watkins[1], T. Pastinen[1,3], D.S. Rosenblatt[1]*. 1) Department of Human Genetics, McGill University, Montreal, Quebec; 2) Children's Hospital of Philadelphia, Philadelphia, PA; 3) McGill University and Genome Quebec Innovation Centre, Montreal, Quebec.

The *cblC* complementation group, caused by mutations in the *MMACHC* gene, is the most common inborn error of vitamin B_{12} (cobalamin) metabolism with over 500 cases known. We describe a patient who came to medical attention at 59 years of age because of squamous cell carcinoma of the tonsil. A metabolic analysis unexpectedly revealed elevated serum methylmalonic acid (MMA) and homocysteine. His serum vitamin B_{12} level was not low. The excretion of metabolites corrected with vitamin B_{12} supplementation. Commercial sequencing of the *MMACHC* gene showed only one novel heterozygous missense mutation, c.158T>C (p.L53P). Somatic cell studies using cultured fibroblasts showed decreased incorporation of both [^{14}C]propionate and [^{14}C]methyltetrahydrofolate into cellular macromolecules, demonstrating abnormalities in the function of both cobalamin-dependent enzymes, methylmalonyl CoA mutase and methionine synthase. Cells from the patient also had decreased synthesis of both adenosylcobalamin and methylcobalamin. Somatic cell complementation confirmed the diagnosis of *cblC*. Genomic DNA and RNA were isolated from patient fibroblasts. RNA was converted to cDNA using reverse transcriptase. Mutations in the *MMACHC* gene were searched for using PCR and Sanger sequencing. Sequencing of the genomic DNA confirmed the presence of the heterozygous c.158T>C (p.L53P) missense mutation. This mutation is at a conserved residue and predicted to be possibly damaging by Polyphen. The c.158T>C mutation was seen in a separate *cblC* patient in compound heterozygosity. This supports its role as a causal mutation in *cblC*. Sequencing of the cDNA identified the c.158T>C mutation in apparent hemizygosity. To date, no second mutation has been identified that could be causing the absence of one allele at the mRNA level. No mutations were found in a putative promoter site (1.8kb upstream the ATG start site) as well as 5' and 3' untranslated regions. These results suggest that the second mutation may be located in a non-coding region of the *MMACHC* gene, and that this mutation is responsible for eliminating expression of one allele.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

706T

Exome sequencing reveals a homozygous *ABCD4* mutation in an adolescent with hyperpigmentation, ischemia, hyperhomocysteinemia, and methylmalonic aciduria. *J.C. Kim[1], A. Brebner[1], N.C. Lee[2], W.L. Hwu[2], Y.H. Chien[2], S. Fahiminiya[3], J. Majewski[1,3], D. Watkins[1], D.S. Rosenblatt[1]*. 1) Department of Human Genetics, McGill University, Montreal, Quebec, Canada; 2) Department of Pediatrics and Medical Genetics, National Taiwan University Hospital, Taipei, Taiwan; 3) McGill University and Genome Québec Innovation Centre, Montreal, Canada.

The proband, a 14-year-old boy, was noted to have hyperpigmentation and gray hair at the age of 4 years. At age 7 years, he had a transient ischemic attack with left limping gait for 1–2 months. MRI/MRAI of the brain revealed decreased flow in the left middle cerebral artery (M2) region. Hyperhomocysteinemia (71.47 μM, reference range 3.4–15.6 μM) with methylmalonic aciduria and low serum vitamin B_{12} (cobalamin) level (152 pg/ml; reference range 179–1132 pg/ml) were noted, suggesting an inborn error of cobalamin metabolism. The incorporations of both [^{14}C]propionate and [^{14}C]methyltetrahydrofolate in cultured fibroblasts were within reference ranges and too high to allow for complementation analysis. However, there was decreased synthesis of both adenosylcobalamin and methylcobalamin and accumulation of unmetabolized cyanocobalamin. Size exclusion chromatography showed that 8–9% of the patient's cobalamin was enzyme bound (reference >50%). Exome sequencing of the patient's genomic DNA demonstrated that he was homozygous for a novel missense mutation (c.423C>G, p.N141K) in the *ABCD4* gene, indicating that the patient falls into the *cblJ* class of inborn error of cobalamin metabolism. We have recently discovered that the *cblJ* and *cblF* defects are phenocopies, and that both result in the failure of cobalamin to exit the lysosome. The c.423C>G mutation affects an asparagine residue that is conserved to zebra fish. It was predicted to be probably damaging by the mutation prediction software Polyphen-2. Segregation analysis showed that the patient's father, mother, aunt and two unaffected siblings were heterozygous carriers of the c.423C>G mutation. Clinically the proband appears to be less severely affected than other reported *cblF* and *cblJ* patients. At the cellular level, high propionate and methyltetrahydrofolate incorporation precluded complementation analysis. Accumulation of unmetabolized cyanocobalamin was less marked than in other *cblF* and *cblJ* patients, the synthesis of adenosylcobalamin and methylcobalamin was greater, and more cobalamin was bound to enzymes. These findings suggest that the c.423C>G mutation is less damaging than previously reported *ABCD4* mutations and that the range of biochemical and clinical phenotypes of inborn errors of cobalamin are broader than previously recognized.

707F

Molecular characterization of the GALK1 gene in Filipino patients with Non-Classical Galactosemia. *C.L.T. Silao[1,2], K.N. Hernandez[1], D.M. Canson[1], S.C. Estrada[1,2]*. 1) Inst of Human Genetics, Manila, Philippines; 2) Dept of Pediatrics, University of the Philippines College of Medicine-Philippine General Hospital, Manila, Philippines.

Introduction Galactosemia is an autosomal recessive disorder resulting from a deficiency in one of the three enzymes (galactokinase, epimerase, or galactose-1-phosphateuridyltransferase) responsible for the metabolism of galactose. Galactokinase (GALK) deficiency is characterized by elevation of blood galactose concentration and diminished galactose-1-phosphate resulting in the production of galactitol. Unlike galactose-1-phosphateuridyltransferase deficiency, GALK deficiency patients do not present with severe manifestations in infancy. Clinical manifestations are relatively mild and include neonatal and childhood cataracts, as a result of accumulation in the lens of galactitol, without hepatocellular dysfunction. The disorder is caused by mutations in the GALK1 gene, located in chromosome 17q24. GALK deficiency has an unknown incidence in the general population but is probably less than 1 in 100,000. No study has been done, so far, to determine the molecular characteristics of GALK1 in Filipino patients with GALK deficiency. Methodology The study included six (10) unrelated Filipino non-classical galactosemia patients detected by the Philippine Newborn Screening Program from 1998 to 2012. To investigate the molecular defects of the GALK1 gene, polymerase chain reaction with subsequent direct sequencing were performed on genomic DNA extracted from dried bloodspots or from peripheral blood of the patients using the Qiagen QIAamp DNA Blood Mini/Midi Kit (Qiagen, SantaClara, CA). Results and Discussion Three novel missense mutations in the GALK1 gene (p.G65S, p.G136R, and p.E284K) were identified. Two patients were found to be homozygous for p.E284K while three were homozygous for p.G136R. Two patients were compound heterozygotes for different mutations (p.G65S/p.G136R and p.G136R/p.R256W). Two patients were identified as heterozygotes for the previously reported c.-22 T>C promoter region mutation, one of whom is the p.G136R/p.R256W heterozygote mentioned above. One patient was found to be heterozygous for an unreported intronic mutation (IVS2 -14 G>A). Conclusion The present results provide information regarding the GALK1 mutations that occur in affected Filipinos and further emphasize the heterogeneity of Galactosemia at the molecular level.

708T

Exome sequencing reveals a frequent Moroccan founder mutation in SLC19A3 as a new cause of early-childhood fatal Leigh syndrome. *H. Smeets[1,2], M. Gerards[1,2], R. Kamps[1,2], I. Boesten[1], E. Jongen[1], B. de Koning[1], K. Schoonderwoerd[3], A. Sefiani[4], I. Ratbi[4], W. Coppieters[5], L. Karim[5], I. de Coo[6], B. van den Bosch[1]*. 1) Unit Clinical Genomics, Dept Clinical Genetics, Maastricht UMC, Maastricht, Netherlands; 2) Research Schools GROW and CARIM, Maastricht UMC, Maastricht, Netherlands; 3) Dept of Clinical Genetics, Erasmus MC, Rotterdam, Netherlands; 4) Centre de Génomique Humaine, Université Mohammed V Souissi, Rabat, Morocco; 5) GIGA, Université de Liège, Liège, Belgium; 6) Dept of Neurology, Erasmus MC, Rotterdam, Netherlands.

Leigh syndrome is an early onset, often fatal progressive neurodegenerative disorder caused by mutations in the mitochondrial or nuclear DNA. Until now mutations in more than 35 genes have been have been reported to cause Leigh syndrome, indicating an extreme genetic heterogeneity for this disorder, but still only explaining part of the cases. The possibility of whole exome screening enables not only mutation detection in known candidate genes, but also the identification of new genes associated with Leigh syndrome in small families and isolated cases. Exome sequencing was combined with homozygosity mapping to identify the genetic defect in a Moroccan family with fatal Leigh syndrome in early childhood and typical MRI abnormalities. We detected a homozygous nonsense mutation in the thiamine transporter SLC19A3 (p.Ser7Ter) leading to nonsense mediated decay. Seventeen additional patients with Leigh syndrome were screened for mutations in SLC19A3. Two unrelated patients, both from Moroccan origin and one from consanguineous parents, were homozygous for the same p.Ser7Ter mutation. One of these patients showed the same MRI abnormalities as the patients from the first family. Although unrelated, all patients came from the same area in North-Morocco. Marker analysis of the region surrounding the mutation revealed the same rare haplotype on 5 chromosomes and half the haplotype on the sixth, confirming a founder effect. Strikingly, the only patient in the second family treated with biotine had a much better life expectancy. Our data shows that SLC19A3 is a new candidate for mutation screening in patients with Leigh syndrome, who might benefit from high doses of biotine and/or thiamine. Especially Moroccan Leigh patients should be tested for the c.20C>A founder mutation in SLC19A3.

709F

SATB2 Acts as an Activator of the UPF3B Gene. *P. Leoyklang[1,2,3], K. Suphapeetiporn[2,3], C. Srichomthong[2,3], S. Tongkobpetch[2,3], H. Dorward[4], A.R. Cullinane[4], M. Huizing[4], W.A. Gahl[4], V. Shotelersuk[2,3]*. 1) 1Biomedical Science Program, Faculty of Graduate School, Chulalongkorn University, Bangkok, Thailand; 2) 2The Center of Excellence for Medical Genetics, Department of Pediatrics, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand; 3) 3Excellence Center for Medical Genetics, King Chulalongkorn Memorial Hospital, Thai Red Cross, Bangkok, Thailand; 4) Medical Genetics Branch, NHGRI (NIH), Bethesda, MD.

Two syndromic cognitive impairment disorders have very similar craniofacial dysmorphisms. One is caused by a heterozygous nonsense mutation of SATB2, a transcription regulator, and the other by heterozygous mutations leading to premature stop codons in UPF3B, encoding a member of the nonsense-mediated mRNA decay complex. We hypothesized that the products of these two causative genes functioned in the same pathway. Here, we demonstrated that the SATB2 nonsense mutation identified in our patient led to a truncated protein that localized to the nucleus, formed a dimer with the wild-type SATB2 and interfered with its normal activity. This suggested that the SATB2 nonsense mutation had a dominant negative effect leading to more severe clinical manifestations, compared to patients with SATB2 haploinsufficiency. We also showed that the patient's leukocytes had significantly decreased UPF3B mRNA compared to controls. A chromatin immunoprecipitation assay demonstrated that the SATB2 protein bound to the promoter of the UPF3B gene. A luciferase reporter assay revealed that recombinant SATB2 protein significantly activated gene transcription using the UPF3B promoter. These findings indicate that SATB2 acts as an activator of the UPF3B gene through binding to its promoter. This study emphasizes the benefit of linking disorders with similar clinical phenotypes to elucidate previously unrecognized mechanisms of interaction.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

710T

Low bone strength is a manifestation of phenylketonuria in mice and is attenuated by a glycomacropeptide-based diet. *R. D. Blank[1, 2], P. Solverson[3], S. G. Murali[3], S. J. Litscher[1, 2], D. M. Ney[3]*. 1) Medicine/Endocrinology, Univ of Wisconsin School of Medicine & Public Health, Madison, WI; 2) GRECC, William S. Middleton Veterans Hospital, Madison, WI; 3) Nutritional Sciences, Univ of Wisconsin, Madison, WI.

Purpose: Phenylketonuria (PKU), caused by phenylalanine (phe) hydroxylase loss of function mutations, requires a low-phe diet plus amino acid (AA) formula to prevent cognitive impairment. Skeletal fragility is a poorly understood chronic complication of PKU. Glycomacropeptide (GMP), a low-phe protein isolated from cheese whey, provides a palatable alternative to AA formula. We sought to determine the impact of the PKU genotype and dietary protein source on bone biomechanics.

Methods: Wild type (WT; $Pah^{+/+}$) and PKU ($Pah^{enu2/enu2}$) mice on a C57BL/6J background were fed high-phe casein, low-phe AA, and low-phe GMP diets between 3 to 23 weeks of age. Following euthanasia, femur biomechanics were assessed by 3-point bending and femoral diaphyseal structure was determined. Femoral *ex vivo* bone mineral density (BMD) and bone mineral content (BMC) were assessed by dual-enengy x-ray absorptiometry (DXA). Sixteen whole bone parameters were used to perform prinicipal component analysis. Data were analyzed by 3-way ANOVA with genotype, sex, and diet as the main factors.

Findings: Regardless of diet and sex, PKU femora were more brittle, as manifested by lower post-yield displacement, weaker, as manifested by lower energy and yield and maximal loads, and showed reduced BMD and BMC compared with WT femora. Regardless of genotype and sex, the AA diet reduced femoral cross-sectional area with a consequent reduction of maximal load compared with the GMP diet. Four principal components accounted for 87% of the phenotype variance and confirmed the genotype and diet effects.

Conclusions: Skeletal fragility, as reflected in brittle and weak femora, is an inherent feature of PKU in mice. This PKU bone phenotype is attenuated by a GMP-based rather than an AA-based diet.

711F

Natural History Study of Patients with Hereditary Inclusion Body Myopathy (HIBM). *J.K. De Dios[1], J. Shrader[2], G. Joe[2], C. Ciccone[1], A. Mankodi[3], J. Dastgir[3], C. Bonnemann[3], M. Bevans[2], D. Draper[1], J. McKew[4], M. Huizing[1], W.A. Gahl[1], N. Carrillo-Carrasco[4]*. 1) Medical Genetics Branch (MGB), National Human Genome Research Institute (NHGRI), National Institutes of Health (NIH), Bethesda, MD; 2) Clinical Center, NIH, Bethesda, MD; 3) National Institute of Neurological Disorders and Stroke (NINDS), NIH, Bethesda, MD; 4) Therapeutics for Rare and Neglected Diseases (TRND), National Center for Advancing Translational Sciences (NCATS), NIH, Bethesda, MD.

Background: HIBM, a rare disorder characterized by progressive muscle weakness, is caused by *GNE* mutations, impairing sialic acid biosynthesis. Weakness, apparent in early adulthood, progresses slowly from distal to proximal, first in lower and subsequently in upper extremities, with relative sparing of the quadriceps. There is no therapy available and patients eventually require a wheelchair. **Methods**: We have evaluated 17 patients (age range, 28 to 58 years) in a prospective, longitudinal, single-center study of HIBM (CT.gov: NCT01417533) to define the natural history, genotype-phenotype correlations, prognostic factors, associated manifestations and potential biomarkers and outcome measures. **Results**: The mean age of presentation was 27 years with a mean diagnostic delay of 10 years (range: 2–19 years). Initial manifestations included falls, gait disturbance, foot drop, and overuse-related hand weakness (n=1). Various genotypes were seen, the most common being homozygous p.M712T (n=3). Decreased mean predicted ankle dorsiflexion (9 %), grip (34%) and knee extension (69%) strength were seen. Grip strength deficits emerged around 7.5 years after onset of first symptoms. The 6-minute walk was 68% of mean predicted distance and directly correlated with Activities-specific Balance Confidence Scale Scores. Less strength for selected muscle groups was also related to less balance confidence and longer duration of disease. MRI of lower extremities showed atrophy of muscles with relative sparing of the quadriceps. Prolonged QTc (n=1) and asymptomatic cardiomyopathy (n=1) were seen, but their association to HIBM remains unclear. CPK levels ranged from 161–1152 and decreased with progression of the disease. Mild elevation of ALT (n=2) and hypoalbuminemia (n=12) were seen; all patients had normal renal function tests. **Conclusions**: The characteristic presentation may be modified by muscle overuse. More data are needed to establish genotype/phenotype correlations, but variability among siblings suggests involvement of other factors. Muscle strength and volume decreased with progression and correlated with functional measures and balance confidence. The heart was the only other organ involved as previously reported (Chai et al 2011). Planning clinical trials for potential therapies (ManNAc and sialic acid) is complicated by diagnostic delays and lack of natural history data. Selection of appropriate biomarkers and clinical outcome measures is under evaluation.

712T

Hepatic carnitine palmitoyltransferase 1A deficiency: Adulthood complications and a Durango, Mexico founder mutation. *N. M. Gallant[1,2], N. Vatanavicharn[1,3], D. Salazar[4], W. R. Wilcox[1,2]*. 1) Medical Genetics Institute, Cedars-Sinai Medical Center, Los Angeles, CA; 2) Department of Pediatrics, UCLA School of Medicine, Los Angeles, CA; 3) Division of Medical Genetics, Department of Pediatrics, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand; 4) Biochemical Genetics Laboratory, Quest Diagnostics Nichols Institute, San Juan Capistrano, CA.

Background: Hepatic carnitine palmitoyltransferase (CPT1A) deficiency is a rare disorder of mitochondrial long-chain fatty acid transport typically presenting with sudden episodes of hypoketotic hypoglycemia, liver dysfunction and hepatic encephalopathy during times of fasting or illness. Scant information is known about complications in adulthood and long-term outcome. Methods: We report four patients in three families with CPT1A deficiency and complications not described previously. Molecular sequence analysis of CPT1A and haplotype analysis were performed. Results: Novel clinical and radiographic findings included intrahepatic cholestasis, recurrent pancreatitis, and abnormal brain MRI with T2-hyperintensity in the bilateral basal ganglia and brainstem. Two patients had persistent renal tubular acidosis in spite of therapy with MCT oil. Three patients from two unrelated families were of Durango, Mexican descent. Molecular sequence analysis of CPT1A showed homozygosity for a common disease-causing variant (Q100X), and all shared the same haplotype on chromosome 11q13 from D11S4155-6 to D11S4095-15a. Conclusions: We describe four patients with CPT1A and novel clinical and radiographic findings. We speculate that these new findings may be due in part to a toxic effect of long-chain fatty acids or their acyl-CoA esters in the brain, liver, pancreas and kidney. Two unrelated families from Durango, Mexico share a common haplotype on chromosome 11q13, suggesting their common CPT1A disease-causing variant Q100X to be a founder mutation.

713F

Basis of immune dysregulation in Gaucher disease. *C. Martin[1], L. Kozhaya[2], T. Taber[1], D. Unutmaz[2], O. Goker-Alpan[1]*. 1) LSD Research and Treatment Unit, Center for Clinical Trials, O & O Alpan, Fairfax, VA; 2) New York University School of Medicine, Department of Microbiology, New York, NY.

Gaucher disease (GD), the inherited deficiency of the lysosomal enzyme glucocerebrosidase, is the most common Lysosomal Storage Disorder. Activated lipid engorged macrophages are the hallmark for GD. Although the mechanisms for macrophage activation in GD is unclear, there is clinical evidence of immune dysregulation, and a non-specific inflammatory response, characterized with poor wound healing, frequent infections, and predisposition to cancers, are frequently encountered in patients with GD. The basis of immune dysregulation was explored in patients with Gaucher disease. Under an IRB approved protocol, 14 subjects (4M: 10F), (age range 34–56) were evaluated and immune profiling analysis were performed and results were compared to healthy controls. The proportion CD14+ monocytes, CD4+ lymphoid dendritic cells and CD1d-restricted NKT cells were greatly reduced in GD patients ($p < 0.05$). In addition, while the total NK cell numbers were reduced, the proportion of NK cells displayed more activated/differentiated phenotype. There was also a profound decrease in naïve T cells. Clinically, these findings correlated with the disease burden and response to enzyme replacement therapy. Elucidating the link between immunological dysregulation in Gaucher disease will not only provide insights into novel functions of this lysosomal enzyme, but also further our understanding of the pathways that establish the cross talk between immune cells.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

714T

Longitudinal Change in Brain Volumes and Cognitive Function in MPS IIIA. *I. Nestrasil[1], K. Delaney[1], B. Yund[1], A. Ahmed[1], K. Rudser[1], V. Kovac[1], P. Haslett[2], C. Richard[2], C. Whitley[1], E. Shapiro[1]*. 1) University of Minnesota, Minneapolis, MN; 2) Shire HGT.

Background: MPS IIIA (Sanfilippo syndrome type A) is primarily a neurological disease. Quantitative MRI and neurocognitive function may provide understanding of the disease and provide information about rate of decline. In an ongoing study of the natural history of this disease (NCT01047306), we previously reported that baseline cognition is significantly associated with age and grey matter volumes. Goals: To present longitudinal data in MPS IIIA patients and to estimate the rate of decline in both cognition and MRI over 12 months. Methods: Longitudinal cognitive and quantitative MRI data were collected from 24 children with documented Sanfilippo syndrome type A. Volumetric analysis by automated segmentation, and developmental quotient (DQ) and age equivalent scores by cognitive assessment were obtained. Results: For a subset of 19 patients with classic disease diagnosed before age six, all but one participant showed no development past a ceiling of 28 months. Younger patients continued to show slow development until 28 months; those over two showed no growth or decline in skills. For patients diagnosed after age six, patterns were variable. Decrements in developmental quotients were associated with the age of the child such that younger children showed steeper declines than older children. Gray matter volumes were similarly associated with age with larger decrements in volumes in younger children. When the two were associated a robust correlation was found between developmental quotient and total gray matter volume ($r = 0.86$. $p < .01$) Conclusions: Loss of grey matter volume is the primary MRI pattern linked to decline in cognitive function in MPS IIIA patients. We have demonstrated that both Developmental Quotient and Gray Matter Volume are markers of disease progression in MPS IIIA, and that they are closely associated. We have also shown that MPS IIIA patients usually do not exceed a ceiling of development at 28 months age equivalent.

715F

Is the Profile of Medical Signs and Symptoms Associated with Specific Genotype in Mucopolysaccharidosis Type I? *A. Ahmed, R. Cooksley, C. Whitley, E. Shapiro.* Department of Pediatrics, University of Minnesota, Minneapolis, MN.

Mucopolysaccharidosis Type I or MPS I is an autosomal recessive lysosomal disorder due to deficiency or absent of the enzyme alpha-L-iduronidase. Over 100 mutations have been identified on the IDUA gene located on chromosome 4p16.3. MPS I affects almost all organ systems of the body. Previous research has focused on the correspondence of mutation with severity of disease. In this study, our goal is to determine the association of mutation (nonsense, missense, splice site and deletions/insertions) with the profile of organ system involvement. Methods: Medical history data and genotype were collected on 26 severe and 14 attenuated MPS I participants as part of a NIH supported longitudinal study. Medical history included skeletal, cardiorespiratory, somatic, vision, hearing, neurological, and cognitive status. For severe patients, all but one had hematopoietic cell transplant; for attenuated patients, all were on enzyme replacement therapy. Data were analyzed to determine the probability of specific organ system involvement in particular mutation types. Results: In the severe group, 18 had two nonsense mutations and 6 had abnormalities of either splice site or a deletion; 2 had one nonsense and one missense mutation. In the attenuated group, 8 were nonsense/missense, and the remaining 6 had combinations of splice site/missense/deletions. All patients had skeletal and most had cardiac abnormalities. Most had corneal opacities but they varied in severity; multiple episodes of otitis media/with tube placement were common but not universal. Frequency of surgeries was higher in nonsense mutations. The probability of cognitive impairment was higher for the nonsense mutations. Probabilities of organ involvement for each type of mutation will be reported. We conclude that examining the probability of specific organ system involvement in particular mutation types can contribute important data for understanding the development of signs and symptoms in MPS I.

716T

Prevalence of Fabry disease in patients with chronic pain. *C. Goizet[1,2], V. Dousset[3], M. Cadenne[3], E. Maurat[2], S. François[2], C. Blanchard[2], P. Poulizac[4], S. Tabuteau[4], C. Hubert[5], S. Laborde[3], P. Casenave[6], X. Ferrer[2,6], G. Solé[2,6], I. Coupry[2], F. Bauduer[2,7], T. Levade[8], D. Lacombe[1,2]*. 1) Medical Genetics, CHU Pellegrin, Bordeaux, France; 2) Université Bordeaux Segalen, Laboratoire Maladies Rares: Génétique et Métabolisme (MRGM), EA4576, Bordeaux, France; 3) CHU Bordeaux, Hopital Pellegrin, Pôle des Neurosciences Cliniques, Centre Douleurs Chroniques, Bordeaux, France; 4) CHU de Bordeaux, Direction de la Recherche clinique et de l'Innovation, Talence, France; 5) Université Bordeaux Segalen, Plateforme Génome Transcriptome, Centre de Génomique Fonctionnelle de Bordeaux, 33076 Bordeaux, France; 6) CHU Bordeaux, Groupe Hospitalier Sud, Service de Neurologie, Pôle des Neurosciences Cliniques, Pessac, France; 7) Centre hospitalier de la Côte Basque, Service d'Hématologie, Bayonne, France; 8) CHU Toulouse, Laboratoire de Biochimie Métabolique, Toulouse, France.

Fabry disease (FD) is an X-linked recessive lysosomal disease caused by α-galactosidase A (α-gal) deficiency related to a deleterious mutation in the GLA gene, causing progressive glycosphingolipid storage in various organ systems. Chronic pain is classically considered as a frequent and early feature. To our knowledge, despite several studies of FD's prevalence in various populations with organ deficiencies that may occur in the course of the disease, there are no reports about the prevalence of FD in patients with chronic pain of undetermined causes. The objective of this study was to determine FD prevalence, as assessed by alpha-gal activity and genetic tests, in patients suffering from chronic pain of undetermined causes. We included 119 consecutive patients with such a diagnosis followed by the Centre of Chronic Pain in our University Hospital. Male patients underwent biochemical analyses to evaluate peripheral α-gal activity in leucocytes, although a search for GLA mutations was directly performed in female patients. Of 119 consecutive screened subjects (81 females, 38 males), all enzymatic assays were normal but one previously unknown GLA mutation p.Arg363His was found in a 62-year-old woman affected by asthenia since several years and presenting with a 2-years history of distal chronic pain in the limbs. Evaluation for additional features of FD in this patient is in progress. This is the first study reporting FD prevalence in a consecutive cohort of patients with chronic pain of undetermined causes. We found a relatively high prevalence among the included patients. These findings need to be confirmed in a larger cohort of patients in order to determine whether a systematic screening for FD is indicated in all patients with chronic pain, since it is a treatable genetic disorder.

717F

Infantile onset severe phenotype of cblG disorder in a patient and abnormal neurotransmitter metabolites. *R. Jobling[1,2], S. Hewson[1,2], S. Mercimek- Mahmutoglu[1,2]*. 1) The Hospital for Sick Children, Toronto, Canada; 2) University of Toronto, Toronto, Canada.

Cobalamin G disorder (*cbl*G) is an autosomal recessive inborn error of cobalamin metabolism caused by mutations in the *MTR* gene, encoding methionine synthase. It is characterized by developmental delay, seizures, hyperhomocysteinemia, hypomethioninemia and low intracellular methylcobalamin.

Case report: This is an 11-month-old girl, first child of consanguineous Somali parents, who presented with developmental regression and microcephaly at age 3 months. Her condition progressed to central hypotonia and epileptic encephalopathy at age 5 months. She had total a homocysteine of 100 μmol/L (reference: 0.5–11 μmol/L) and methionine of 9 μmol/L (reference: 15–21 μmol/L) with normal methylmalonic acid in urine. She was started on intramuscular hydroxycobalamin 1 mg/day, oral betaine 250 mg/kg/day and folic acid (5 mg/day). Total homocysteine decreased to 28 μmol/L within 6 days of therapy. The synthesis of methylcobalamin was low in cultured skin fibroblasts by complementation analysis, confirming cblG disorder. She was homozygous for a known intronic mutation (IVS-166 A>G) in the MTR gene. CSF neurotransmitter analysis showed low levels of 5 -methyltetrahydrofolate (MeTHF) (28 nmol/L; reference:40–240), homovanillic acid (245 nmol/L; reference:450–1132) and 5 -hydroxyindolacetic acid (109 nmol/L; reference:179–711).

Discussion: This patient presented with a severe phenotype of cblG disorder. The intronic mutation identified has been previously described as a null-type mutation in two siblings with severe neonatal onset in compound heterozygous state. To our knowledge, this is the first patient with cblG disorder in whom low levels of neurotransmitters and cerebral folate have been found.These abnormalities are of particular interest given the known neurological and psychiatric manifestations of this condition. The findings in this case are a potential source of new insight into the pathophysiology of *cbl*G.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

718T

Abnormal central tegmental tract signal in a patient with mild glycogen storage disease type 1a (GSD1a). *D.L. Renaud[1, 3], S.T. Janousek[2], D.K. Freese[3], W.F. Schwenk[3].* 1) Div Child/Adolescent Neurology, Mayo Clinic, Rochester, MN; 2) Noran Neurological Clinic Minneapolis, MN; 3) Department of Pediatrics, Mayo Clinic Rochester, MN.

Case Report A 15 month old boy, born to non-consanguineous parents, presented for evaluation of spells. His development and neurological examination were normal. He had hypoglycemia as a newborn. No subsequent episodes of hypoglycemia were documented. His EEG was normal. The MRI revealed abnormal signal in the central tegmental tracts bilaterally. Evaluation in the Neurometabolic Clinic revealed an elevated uric acid of 11.1 (2.4–5.4) and lactic acid at 7.8 (normal <3.2). Hyperlipidemia was present. The ALT was increased at 479. Neutropenia was absent. Urine organic acids revealed 3-methylglutaconic acid and 3-methylglutaric acid. An abdominal ultrasound demonstrated marked hepatomegaly with normal hepatic echogenicity. Kidneys were mildly enlarged. The biochemical and clinical profile were suggestive of GSD1a despite the limited history of hypoglycemia. Enzyme analysis of glucose-6-phosphatase activity in liver revealed a significantly decreased (0.9 micromole/min/gram tissue) but not absent enzyme activity (normal 3.50 +/− 0.8). Sequencing revealed a homozygous mutation, c.648G>T, which results in the creation of a new splice donor site in exon 5 and premature translation termination. Conclusions Central tegmental tract abnormalities have been described in patients with disorders of amino acid, organic acid and tetrahydrobiopterin metabolism. GSD1a should be added to the differential diagnosis of this MRI finding.

719F

Mild clinical and atypical somatic cell findings in a *cblF* patient detected by expanded newborn screening. *D. Rosenblatt[1], A. Brebner[1], I.R. Miousse[1], J.C. Kim[1], M.T. Geraghty[2], D. Watkins[1], C. Armour[3].* 1) Department of Human Genetics, McGill University, Montreal, Quebec; 2) Department of Pediatrics, Children's Hospital of Eastern Ontario and University of Ottawa, Ottawa, Ontario; 3) Medical Genetics Unit, Department of Pediatrics, Queens University, Kingston, Ontario.

A full term female infant came to medical attention following a positive newborn screen due to a borderline elevation of C3 and an increased ratio of C3:C2. The family history was unremarkable and she was of mixed European Caucasian descent. At the time of ascertainment, she was clinically well, with unremarkable physical exam. Follow-up testing at age 12 days demonstrated a continued elevation of C3 at 2.55 (reference value < 0.65), with increased excretion of methylmalonic acid (MMA), without methylcitric acid. The total serum homocysteine was elevated (39 μmol/L, reference range: 4–15μmol/L). The hemoglobin and MCV values were within reference limits. The serum cobalamin level was low (62 pmol/L, reference range 165–740 pmol/L). The urinary quantitative MMA level was 54mmol/mole creatinine (reference upper limit: 6.6mmol/mole creatinine). The infant responded first to cyanocobalamin and then to hydroxycobalamin (1mg every 2 weeks) and all values normalized. At 2.5 years of age she remains healthy and developing appropriately. Somatic cell studies using cultured fibroblasts studies showed that incorporation of [^{14}C]-propionate into cellular macromolecules was at the lower end of the reference range but too high to allow for complementation analysis. The incorporation [^{14}C]-methyltetrahydrofolate was well within the reference range. Fibroblasts from the patient showed decreased synthesis of both adenosylcobalamin and methylcobalamin and an excess of the labelled substrate, cyanocobalamin. Mutation analysis of *LMBRD1* revealed one copy of the common c.1056delG (p.L352fsX18) and one copy of a splice mutation c.1339–1G>T (p.T447fsX33) leading to a cDNA product lacking exon 14 (Miousse IR et al. Mol Genet Metab 102:505–507, 2011). Thus, expanded newborn screening has detected a *cblF* patient with a mild clinical phenotype and atypical cellular findings. Molecular analysis of the candidate gene allowed for the correct diagnosis and prompt therapy likely prevented any permanent neurological or physical sequelae.

720T

Fumarase deficiency. Expanding the phenotype. *K. Siriwardena[1], W. Al-Hertani[1], J. Kronick[1], D. Chitayat[1,2].* 1) The Hospital for sick Children, Department of Pediatrics, 1)Division of Clinical and Metabolic Genetics, University of Toronto, Ontario,Canada; 2) Mount Sinai Hospital, Department of Obstetrics and Gynecology, 2The Prenatal Diagnosis and Medical Genetics Program,; University of Toronto, Toronto, Ontario, Canada.

Fumarase is the enzyme that converts fumarate to malate in the Krebs cycle. The enzyme is encoded by the FH gene mapped to chromosome 1q42.1and deficiency of fumarase activity causes a rare autosomal recessive condition known as fumaric aciduria, or fumarase deficiency that is associated with severe neurological impairment and seizures. We report on a 24-month-old male, born to non-consanguineous parents with a non contributory family history who presented at 8 months of age with global developmental delay, failure to thrive, microcephaly, dystonia, chorea, intractable epilepsy, osteopenia and facial dysmorphism. Brain MRI showed delayed myelination with diffuse T2 hyperintensity in the supratentorial white matter with normal MRS. Urine organic acid analyses showed markedly increased excretion of fumaric acid, leading to the diagnosis of fumarase deficiency. DNA analysis revealed compound heterozygotic novel variants in the fumarate hydratase (FH) gene. Results of patient's enzyme analysis and parental DNA analysis are pending. There are less than 50 cases of Fumarase deficiency reported worldwide. This case report expands on the phenotype described for this disease. Inactivation of fumarate hydratase (FH) drives a metabolic shift to anaerobic glycolysis in the absence of hypoxia in FH-deficient kidney tumors and cell lines from patients with hereditary leiomyomatosis renal cell cancer (HLRCC), resulting in decreased levels of AMP-activated kinase (AMPK) and p53 tumor suppressor. Further studies are required to find if parents of patients with fumarase deficiency are at increased risk for developing hereditary leiomyomatosis renal cell cancer (HLRCC).

721F

Renal Growth in Isolated Methylmalonic Acidemia (MMA). *P. Kruszka, I. Manoli, J. Sloan, C.P. Venditti.* Genetics and Molecular Biology Branch, NHGRI, NIH, Bethesda, MD.

OBJECTIVES: A major complication of methylmalonic acidemia (MMA) is the development of tubulointerstitial nephritis and renal failure. The effect of MMA on kidney size has not been studied. This study examines kidney size variation with age and other measureable predictors and develops a kidney length nomogram. DESIGN: Renal ultrasounds, anthropometric measurements, and laboratory evaluations including serum creatinine (Cr), serum and urine methylmalonic acid, and cystatin C were conducted on 50 non-transplanted patients (ages 2.2 to 36.3 years) with enzymatic and mutation confirmed MMA. The patients were followed prospectively from 2004 to 2011, with some patients imaged sequentially, yielding 84 total ultrasounds. MMA patients were compared to a control group of 209 patients aged 0 to 19 years using the paired t-test. Linear and multiple regression analysis were used to study relationships between renal length and other clinical variables. RESULTS: Comparisons with age-matched controls showed a significant difference in renal length ($p < 0.05$) for most age groups. The regression equation for the control group was renal length (cm) = 6.79 + 0.22 * age (years) for children older than 1 year and for our study patients was 6.80 + 0.09 * age (years); $p < 0.001$ for both the constant and β_{age} coefficient ($R^2 = 0.399$). The most highly correlated single predictor model for MMA patients used height instead of age ($R^2 = 0.490$). When serum Cr values were in the normal range (Cr < 1.3), the most predictive multiple regression model found height, age, and serum methylmalonic acid (MMA_S) to be the optimal independent variables ($R^2 = 0.720$). Cystatin C was comparable to MMA_S in the multiple regression model and correlated closely with Cr and MMA_S (Pearson correlation coefficients = 0.7901 and 0.7684, respectively). Urine methylmalonic acid concentration was not a significant predictor of renal length when controlling for age, height and Cr ($p = 0.849$). CONCLUSIONS: Renal length in MMA patients is significantly decreased compared to normal controls and predicted by a multiple regression model that uses the clinical variables of height, age, and MMA_S when Cr is in normal range. In addition to generating a clinically useful nomogram, our results capture a marker of the natural history of MMA renal disease. The use of a simple index such as renal length will contribute to the study of therapeutic interventions designed to target the kidney in this population.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

722T

Plasmalogen deficiency in the Pex7 hypomorphic mouse, a model of Rhizomelic Chondrodysplasia Punctata, causes neonatal lethality associated with late lung maturational defects. *N. Braverman[1], X. He[1], S. Jiralerspong[1], G. Nimmo[1], Y. Chen[1], J. Hacia[2], A. Moser[3], W. Cui[1].* 1) Dept Human Gen & Pediatrics, McGill Univ, MCH Res Inst, Montreal, PQ, Canada; 2) Dept Biochemistry & Molecular Biology, Univ Southern California, Los Angeles, CA, USA; 3) Dept Neurogenetics, Kennedy Kreiger Institute, Baltimore, MD, USA.

The biological functions of plasmalogens (PLs), a specialized class of membrane glycerophospholipids, are largely unknown. Cell-based studies implicate roles in oxidant protection, inflammatory responses, vesicle formation and signal transduction. PLs are also integral components of surfactant. Inherited defects of PL synthesis cause rhizomelic chondrodysplasia punctata (RCDP), a heterogeneous disorder commonly due to defects in the peroxisome transporter, PEX7. The clinical features of RCDP include growth and developmental retardation, cataracts and skeletal dysplasia. Patients suffer chronic respiratory compromise that ultimately causes death. To investigate PL functions, we generated hypomorphic Pex7 mouse models that survived on a mixed 129/C57Bl6 background, but died within a few days of life on a C57Bl6 background. Morphometric studies showed that death was associated with increased saccular wall tissue and decreased air spaces, indicating a delay, or arrest, in alveolarization. In addition, we found increased numbers of terminal bronchiole epithelial cells (Clara cells) and alveolar type II epithelial cells, as well as increased lipid peroxidation protein adducts. We suggest that (1) PL deficiency may affect cell signaling in late lung maturation and independently, also promote oxidant damage, (2) tissue effects of PL deficiency may be modified by strain background and, (3) PL deficiency contributes to the chronic lung disease observed in RCDP patients.

723F

Phenylketonuria and defective metabolism of tetrahydrobiopterin. *A. Al-Aqeel.* Department of Pediatrics, Riyadh Military Hospital, Stem Cell Therapy program, King Faisal Specialist Hospital and Research Centre, Riyadh, Saudi Arabia.

Variant forms of hyperphenylalaninemia result from abnormalities of tetrahydrobipterin (BH4) cofactor synthesis. BH4 is an essential cofactor not only for phenylalanine hydroxylase, but also for tyrosine and two tryptophan hydroxylases, three nitric oxide synthases, and glyceryl-ether monooxygenase. Defective activity of tyrosine and tryptophan hydroxylases explains the neurological deterioration in patients with BH4 deficiency with progressive mental and physical retardation, central hypotonia and peripheral spasticity, seizures and microcephaly Five separate genetic conditions affect BH4 synthesis or regeneration: deficiency of GTP cyclohydrolase I, 6-pyruvoyl tetrahydropterin synthase (6-PTS), sepiapterin reductase, dihydropteridine reductase (DHPR) and pterin-4alpha-carbinolamine dehydratase. Only the latter of these conditions is relatively benign and is associated with transient hyperphenylalaninemia, however the clinical manifestations of the other disorders are quite similar. All these conditions can be identified in newborns by an elevated phenylalanine, with the exception of sepiapterin reductase and the dominant form of GTP cyclohydrolase I deficiency that results in biopterin deficiency/insufficiency only in the brain. Diagnosis relies on the measurement of pterin metabolites in urine, dihydropteridine reductase in blood spots, neurotransmitters and pterins in the CSF and on the demonstration of reduced enzyme activity (red blood cells or fibroblasts) or causative mutations in the relative genes. In Saudi Arabia newborn screening metabolic disorders was started in 2005 with overall incidence of 1:900, with an incidence of PKU cases of 1:4000. Deficiency of 6-PTS is the most common of the defects in biopterin metabolism, with almost 50% of all cases of PKU detected on newborn screening 6-PTS gene is an 8Kb span with 6 exons, with over 40 mutations reported. Four unique mutations are reported in Saudi Arabia, one of which with 23 base pairs deletions. In conclusion 6-PTS deficiency PKU is common in Saudi Arabia with novel mutations.

724T

Two Cases of Neonatal Onset Type II Citrullinemia Diagnosed by Urine Organic Acids Based Newborn Screening. *Y. Watanabe[1, 2], K. Tashiro[1], K. Aoki[1], T. Inokuchi[1], Y. Seki[2], T. Yanagi[2], T. Mizuochi[2], J. Okada[2], M. Yoshino[2], T. Matsuishi[1, 2].* 1) Research Institute of GC/MS, Kurume University, Kurume, Japan; 2) Dept. Pediatrics, Kurume University, Kurume, Japan.

Background: Citrullinemia type II caused by citrine deficiency, can manifest in newborns as neonatal intrahepatic cholestasis (neonatal intrahepatic cholestasis caused by citrin deficiency: NICCD). Citrin is encoded by SLC24A13 and is a mitochondrial membrane protein consisting of the aspartate-glutamate carrier. Citrin deficiency is an autosomal recessive disorder with the higher incidence in far eastern Asian countries. Citrin deficiency is recognized as one of the common genetic causes of neonatal intrahepatic cholestasis as are galactosemia and tyrosinemia type1. Symptoms associated with NICCD including intrahepatic cholestasis, liver dysfunction due to hepatic fibrosis and fatty infiltrations generally resolve within one year with appropriate treatment. However, some infants require liver transplantation. Majority of patients with citrin deficiency present with high blood citrullin levels in newborn screening (NBS). A pilot study of urinary organic acid based NBS has been conducted since 1996 in the southwest area (Kyushu Island) of Japan. Two cases of NICCD were missed by blood amino acid based NBS by tandem mass spectrometry, but they were identified by urine organic acid based NBS. Patients and Methods: Case 1 was a 2784g, term male infant. Blood and urine specimens were collected on age 4 days for NBS. A high urine 4-hydroxyphenyllactic acid (4OH-PLA) was noted. Blood based NBS was unremarkable. Case 2 was an 1808g, 33 weeks, preterm male infant born by caesarian section. High 4OH-PLA and 4OH-phenylpyruvic acid (4OH-PPA) levels were noted in urine specimens collected on age 18 days. A high blood galactose was noted on age 21 days although blood based NBS on age 5 days was unremarkable. Conclusions: Because of the high incidence, particularly in the far eastern Asian countries, citrin deficiency is an important genetic condition for NBS. Biochemical analytes including blood ammonia, serum amino acids including citrulline, arginine, methionine, phenylalanine, tyrosine, and threonine, as well as serum pancreatic secretory trypsin inhibitor, serum alfa-feto-protein, and galactose have been reported to be elevated in patients with citrin deficiency. Although high levels of urine 4OH-PLA and 4OH-PPA are known to be non-specific findings and often seen in patients with hepatic dysfunction of many different etiologies, they might serve as useful early biomarkers for citrin deficiency in the newborn period.

725F

Translational research in newborn screening: implementation of a consensus data set for the long-term follow-up of patients to improve scientific understanding and optimize health outcomes. *S. Berry[1], K. Hassell[2], A. Brower[3], Joint Committee of the NBSTRN Clinical Centers Workgroup and NCC/RC Long-Term Follow-Up Workgroup.* 1) Dept Pediatrics, Univ Minnesota, Minneapolis, MN; 2) University of Colorado, Aurora, CO, 80045, USA; 3) American College of Medical Genetics, Bethesda, MD, 20814.

Background: The goal of translational research in newborn screening is to accelerate the implementation of new discoveries into public health programs and routine medical practice. Innovations in analytical technologies and the development of novel therapies and management approaches trigger validation activities across the spectrum of the screening, diagnosis, treatment and long-term follow-up. While analytical validation takes place within the state-based newborn screening programs and quality control programs are in place, clinical validation occurs in the course patient care in specialty centers and there is no mechanism to collect, aggregate and analyze information. We report on an innovative effort to implement a system to capture health information across the lifespan for newborn screen identified patients. Methods: A consensus data set developed through the efforts of the Joint Committee, described above, and accompanying informatics tools developed by the NBSTRN, were implemented within the context of two research projects. The aims of the first research project was to further understanding of the natural history of metabolic conditions included in routine newborn screening efforts, while the objective of the second study was to develop and pilot a novel newborn screening technology and therapeutic intervention for a condition that is not included in newborn screening. The structure and goals of the two projects were different but both utilized the same consensus set of data elements to understand the natural history and impact of newborn screening on diverse conditions. In addition, each project implemented disease specific data elements. Results: A consensus common data set of approximately 170 elements for use across all disorders, and individual disease specific data sets with a range between 300 and 600 individual elements were implemented in two ongoing research studies. Data capture, aggregation and analysis occurred within the context of routine clinical care and included data derived from public health screening. Conclusions: The implementation of a consensus data set for the long-term follow-up of newborns facilitates the translation of new discoveries into clinical practice to improve health outcomes and further scientific understanding. The use of this approach in ongoing research studies serves a model for future initiatives as genomic technologies begin to be applied to disease screening and management across the lifespan.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

726T

Newborn screening for infantile Pompe disease: Report of a pilot study in National Center for Child Health and Development. *M. Kosuga, K. Kida, N. Fuji, T. Okuyama.* Department of Clinical Laboratory Medicine, National Center for Child Health and Development, Tokyo, Japan.

Newborn screening (NBS) for Pompe disease has been initiated in several countries and regions, and is reportedly successful. However, the comparatively high frequency of pseudodeficiency allele makes NBS for Pompe disease complicated in Japanese population. We have demonstrated our modified screening procedure was able to distinguish pseudodeficiency individuals from Pompe disease patients. Based on this research, we have started a pilot study of NBS for infantile Pompe disease in our hospital and reported the result of this study. Healthy newborns born from Jan. to May in 2011 in National Center for Child Health and Development were screened. Informed consent was obtained from all newborn's family members. Dried blood spot (DBS) samples were obtained from newborns in 3–5 days after deliver. Samples with 30% of the normal GAA activity mean, % inhibition 60%, and NAG/GAA ratio 30 were considered as 1st screening positive.361(Male: 185, Female: 176) neonates were screened in this pilot study. 15 neonates were found with less than 30% of normal GAA activity level. 14 of 15 newborns were diagnosed as pseudodeficiency with the result of gene analysis. The frequency of pseudodeficiency was 3.9%. One newborn was underwent the measurement of GAA activity in lymphocytes and GAA gene analysis. This newborn was diagnosed as normal healthy individual in the end. Specificity of the pilot study was 95.8%.

727F

The identification of infants with Fabry, Pompe, and Mucopolysaccharidosis-I from newborn blood spots. *C.R. Scott[1], S. Elliott[1,2], N. Buroker[1], L.I. Thomas[1], M. Glass[2], F. Turecek[3], M.H. Gelb[3].* 1) Dept Pediatrics, 356320, Univ Washington Sch Med, Seattle, WA; 2) Washington State Dept of Health Newborn Screening Laboratory, Shoreline, WA; 3) Dept Chemistry, Univ Washington, Seattle, WA.

Can infants with lysosomal storage diseases (LSDs) be identified from newborn blood spots? With an appropriate buffer, substrates and internal standards, quantitation of the enzymes responsible for Fabry, Pompe, and MPS-I can be simultaneously assayed and quantitated by MS/MS. Here we report the results from screening 100,000 samples obtained from newborns. A stipulation of IRB was that screening be performed on anonymized samples. From the 100,000 samples, we identified infants with low enzyme activity, and confirmation was performed by mutation analysis of DNA isolated from a 3.2-mm blood spot.

For Fabry, a total of 16 samples were identified with low enzyme activity. Eight males were confirmed to have mutations consistent with Fabry. Thus, the prevalence for Fabry is ~1/6,800 males. The method has a positive predictive value (PPV) of 0.5 and a false positive rate (FPR) of ~1/14,000.

For Pompe, 17 infants were identified with low activity. Four were confirmed to have nucleotide changes consistent with clinical Pompe. All were of "late-onset" phenotype. Of the remaining 13 samples, four were heterozygote/wt, three were heterozygote/pseudo def, six were pseudo def/wt. The prevalence for Pompe was ~1/28,000, with a PPV = 0.24 and FPR = ~1/8,600.

Nine samples were identified with low activity for α-iduronidase. Three had mutations consistent with the diagnosis of MPS-I. One was homozygous for a null-mutation (Hurler), two were consistent with an "attenuated" phenotype. Of the six unaffected samples, one was a heterozygote/wt, two were poor punches, and three were wt/wt. The prevalence for MPS-I was ~1/35,500, with a PPV = 0.33 and FPR = ~1/15,000

The prevalences of these LSDs is 2–6x greater than reports from clinical studies. The MS/MS assay is effective for measuring multiple enzymes with a low false-positive rate. The assay can be performed by any screening laboratory that is interested in expanding their repertoire for the detection of genetic disease.

Support: NICHD #HHSN2672006 and Genzyme Corp.

728T

Bare lymphocyte syndrome type III, a novel mutation in *RFXAP* and failure to identify with newborn screening for severe combined immunodeficiency. *L. Randolph[1], J. Chase[2], M. Aguilera[3], J. Siles[4], J. Church[5].* 1) Division of Medical Genetics, Children's Hospital Los Angeles and Keck School of Medicine, University of Southern California, Los Angeles, CA; 2) Division of General Pediatrics, Children's Hospital Los Angeles and Keck School of Medicine, University of Southern California; 3) Molecular Biology Center GENETAQ, Málaga, Spain; 4) Molecular Biology Center GENETAQ; 5) Division of Clinical Immunology and Allergy, Children's Hospital Los Angeles and Keck School of Medicine, University of Southern California.

Newborn screening (NBS) for severe combined immunodeficiency (SCID) is performed by measuring T-cell receptor excision circles (TREC), episomal products of T-cell maturation in the thymus, from dried blood spots obtained at birth. However, there are genetic forms of SCID that theoretically would not be diagnosed by this method. Our patient was 14 months of age at the time immune deficiency was suspected. The son of second cousins, he had FTT, global DD and multiple atypical respiratory infections (e.g. rhinovirus pneumonia) and B. pertussis pneumonia, despite having been properly immunized. Serum IgG, IgA, and IgM were undetectable. Flow cytometry revealed T- and B-lymphocytosis but a reversed CD4/CD8 ratio. His T-cells responded in vitro to non-specific mitogens but not to tetanus or candida antigens. Retrospective review of his NBS TREC result was normal. Single nucleotide polymorphism (SNP) chromosomal microarray was initially performed because of DD. It showed a deletion on 5p15.2 including the catenin delta 2 (*CTNND2*) gene, which predicts moderate to severe mental retardation. Loss of heterozygosity was also noted in several areas, including a region encoding the regulatory factor X-associated protein (*RFXAP*) gene in 13q13.3. Mutations in this gene cause bare lymphocyte syndrome (BLS) types II and III. This form of SCID results from the lack of expression of MHC II (BLS II) or MHC I and II (BLS III). Further investigation of our patient demonstrated that his lymphocytes lack expression of HLA A, B and C (MHC I) and HLA DP, DQ and DR, resulting in a diagnosis of BLS III SCID. *RFXAP* was sequenced, and a novel homozygous mutation was identified: c.323T>A in exon 1, a nonsense variant predicted to lead to a premature stop codon. **Summary**: Lack of MHC expression profoundly impairs antigen-specific T- and B-cell functions. Thymus production of mature T-cells is normal in BLS III; this results in a form of SCID that will not be identified by current NBS testing. SNP arrays can be useful in identifying candidate genes in complex disorders such as these.

729F

Standard value of urine HVA/VMA ratio in the early neonatal period to screen for Menkes disease. *M. Yagi, N. Kusunoki, T. Lee, H. Awano, T. Yokota, A. Miwa, A. Shibata, I. Morioka, Y. Takeshima, K. Iijima.* Dept Pediatrics, Kobe Univ Graduate Sch Med, Kobe, Hyogo, Japan.

[Background] Menkes disease (MD) is a lethal, X-linked recessive disorder of copper metabolism dominated by neurodegenerative symptoms and connective tissue disturbance. Early treatment with parenteral copper-histidine has been shown to improve clinical outcomes. However, early detection is difficult because clinical abnormalities are subtle in early neonatal period. Previously, it was reported that decreased activity of dopamine β-hydroxylase, a copper-dependent enzyme, leads to increased urine homovanillic acid/vanillylmandelic acid (HVA/VMA) ratio in MD patients, and the urine HVA/VMA ratio (cut-off value:4.0) is a useful screening method for MD [Matsuo et al. *J Inherit Metab Dis.* 2005]. However, there is no data of urine HVA/VMA ratio in neonates. **[Purpose]** To examine the standard value of urine HVA/VMA ratio in early neonatal period. **[Methods]** Concentration of urine HVA and VMA from 54 neonates with a gestational age from 32 to 41 weeks at 1 or 4 days of age were measured by high-performance liquid chromatography. **[Results & Discussion]** Among all samples, HVA concentrations ranged from 5.0 to 245.1μg/mg creatinine (mean ± SD=26.4 ± 43.0), VMA concentrations ranged from 3.9 to 23.7μg/mg creatinine (mean ± SD= 7.34 ± 2.93), and HVA/VMA ratios ranged from 0.725 to 29.1(mean ± SD= 3.35 ± 4.28). Three samples had extremely high HVA/VMA ratios; 9.21, 17.7 and 29.1. All of three were collected from the cases treated with catecholamine on their collection date. The treatment with catecholamine influences HVA and VMA measurements. Among 51 samples except the three with catecholamine treatment, HVA/VMA ratios ranged from 0.725 to 4.66 (mean ± SD=2.45 ± 0.75). These values were similar to previously reported urine HVA/VMA ratios at 6 months of age which was used as a control data to determine the cut-off value. These data suggest that the urine HVA/VMA ratio can be an effective screening method for MD in early neonatal period as well.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

730T

Correlation of phenylalanine concentration with neuropsychiatric symptom scores in phenylketonuria patients. *T. Morgan, N. Owen, A. Bawcom, G. Wey, L. Ashford, J. Phillips.* Pediatrics, Medical Genetics, Vanderbilt Univ Sch Med, Nashville, TN.

Background. In Phenylketonuria (PKU), the central nervous system may be adversely affected by accumulation of the toxic precursor (Phe) as well as a deficit of the product (Tyr), which is a precursor for Dopa and Dopamine. Lifetime Phe:Tyr ratio (PTR) correlates with executive cognitive function in PKU patients, but less is known about the relationship between PTR and neuropsychiatric symptoms such as anxiety, depression, and attention deficit-hyperactivity disorder (ADHD). Thus, we hypothesized that amino acid concentrations, which are more convenient than lifetime measures, may reflect current neuropsychiatric status. To test this hypothesis we determined the correlation between concomitant neuropsychiatric symptom scores and Phe, Tyr, and PTR in PKU patients. Methods. Subjects were enrolled in an observational, single-site, pilot study nested within a multi-center, double-blind, placebo-controlled, randomized trial (PKU-016) to evaluate the safety and efficacy of sapropterin hydrochloride (Kuvan®) for neuropsychiatric symptoms in patients with PKU. Subjects had PKU with stable diets and medication intakes. Baseline and serial fasting amino acid levels, dietary histories, and neuopsychiatric testing using HAM-A, HAM-D, and ADHD standardized testing scales were obtained. Symptom scores were summed across all tests, and correlations with PTR, Phe and Tyr were tested for two-tailed statistical significance by the Spearman rank-sum correlation coefficient. Results. Eleven pediatric and adult subjects were enrolled at the Vanderbilt study site. Mean sum of scores (HAM-A, HAM-D, ADHD RS/ASRS) was 35.1 ± 16.5 (standard deviation). There was a correlation ($r = 0.66$; $P = 0.026$) between spot PTR and total sum of scores. Phe concentration had a similar correlation ($r = 0.67$; $P = 0.02$), but Tyr concentration was not correlated ($r = 0.22$; $P = 0.52$). Discussion. Despite the modest sample size of our pilot study, we found a strong, statistically significant correlation between neuropsychiatric symptoms in stable PKU patients and spot PTR and Phe but not Tyr concentrations. A larger, multi-site study is needed to determine optimal, clinically relevant, PTR and Phe cutoffs at which neuropsychiatric testing is warranted. In addition, the extent to which neuropsychiatric scores track with PTR and Phe over time, as well as the efficacy of sapropterin hydrochloride for optimizing PTR, Phe, and neuropsychiatric scores, should be determined.

731F

A rare disease - simple treatment: Sepiapterin Reductase Deficiency with a novel mutation. *I.M. Zamir, B. Ben-Zeev, A. Mimouni-Bloch, S. Msilaty-Gross, Y. Anikster.* The Edmond and Lily Safra Children's Hospital and Sackler Faculty of Medicine, Tel Aviv University, Tel Hashomer, Israel.

Sepiapterin Reductase catalyzes the last step in the biosynthesis of Tetrahydrobiopterin (BH_4), which is important in the metabolism of phenylalanine, catecholamines and serotonin. SPR deficiency patients were described as having a Dopa-responsive dystonia, progressive psychomotor retardation, hypersomnia and microcephaly, with the classic triad of oculogyric crises, paroxysmal stiffening and hypotonia and an autosomal recessive inheritance pattern. No Hyperphenylalaninemia is noted, which makes this condition difficult to recognize with regular newborn screening. The rational for therapy is substituting with L-Dopa and 5-Hydroxytryptophan. We report a 20-month-old Russian origin girl that presented with marked global developmental delay and oculogyric crises. She is the first born of unrelated healthy parents. Her weight was 8kg at age 17 months. Upon examination the patient exhibited head lag with opisthotonus and oculogyri crises during crying, weakness of neck and back with spastic quadriparesis, no eye contact was maintained and cogwheel rigidity in all limbs was present with positive Babinski's sign. Imaging was normal. Extensive workup revealed increased Sepiapterin and Biopterin levels, low levels of Dihydropteridine Reductase, with no Hyperphenylalaninemia. CSF levels of 5-Hydroxyindolacetic acid (5-HIAA), Homovanillic acid (HVA) and L-Dopa were all markedly reduced. Sepiapterin Reductase Deficiency was diagnosed. The patient was started on LevoDopa titrated up to 40 mg/d, 5-Hydroxytryptophan 8 mg/kg/week and Folinic Acid 15 mg/d. Great improvement was noted with higher levels of arousal and absence of opisthotonus and oculogyri crises, markedly improved tonus, reflexes and speech with movement in all limbs. Subsequent *SPR* gene sequencing revealed that the patient is a compound heterozygote for 2 mutations: 1) c.448A>G/N (p.R150G) in exon 2 - was carried by the mother and was previously described in Dopa-responsive SPR deficiency patients; 2) A novel mutation - c.26–36dup/N (p.T13CfsX4) in exon 1 - was carried by the father. We describe a patient with the classic manifestations of SPR deficiency, with a novel mutation in the *SPR* gene. It is of utmost importance to diagnose this potentially treatable condition early as there is a good chance of normal development when promptly treated.

732T

A Phase 2a Study to Investigate the Effects of a Single Dose of Migalastat HCl, a Pharmacological Chaperone, on alpha-Gal A Activity in Subjects with Fabry Disease. *F.K. Johnson[1], P.N. Mudd Jr.[2], S. Sitaraman[3], R. Winkler[3], J. Flanagan[4], R. Khanna[4], K.J. Valenzano[4], D.J. Lockhart[4], P. Boudes[3] On Behalf of Study AT1001–013 Principal Investigators.* 1) Clinical Pharmacology and Pharmacokinetics, Amicus Therapeutics, Cranbury, NJ; 2) Clinical Pharmacology, Glaxo Smith-Kline, Research Triangle Park, NC; 3) Clinical Research, Amicus Therapeutics, Cranbury, NJ; 4) In Vivo Pharmacology, Amicus Therapeutics, Cranbury, NJ.

Fabry disease is an X-linked glycosphingolipid metabolism disorder caused by mutations in the GLA gene which results in a deficiency in lysosomal enzyme alpha-galactosidase A (alpha-Gal A). Migalastat HCl (AT1001, GR181413A) is a pharmacological chaperone that binds and stabilizes mutant and wild-type forms of alpha-Gal A. In preclinical studies, co-administration of migalastat HCl with gene-activated or recombinant human alpha-Gal A (generically known as enzyme replacement therapies (ERTs), agalsidase alfa and beta, respectively) to rodents resulted in greater increases in skin, heart, kidney, and plasma alpha-Gal A activity than seen with agalsidase alone. Co-administration with migalastat HCl also resulted in greater reductions in tissue globotriaosylceramide (GL 3) in a Fabry mouse model. Study AT1001–013 (NCT # 01196871) is a Phase 2a study to evaluate the safety and pharmacokinetic (PK) effects of oral administration of migalastat HCl two hours prior to initiation of intravenous infusion of agalsidase alfa or beta in male patients with Fabry disease. The primary objective of the study is to characterize the effects of a single administration of 150 mg or 450 mg migalastat HCl on alpha-Gal A activity in plasma, skin, and peripheral blood mononuclear cells (PBMCs). This is an open-label, single-dose, fixed-sequence study in two sequential stages. Stage 1 is comprised of 3 periods: algalsidase alfa or beta ERT alone, 150 mg migalastat HCl followed by initiation of ERT, then 150 mg migalastat HCl alone. Stage 2 is comprised of 2 periods: ERT alone, then 450 mg migalastat HCl followed by initiation of ERT. A minimum 14-day interval separates each treatment period. Results for alpha-Gal A activity and total alpha-Gal A protein in plasma, skin, and PBMCs, and plasma migalastat PK will be presented at the meeting.

733F

A novel drug target site for medium chain acyl-CoA dehydrogenase deficiency: Implications for future drug therapy. *H. Kang[1], J. Vockley[1,2], A-W. Mohsen[2].* 1) Department of Human Genetics, University of Pittsburgh, Pittsburgh, PA; 2) Department of Pediatrics, University of Pittsburgh, Pittsburgh, PA.

Medium chain acyl-CoA dehydrogenase deficiency (MCADD) is one of the most common biochemical genetic disorders in the US. Patients are asymptomatic at birth but are at risk for episodes of life threatening metabolic decompensation. Episodes first occur between 3 and 24 months of age in association with physiological stress. Mortality rate during an acute crisis in undiagnosed patients can be as high as 20%. With newborn screening, MCADD can now be identified pre-symptomatically, nearly eliminating mortality. However, treatment requires lifelong dietary monitoring, and significant morbidity still occurs due to hospitalizations for IV glucose therapy. A medication capable of relieving the metabolic block would be of great benefit. A single point mutation in the ACADM gene (A985G) substituting K304 with a glutamate, has been identified in 90% of the alleles in deficient patients. This amino acid change in turn leads to impaired folding and/or stability of MCAD. Substrate binding to MCAD K304E mutant has previously been shown to improve thermal stability of the protein. To improve the MCAD K304E mutant stability and activity, we explored the possibility that other ligands can confer similar protective effect. The electron transfer flavoprotein (ETF), the second substrate/electron acceptor, binds to the MCAD:CoA ester ternary complex to complete the α,β-dehydrogenation reaction with concomitant release of the enoyl-CoA product. X-ray structure of the MCAD:ETF complex published data was used to examine ETF docking site and its pharmacophore characteristics was confirmed. Different 12-mer peptides were synthesized based on the structure of the ETF docking peptide. To test whether these peptides can influence enzyme activity, the purified recombinant MCAD K304E mutant protein was coincubated with these peptides and enzyme assays were performed at different temperatures. The presence of wild type peptide, ETF βR191-βK202, has shown improved MCAD mutant activity at higher temperatures. A derivative of the WT peptide, DP193, showed better improvement (~2.5°C) of the MCAD mutant activity with increasing temperatures, compared to WT peptide. CD spectroscopy protein melting assays confirmed a shift of the Tm of 4°C with coincubation of DP193 peptide and MCAD mutant protein. This suggests that ETF docking site can be targeted for MCADD drug design using a structural based approach and that peptide 193 can be used as a scaffold for fragment based drug design approach.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

734T

Immune Response against Enzyme Replacement Therapy (ERT) Alters Enzyme Distribution in Mucopolysaccharidosis I Mice. *S. Le, S-H. Kan, M. Vera, P. Dicskon.* Dept Med Gen, LA Biomed Harbor-UCLA, Torrance, CA.

Mucopolysaccharidosis type I (MPS I) is an inherited and progressive lysosomal storage disease. Enzyme replacement therapy (ERT) with recombinant human alpha-L-iduronidase (rhIDU) is available, but studies in MPS I dogs suggest that anti-iduronidase antibodies may reduce the effectiveness of treatment.

This study aimed to characterize the tissue and cellular distribution of rhIDU in naïve and sensitized MPS I (*Idua*$^{-/-}$) mice. *Idua*$^{-/-}$ mice (n=13)were sensitized to rhIDU by weekly exposure to the antigen via tail-vein injection from 4 to 15 weeks of age. At week 16, a single dose of rhIDU labeled with the near-infrared fluorophore AlexaFluor 680 (AF680-IDU) was administered to the sensitized mice and 6 MPS I mice that were naïve to rhIDU. *In vivo* and *ex vivo* AF680-IDU distribution was quantified using a CRi Maestro EX fluorescent small animal imager.

Three of the sensitized mice died prior to the completion of the experiment, apparently from infusion reactions. Of the ten surviving sensitized mice, five had anti-iduronidase IgG antibody titers ranging from 3 to 30 OD units/ml and were designated “high-titer,” while five had titers <1 OD units/ml and were designated “low-titer”. Incubation of MPS I fibroblasts with serum from 4 high-titer mice let to an approximate 40% reduction in rhIDU uptake compared with antibody-free serum.

We observed a difference in the distribution of AF680-IDU to all organs in high-titer vs. naïve MPS I mice. We evaluated distribution in organs with high reticuloendothelial (RE) content (liver, spleen, thymus) and organs with low RE content (lung, kidney, heart, brain). We found that the sensitized, high-titer mice had an average fluorescence intensity that was ~33–45% lower in high RE organs and ~33–70% lower in low RE organs compared to naïve mice. This difference reached statistical significance for the heart and liver (p=0.015). Studies of the effect of the immune response against rhIDU on cellular and sub-cellular distribution are underway. Experiments in MPS I mice confirm findings in dogs that anti-iduronidase antibodies may alter the distribution of rhIDU.

735F

Hepatic gene transfer of TFEB results in clearance of mutant alpha-1-antitrypsin. *N. Pastore[1], P. Piccolo[1], F. Annunziata[1], R.M. Sepe[1], F. Vetrini[2], D. Palmer[2], P. Ng[2], K. Blomenkamp[3], J. Teckman[3], A. Ballabio[1,2,4,5], N. Brunetti-Pierri[1,5].* 1) Telethon Institute of Genetics and Medicine, Naples, Italy; 2) Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA; 3) Department of Pediatrics, Washington University School of Medicine, St. Louis Children's Hospital, St. Louis, MO, USA; 4) Jan and Dan Duncan Neurological Research Institute, Texas Children's Hospital, Houston, TX, USA; 5) Department of Pediatrics, Federico II University, Naples, Italy.

Alpha-1-antitrypsin deficiency caused by missense mutation (lysine for glutamate at amino acid position 342) altering protein folding is the most common genetic cause of liver disease in children. The only curative treatment currently available is liver transplantation. Transcription factor EB (TFEB) is a master regulatory gene affecting the number and function of lysosomes and autophagy. In the present study, we have investigated efficiency of hepatic gene transfer of TFEB at increasing clearance of mutant hepatotoxic alpha-1-antitrypsin (ATZ). We injected a helper-dependent adenoviral (HDAd) vector expressing the TFEB under the control of a liver-specific promoter (HDAd-TFEB) in the PiZ mouse model, a transgenic mouse expressing the human ATZ gene and recapitulating the features of liver disease observed in humans. Three-month old PiZ mice were injected intravenously with HDAd-TFEB at the dose of 1x10e13 vp/kg or, as controls, with the same dose of a HDAd vector expressing the unrelated alpha-fetoprotein (AFP) gene under the control of the same expression cassette (HDAd-AFP) or saline. Compared to saline or HDAd-AFP, mice injected with HDAd-TFEB showed a dramatic reduction in hepatic ATZ accumulation, as demonstrated by marked reduction of periodic acid-Schiff (PAS) staining and ATZ-containing globules. Moreover, a marked, statistically significant decrease in both ATZ monomer and polymer was observed in HDAd-TFEB injected mouse livers as compared to either saline or HDAd-AFP injected control mice, thus indicating that TFEB hepatic expression enhances disposal of both insoluble and soluble hepatic ATZ. TFEB gene transfer resulted in an increase of hepatic LC3, a marker of autophagic activity, and increased ATZ degradation by autophagolysosomes. Taken together, these results demonstrate that hepatic gene transfer of TFEB reduces accumulation of ATZ by enhancement of autophagy. In addition, we showed a reduction of hepatocyte apoptosis and hepatic fibrosis, which are key features of the hepatic disease of AAT deficiency, in HDAd-TFEB injected mice. In summary, we showed that TFEB-mediated hepatocyte expression results in clearance of ATZ, improvement of the liver phenotype and therefore, is an attractive gene-based strategy for the treatment of alpha-1-antitrypsin deficiency hepatic disease.

736T

Neuronal differentiation of induced pluripotent stem cells from Gaucher disease fibroblasts. *J.B. Florer[1], C.N. Mayhew[2,3], Y. Sun[1,3], G.A. Grabowski[1,3].* 1) Division of Human Genetics, Cincinnati Children's Hospital Research Foundation, Cincinnati, OH; 2) Division of Developmental Biology, Cincinnati Children's Hospital research Foundation, Cincinnati, OH; 3) Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati OH.

Gaucher disease (GD) is caused by inherited deficiencies of acid β-glucosidase (GCase). Two major GD variants have been delineated: Type 1 is the nonneuropathic, and types 2 and 3 are the acute and subacute neuronopathic variants. Parkinson's disease and synucleinopathies have been linked to GD. To understand the pathogenesis of the neuronopathic involvement in GD and to generate patient specific pluripotent stem cells for potential treatment, induced pluripotent stem cells (iPSC) were derived from fibroblasts of a GD patient carrying mutation L444P/P425R and an unaffected individual. The iPSCs were differentiated to neural cell types using a neural aggregate method. Populations of the induced neural cells expressed nestin and sox2, indicating generation of neural progenitor cells (NPC). The NPCs were grown on polyornithine/laminin coated plates in defined media with bFGF and were maintained for more than 14 passages. The differentiation of neurons from NPCs was carried out by replacing the bFGF with FGF8 and SHH. The derived cells exhibited neuronal morphology with extended neurites and expressed neuron markers Map2 and Tuj1. A subset of cells stained positive for tyrosine hydroxylase indicating generation of dopaminergic neurons. GCase activity was deficient in fibroblasts, iPSCs and NPCs derived from the GD patient. The glucosylceramide species profiles were specific to each cell type in GD and the control. There were increased amounts of glucosylceramide in the GD fibroblasts, and increased glucosylsphingosine in the GD fibroblasts and GD derived NPC and neurons. Similar cell culture systems are being developed using our GD mouse models. The results demonstrated that iPSC-derived NPCs, neurons, and striatal neurons from various GD human and mouse fibroblasts were viable, which indicates the cell non-autonomous nature of neuronal death in the CNS variants of Gaucher disease.

737F

Resveratrol and bezafibrate both dramatically induce transcription of aspartoacylase - a potential new treatment for Canavan Disease? *B.S. Andresen[1], T.K. Doktor[1], J.O. Sass[2,3], J. Bastin[4], F. Djouadi[4], T.J. Corydon[5], E. Christensen[6], A.M. Lund[6], H.S. Andersen[1].* 1) Dept Biochem & Molec Biol, Univ Southern Denmark, Odense M, Denmark; 2) Department of Clinical Chemistry and Biochemistry, Kinderspital Zürich, Zürich, Switzerland; 3) Universitätsklinikum Freiburg, Germany; 4) INSERM U747, Univ Paris Descartes, France; 5) Dept. of Biomedicine, Aarhus University, Aarhus, Denmark; 6) Dept Clin. Genet., Rigshospitalet, Copenhagen, Denmark.

Bezafibrate (BZF) and Resveratrol (RSV) are potent modulators of gene expression, and are currently being tested as drugs for treatment of fatty acid oxidation defects. The rationale being that BZF or RSV may increase expression sufficiently to alleviate disease manifestation in patients who have missense mutations with some residual enzyme activity. We employed Next Generation Sequencing of RNA (RNA-seq) from fibroblasts stimulated with RSV or BZF to shed light on their global effects on gene expression. Interestingly, aspartoacylase (ASPA) is dramatically up-regulated by both treatments. We validated this by QPCR including other cell types. The effect of RSV could be partly antagonized by sirtinol, indicating Sirt-1 involvement. Moreover, RNA-seq and QPCR/PCR validation revealed RSV and BZF mediated induction of transcription factors with binding sites in the ASPA promoter and up-regulation of proteins important for myelin synthesis and function. Aspartoacylase is important for myelin synthesis in the brain and mutations in ASPA cause Canavan Disease (CD). RSV stimulation of fibroblasts from CD patients with different missense mutations showed dramatic increase (5–10x) in ASPA transcription. We are currently exploring the stimulatory effects of RSV and BZF on ASPA expression further in order to elucidate their potential as drugs for treatment of Canavan Disease.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

738T

The ingestion of a high protein diet during gestation magnifies the *in vivo* production of metabolites in MMA mice *in utero*. *D. Brown[1, 2], J. Senac[1], J. Sloan[1], I. Manoli[1], C. Venditti[1]*. 1) National Institutes of Health, Bethesda, MD; 2) Washington Hospital Center, Washington, DC.

Objective: To use a mouse model of methylmalonic acidemia (MMA) to show that dietary manipulation during gestation can have an effect on an affected fetus. Introduction: There are several reports of prenatal treatment with vitamin B12 when the fetus has been affected with MMA. However, there have been no studies to examine the effects of maternal diet on the developing fetus. Because both MMA patients and Mut -/- mice display increased methylmalonic acid in utero, we hypothesized that excess of branched chain amino acids during development could stimulate the production of abnormal metabolites and potentially modify the disease phenotype. Materials/Methods: Mut +/− were bred for timed mating studies. Daily plug checks were performed. The mice were assigned to either a regular mouse diet (20% by weight) or a high protein mouse diet (70% by weight). On embryonic day 19, the pregnant females were sacrificed and the feti genotyped. Methylmalonic acid was measured in the amniotic fluid and plasma, and the liver and kidneys were fixed for histochemical studies. Results: On embryonic day 19, Mut -/- pups had methylmalonic acid levels averaged 306 micromolar (N=3) when the mother was fed regular chow whereas the average methylmalonic acid concentration was 468 micromolar (N=2) when the mother ingested a high protein diet. Maternal serum MMA levels were 3.6 and 6.8 micromolar respectively. Conclusions: A trend of higher MMA levels was seen in fetuses exposed to a high protein diet in utero than exposed to a regular diet. With simple dietary manipulation, serum levels of MMA were increased suggesting that maternal diet can accentuate metabolic abnormalities in an affected fetus. These results suggest that prenatal dietary treatment can be important to improving the outcomes of patients with MMA.

739F

Danshen Extract regulates energy metabolism by activating the AMPK-SIRT1-PGC-1α. *Y.H. Cho[1], C.R. Ku[1], H. Lee[1], E.J. Lee[1,2]*. 1) Endocrinology, Brain Korea 21 Project for Medical Science, Institute of Endocrine Research, and Severance Integrative Research Institute for Cerebral & Cardiovascular Disease, Yonsei University College of Medicine, Seoul, South Korea; 2) Endocrinology, Northwestern University Feinberg School of Medicine, Chicago, IL USA.

Danshen, dried root of Salvia miltiorrhiza, is one kind of traditional Chinese medicine that has many effects on metabolic diseases. However, the mechanisms by which Danshen extract (DE) functions as a metabolic regulator remain largely unknown. Here we demonstrate that DE regulates energy homeostasis in brown adipocytes through activation of AMP-activated protein kinase (AMPK) and Sirt 1, resulting in mitochondrial biogenesis function. Furthermore, oral administration of DE to rodents exerts AMPK-Sirt1-PGC-1α levels in hypothalamus and contributes to body weight. AMPK phosphorylation and Sirt1 activation levels were increased by DE treatment in brown adipocytes as well as in hypothalamus, resulting activation of its down stream target, peroxisome proliferator-activated receptor-γcoactivator-1α (PGC-1α). Activation of AMPK and Sirt1by DE in brown adipocytes enhanced mitochondrial biogenesis by induction of key metabolic regulator, PGC-1α. The activation of these key metabolic sensors results in enhancement of uncoupling protein 1(UCP-1), known as thermogenin, accounting for the beneficial metabolic effects of DE. These results indicate that DE regulates energy expenditure through an AMPK-Sirt1-PGC-1α-dependent mechanism in adipocytes and hypothalamus.

740T

Protocols for sick day management and emergency interventions for mitochondrial disorder patients are important for family education, treatment and support. *M. Fredrich, A. Perszyk.* UF, Jacksonville, FL.

Treatment protocols are an important and useful clinical tool for the care of patients with complex medical disorders. Mitochondrial disorders are often unique and individual patient characteristics must be learned by patient and family and by the team of nurses and doctors caring for them. Differentiating signs and symptoms into several steps can help a family manage at home, and know when to call or come in for acute management of the symptoms. We have been building individual protocols for our mitochondrial disorder patients. We find that careful review of systems (ROS) can be helpful at finding patterns and individualize patient treatments. This checklist helps to educate the patient and family about the range of their symptoms and provide a more consistent approach to terminology. We have begun using a checklist and reporting system to help with patient checkup. This helps to review the range of clinical signs and symptoms that are relevant to the particular patient and family. Illustrated in this presentation is our clinical checklist, dietary plan, medications, and patient specific concerns that help the process of learning individual patient characteristics. This greatly facilitates the treatment steps that work for the family at home and in the emergency department and hospital settings.

741F

High Fat Diet Paradoxically Improves Cold Survival in a Mouse Model of Medium-Chain Acyl-coA Dehydrogenase Deficiency. *S.E. McCandless[1], L. Rice[1], C.L. Hoppel[2]*. 1) Department of Genetics, Case Western Reserve University, Cleveland, OH; 2) Center for MItochondrial Diseases, Case Western Reserve University, Cleveland, OH.

Background: The natural history of medium-chain acyl-coA dehydrogenase (MCAD) deficiency has been significantly altered by early diagnosis through newborn screening leading to prompt initiation of therapy. Unfortunately, it is not clear which component of the treatment plan is required for improved outcome. Some clinical centers use a fat reduced diet, which can minimize excretion of metabolites of fatty acid oxidation (FAO), but has not been clearly shown in a clinical or experimental setting to be necessary for improved outcomes. **Objective:** We set out to use the mouse model of mcad deficiency (generated by P. Wood and further refined by us) as a tool to investigate the clinical and biochemical consequences of human MCAD deficiency and to evaluate the effect of a high fat diet on survival during cold challenge at 4° C for 4 hours. **Methods:** The original mouse strain was created on a mixed 129 and C57/B6 background. We have backcrossed from this mixed background >10 generations to produce additional strains with the mcad deficient allele on "pure" B6 (B6 $^{mcad-/-}$) and AJ (AJ $^{mcad-/-}$) backgrounds. Clinical investigations included 4 hour cold challenge (4 hours at 4°C after a 24 hour fast) with the primary outcome measurement of survival (as defined by spontaneous respiration and rectal tempature of >19° C). **Results:** In the 4 hour cold challenge, eating a regular diet, B6 $^{mcad-/-}$ and AJ $^{mcad-/-}$ had much worse survival than wild type littermates (14% vs. 71% and 14% vs. 100%, respectively, n=7 for all groups) and shorter time to death (p<.02) as defined by rectal temperature <19° C. After eating a high fat diet, B6 $^{mcad-/-}$ littermates (n=16 high fat diet) all survived the same cold challenge at 10 weeks post-gestation (p=.0001 compared to regular diet, n=7), regardless of whether the high fat diet began in the mother during gestation or at the time of weaning. Similarly, AJ $^{mcad-/-}$ fed a high fat diet (n=8) had markedly improved survival compared to the regular diet (n=7), with 87.5% vs 14% surviving (p=.02). **Conclusions:** Preliminary studies with two pure background strains of *mcad* deficient mice show evidence of markedly improved tolerance to cold challenge in animals fed a high fat diet compared to those fed normal mouse chow, which is relatively low in fat. Additional studies are underway to explore the possible mechanism for this finding and to identify biomarkers that might be used to look for a similar response in humans.

742T

Sapropterin Treatment for Toddlers with PKU. *N.L. Pallone.* Canadian PKU & Allied Disorders Inc., Toronto, ON, Canada.

Treating PKU Toddlers with sapropterin is still relatively rare, especially in Canada and Europe, and yet this age group can perhaps benefit most from this drug treatment, due to the sensitivity to high or unstable blood-phe levels of the rapidly developing brain. This case report looks at one particular patient (RP) on sapropterin from 20 months to 48 months of age. Information including genotype, phenotype, blood phe levels and stability of levels prior to using sapropterin and during drug treatment is analyzed. This case report will also look at why treatment was initiated when RP was in good metabolic control on dietary treatment alone, and what points should be considered when assessing clinical responsiveness to sapropterin. We will look at what factors can help determine which patients will respond to sapropterin treatment, and why it is important to move forward with this treatment for this young age group of PKU patients. Sapropterin treatment in this patient clearly lowered levels and reduced deviations of levels. An added benefit has been an increased phe tolerance which allows for more natural protein and a better chance of diet adherence once RP becomes a teen and adult. The author will also touch upon the concerns of caregivers in regards to funding criteria of sapropterin specifically as it relates to the risks of normalizing a phe-restricted diet, and how to avoid those risks.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

743F

Ubiquinol preparations are numerous and selection of appropriate mitochondrial supplement requires careful scrutiny by family and doctor. *N. Franklin, A. Perszyk.* UF-Jax, Med. Genetics, Jacksonville, FL.

Ubiquinol and Ubiquinone are two related compounds felt to be beneficial when supplemented in the diets of patients with mitochondrial pathway defects. Ubiquinol is widely available in different formulations with a variety of biochemical, and health-related marketing claims. We reviewed over 100 commercially available Ubiquinol products. Product claims, ingredients and manufacturer supplied information were compared with cost per 100mg. Some product formulations appeared more deserving of consideration due to claims of enhanced mitochondrial functioning, patient compliance, and potential measurable health benefits. Selection of ingredients can be especially relevant to certain patients with absorption issues, or dietary fat restrictions. Our findings uncovered significant disparities in the price of the items sold. The price ranged from $0.12 to $3.12 per 100mg. Out of pocket costs for supportive treatments of mitochondrial disorders can be considerable for a family on a budget. The average consumer and physician would benefit greatly from this information to assist them in determining appropriate product selection. We feel that this compilation of data warrants additional investigations into true product efficacy. Well designed physiological studies relating to the duration of action, timing, dosing and frequency of Ubiquinol products for patients with mitochondrial disorders needs further investigation before a more coherent understanding of intra-cellular Ubiquinol benefits can be assessed.

744T

The role of N-acetylcysteine in treating mitochondrial liver disease. *AK. Niemi, GM. Enns.* Dept of Pediatrics, Division of Medical Genetics, Stanford University, Stanford, CA.

Mitochondrial diseases are often multisystem disorders that can present as liver failure or have nonspecific liver dysfunction as part of the phenotype. They are characterized by increased oxidative stress due to respiratory chain dysfunction and production reactive oxygen species (ROS). Treatment includes antioxidants aiming at decreasing the production of ROS or reducing them to a non-reactive form. Glutathione (GSH) is a major cellular antioxidant protecting cells from free radical damage. Acetaminophen overdose is characterized by elevated transaminases and liver synthetic dysfunction. GSH depletion as the cause of acetaminophen-induced liver failure is well established (Heard 2008, Hinson et al. 2010). The standard treatment is N-acetylcysteine (NAC), which prevents liver injury by restoring hepatic GSH. We describe two patients with a mitochondrial disorder with elevated transaminases and liver dysfunction that improved after treatment with NAC. Patient 1 is a 5-yo girl with mitochondrial DNA deletion syndrome characterized by failure to thrive, bone marrow failure, short stature, exocrine pancreatic insufficiency and panhypopituitarism. At age 5y she was hospitalized due to urinary tract infection and was noted to have elevated AST 500 u/L, ALT 400 u/L and INR 1.4. Extensive viral studies were negative. A week after admission oral NAC was started (150mg/kg/d). Within 24h AST, ALT and INR started improving and normalized in a week. Patient 2 is a 4-yo girl who initially presented with neonatal lactic acidosis, hyperammonemia and elevated transaminases and has had frequent hospitalizations with similar findings. Urine organic acids show abnormal elevations of lactate, fumarate, and malate. An extensive workup has not revealed a diagnosis. She has remained on a mitochondrial antioxidant cocktail because her condition worsens upon weaning these cofactors. At age 3y she had fever, cough and a nasopharyngeal swab positive for influenza A. Transaminases were higher than during other hospitalizations (AST~2,000 u/L, ALT~1,600 u/L), INR was 2.6. NAC was started and within 3 days AST, ALT and INR normalized. Both patients have remained on NAC. Mitochondrial patients have low GSH levels compared to controls (Atkuri et al. 2009) and thus their tissues, including liver, may be more sensitive to any cellular stress that depletes GSH. Therefore, NAC may be a therapeutic intervention worth considering in patients with mitochondrial dysfunction and liver disease.

745F

Phase 3 Study of Migalastat HCl for Fabry Disease: Stage 1 Results. *K. Nicholls[1], C. Feliciani[2], S. Shankar[3], F. Ezgu[4], S.G. Janmohamed[5], S.M. Laing[5], R. Schroyer[6], A.C. Bragat[7], S. Sitaraman[7], V. Sniukiene[7], R.E. Winkler[7], P. Boudes[7].* 1) Royal Melbourne Hospital, Parkville Au, Victoria, Australia; 2) Policlinico Universitario Agostino Gemelli, Roma, Italy; 3) Emory University, Atlanta, GA, USA; 4) Gazi University, Ankara, Turkey; 5) GlaxoSmithKline, Stockley Park, UK; 6) GlaxoSmithKline,Renaissance, PA, USA; 7) Amicus Therapeutics, 1 Cedarbrook Drive, Cranbury, NJ 08512.

Background: Fabry Disease (FD) is a rare X-linked deficiency in the lysosomal enzyme α galactosidase A (α-Gal A). While FD is typically more severe in homozygous males, heterozygous females are also often symptomatic. If untreated, many patients will develop serious end organ complications of kidney failure, cardiac arrhythmias, and/or stroke. There is an unmet need for oral therapies for FD, for which the only currently available treatments are intravenous Enzyme Replacement Therapies (ERT). Migalastat HCl (AT1001/GR181413A) is an investigational pharmacological chaperone that binds to the active site of α-Gal A and increases enzyme stability and trafficking to the lysosome, where it can metabolize its predominant substrate (globotriaosylceramide, GL-3). The relatively slow progression of FD would require prolonged studies to evaluate the effect of therapy on clinical outcomes. Reduction of GL-3 deposition in kidney has previously been used as an indicator of a clinically meaningful impact of therapy on the pathophysiology of FD. Methods: Stage 1 of Study AT1001–011/NCT 00925301 is a 6-month, double-blind, randomized, placebo controlled evaluation of the efficacy and safety of oral migalastat HCl in male and female patients who were treatment-naive or off ERT for at least the preceding 6 months. The study population was enriched by requiring that patients had urine GL-3 levels at least 4 times above upper limit of normal and also had a mutation predicted to be amenable to migalastat HCl, as determined by an in vitro cell-based assay. Patients were stratified by gender and randomized 1:1 to receive either placebo or migalastat HCl 150 mg every other day. The primary objective after 6 months of treatment was to evaluate the effect of migalastat HCl on clearance of kidney GL-3, as assessed by quantitative histological scoring of GL-3 inclusions in peritubular capillary cells (PTCs). Patients were considered responders when they had at least a 50% reduction from baseline to Month 6 in the average number of GL-3 inclusions per PTC. A key secondary endpoint was the effect of migalastat HCl on the percent change from baseline to Month 6 in urine GL-3. Results: Enrollment for this study has been completed (67 patients randomized). Efficacy and safety results after 6 months of double-blind treatment will be summarized.

746T

Differential risk for Parkinson disease in males and females with Gaucher disease. *D. Elstein[1], R. Alcalay[2], T. Dinur[1], G. Altarescu[1], E. Ben-Chetrit[1], A. Zimran[1].* 1) Gaucher Clinic, Shaare Zedek Medical Ctr, Jerusalem, Israel; 2) Department of Neurology, Columbia University Medical Center, New-York, NY, USA.

Background: Male gender is a risk factor in idiopathic Parkinson's Disease (PD), but not in many forms of genetic PD (e.g., LRRK2, PARKIN). Whether gender plays a role in PD risk among Gaucher disease (GD) patients is unknown. Objective: To examine the role of gender in PD penetrance among Type-1 GD patients. Methods: Medical files of patients >18 years of age between 1990–2010 were reviewed for PD signs or clinical diagnosis by a neurologist. Available patients with suspected parkinsonism underwent an additional neurological examination by a movement disorders specialist. Kaplan Meier survival curves were used to estimate age-specific risk for PD among males and females. Age-specific risk for PD was compared between males and females using Cox Hazard ratio model. Results: 510 type 1 GD adults (233 males; 45.7%) had records reviewed. 12 patients with suspected parkinsonism were identified; 11 were confirmed by a movement disorder specialist, using UK PD Brain Bank Criteria for PD (2.2%). PD diagnosis was associated with male gender (81.8% versus 44.9% male, p= 0.027) and older age (mean age, GD/PD=62.8, GD non-PD=47.1, p=0.004). Age-specific risk for PD was higher in males compared to females (p=0.039, Cox Hazard ratio model). At age 65 the cumulative risk for PD among GD among males 7.7% and among females was 1.5%. Conclusion: As opposed to other genetic forms of PD, male gender is a significant risk factor for PD among GD patients.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

747F

Potential cluster of MPS IV A in Northeast Brazil. *F. Kubaski[1,2], A.C. Brusius-Facchin[1,2], P.F.V. Medeiros[3], C.S.E. Gondim[3], R. Giugliani[1,2,4], S. Leistner-Segal[1,2].* 1) Medical Genetics Service, Hospital de Clinicas de Porto Alegre, Porto Alegre, RS, Brazil; 2) Postgraduate Program in Clinical Sciences, UFRGS, Porto Alegre, RS, Brazil; 3) Federal University of Campina Grande, Hospital Alcides Carneiro, Campina Grande, PB, Brazil; 4) Department of Genetics, UFRGS, Porto Alegre, RS, Brazil.

Mucopolysaccharidosis IV A (MPS IV A) or Morquio syndrome type A is an autosomal recessive inborn error of metabolism caused by deficiency of lysosomal enzyme N-acetylgalactosamine-6-sulfate sulfatase (GALNS), which results in storage of keratan sulfate and chondroitin-6-sulfate. MPS IV A is a rare disorder, and precise epidemiologic data in Latin America and Brazil are scarce. Surprisingly, in a specific region of Northeast Brazil, in the state of Paraíba (PB), 7 index cases were diagnosed with MPS IV A. Objectives: Our aim was to characterize the genotype of these patients, through GALNS gene analysis. Methods: We have developed a protocol to identify GALNS common mutations by ARMS-PCR. Five common mutations found in previous studies performed in Brazilian patients (p.S341R, p.L307P, p.N164T, p.G139S, p.G116S) were analyzed for patients included in the MPS BRAZIL NETWORK protocol. Results: These 5 mutations were found in 10 out of 26 MPS IV A patients previously diagnosed by enzyme assay, including 4 patients from PB. The p.S341R mutation was found in homozygosis in 2 patients and in heterozygosis in one patient. One patient was negative for the presence of this mutation. For further 3 patients from this region we tested only for the presence of the p.S341R mutation and the results were 1 homozygous, 1 heterozygous and 1 negative. Thus, for a total of 7 patients from PB we have an allele frequency of 50% for this specific mutation. Conclusions: These 5 mutations can be considered common in Brazilian patients once they were found in 38 % of our cases. This number justifies a strategy based on initial screening for these 5 recurrent mutations in our population. The identification of heterozygous individuals within these families will be important for genetic counseling and for estimating the disease prevalence in this region. Further studies to identify the origin of this mutation, including haplotype and segregation analyses are in progress, and will be evaluated in conjunction with epidemiological data. The presence of a specific mutation for MPS in a cluster was already observed in MPS VI in another area of Northeast Brazil. In this area, a comprehensive population medical genetics approach was implemented to prevent and treat early MPS VI, including training of health agents, education to the community, genetic counseling and targeted newborn screening. A similar approach is under consideration for MPS IV A in this potencial cluster.

748T

Persistent Thrombocytopenia in Type I Gaucher Disease : evaluation and management. *H. Rosenbaum.* HEMATOLOGY, RAMBAM MEDICAL CAMPUS, HAIFA, Israel.

Backgound: Type I Gaucher disease (GD) the non-neuronopathic form is characterized by hepatosplenomegaly, pancytopenia and skeletal complications due to the accumulation of glucocerebroside in macrophages. Thrombocytopenia is usually related to hypersplenism and/or infiltration of bone marrow by the lipid-laden macrophages namely Gaucher cells. Enzyme replacement therapy (ERT) restores the hemoglobin and platelet count in treated GD patients within 12–24 months of treatment. In GD patients, including ERT treated, with persistent low platelet counts other etiological factors should be considered including immune thrombocytopenia(ITP),and concomitant hematological malignancies or metastatic bone marrow by other neoplastic disease. Goals: delineate guidelines and alogarithm for the evaluation of persistent thrombocytopenia and bleeding phenomena in GD. Focus on the approach to treatment especially new modalities including immunomodulation by steroids (high dose Dexamethason) IVIG and Rituximab. Discuss dosage of ERT and switch to other ERT in persistent thrombocytopenia. Emphasize whether steroid therapy is feasible in GD. Debate on the role of thrombopoeitin receptor analogue and splenectomy in patients with ITP and GD. Report successful procedure of splenic artery embolisation which might replace splenectomy in GD pthrombocytopenic patients.This ambulatory therapy might be useful in cases of urgent need to raise platelet counts prior to surgery. Methods: report and discuss cases of GD and thrombocytopenia and summarize the evaluation measures and new therapy strategies.

749F

Chemotactic factors critical for increased immunological cell invasion in Gaucher disease mice. *M.K. Pandey[1, 2], N.A. Jabre[1], G.A. Grabowski[1,2].* 1) Division of Human Genetics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH 45229, USA; 2) Department of Pediatrics, University of Cincinnati, College of Medicine, Cincinnati, Ohio 45229, USA.

Gaucher disease results from mutations in *GBA1* that cause functional disruption of the encoded lysosomal enzyme, acid β-glucosidase (β-D-glucosyl-N-acylsphingosine glucohydrolase, EC 4.2.1.25; GCase). The consequent excess accumulation of glucosylceramide (GC) and glucosylsphingosine (GS) in lysosomes is central to the disease pathogenesis with classical involvement of macrophage (MΦ) lineage cells of visceral organs, bone, or brain. Several studies have implicated the increased secretion of chemokines and increased infiltration of variety of immunological cells, e.g., T lymphocytes, B lymphocytes, MΦs, dendritic cells (DCs), and neutrophils (PMNs) in Gaucher disease. Trafficking of immunological cells to the sites of inflammation requires the presence of chemokines. Although increases of different immunological cells and several chemokines are known in Gaucher disease, the mechanisms and specific molecules involved in causing the increased influx of immunological cells in this disease is unclear. Increased levels of I-309, MCP-5, CXCL2, CXCL-9, CXCL-10, CXCL-11, CXCL-13, and their corresponding leukocytes i.e., MOs (monocytes), MΦs, DCs, PMNs, T, and B cells have been identified in the circulation of 9V/null Gaucher disease mice. 9V/null sera was shown to attract different immunological cells was shown by in vitro chemotaxis in Boyden chambers and by flow cytometery. Enhanced chemotaxis of 9V/null lung, spleen, liver, and bone marrow derived immunological cells, i.e., MΦs ($CD11b^+$ $F480^+$), PMNs ($Gr1^{high}$ $CD11b^+$), DCs ($CD11c^+$ $CD11b^+$), T lymphocytes ($CD3^+$ $TCR\beta^+$), and B lymphocytes ($B220^+$ $CD19^+$) into 9V/null sera supports above chemotactic factors as responsible for increased infiltration of leukocytes in Gaucher disease.

750T

Systematic Screening for treatable inborn errors of metabolism in intellectual disability patients: First Study Results. *S. Stockler[1], R. Salvarinova[1], G. Horvath[1], Y. Lillquist[1], H. Vallance[2], S. Sinclair[2], M. Lafek[1], T. Murphy[2], A. Lehman[4], M. Patel[4], M. Demos[3], M. Connolly[3], C. van Karnebeek[1], TIDE-BC, Treatable Intellectual Disability Endeavor in British Columbia, Vancouver, Canada.* 1) Biochemical Diseases, Dept Pediatrics, BC Children's Hosp, Vancouver, BC, Canada; 2) Biochemical Genetics Laboratory, BC Children's ospital, Vancouver, BC, Canada; 3) Pediatric Neurology, BC Children's Hospital, Vancouver, BC, Canada; 4) Medical Genetics, BC Children's & Women's Hospital, Vancouver, BC, Canada.

Background: Current recommendations to investigate genetic causes of intellectual disability (ID) are based on frequencies of single conditions and yield of tests rather than availability of causal therapy. Early recognition of treatable inborn errors of metabolism (IEM) is crucial for minimizing brain damage and reducing of health-care costs. **Methods & Results:** Our literature review identified 81 IEM, presenting predominantly with ID and amenable to causal therapy. Although evidence is limited, therapies are often effective, safe, accessible. We translated this knowledge into a diagnostic protocol: The 1st tier comprises metabolic screening tests in blood (plasma amino-acids, homocysteine) and urine (creatine metabolites, glycosaminoglycans, oligosaccharides, organic acids, purines/pyrimidines), with potential to identify 62% of treatable IDs. The second tier focuses on remaining disorders, requiring `single test per disease' approach. A freely available App (www.treatable-id.org) supports the protocol. During first 6 months 130 patients were enrolled in a funded study implementing the protocol;a confirmed diagnosis was established in 25% (5 treatable IEM), probable in 17% (4 treatable IEM). We will present the indvidual diagnoses (treatable IEMS) in more detail. **Conclusions:** Our preliminary results indicate that treatable IEM constitute etiology in 4–7% of all ID patients evaluated in tertiary care centres, and account for a considerable proprotion of all diagnoses 16%). Our novel approach may well improve outcomes and convince colleagues/ policymakers to change practice and care for individuals with ID.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

751F

A novel homozygous mutation in EARS2 causing a fatal multisystem infantile disease. *A. Pyle[1], B. Talim[2], H. Griffin[1], H. Topaloglu[2], M. Santibanez-Koref[1], P.F. Chinnery[1], R. Horvath[1]*. 1) Institute of Genetic Medicine, Newcastle University, Newcastle upon Tyne, United Kingdom; 2) Department of Pediatrics, Hacettepe University, Ankara, Turkey.

Mitochondrial disorders are clinical phenotypes, usually a result of mitochondrial dysfunction, caused by mutations in either mitochondrial DNA (mtDNA) or nuclear-encoded (nDNA) mitochondrial genes. The majority of these gene defects manifest as histological and biochemical abnormalities in affected tissues. Clinical phenotypes are typically early-onset, severe and often fatal, implying the importance of efficient mitochondrial translation from birth. mtDNA-related disorders are well defined and the analysis of mtDNA is standardized. More recently, research focus has shifted to nuclear genes and a number of new mutations have been identified in genes involved in mitochondrial protein synthesis. The resulting multiple-respiratory complex disorders are heterogeneous and poorly defined at the molecular level. We present a patient, the third child of consanguineous Turkish parents with severe infantile multisystem disease, involving the brain and liver. The proband died at 3 months of age of necrotising bronchopneumonia. Biochemical and histochemical analysis of skeletal muscle identified a severe, combined, deficiency of respiratory chain complexes I and IV. Subsequent mtDNA analysis excluded mtDNA depletion, deletions or point mutations in affected tissue. Common nDNA causes of mitochondrial dysfunction, including nuclear encoded mitochondrial elongation factors or ribosomal factors were also excluded. Through a combination of homozygosity mapping and whole exome sequencing we identified a homozygous missense mutation in mitochondrial glutamyl-tRNA synthetase (EARS2). Pathogenicity was confirmed by familial segregation and exclusion in ethnically matched controls. Nuclear-mitochondrial disorders are characterised by disturbed mitochondrial translation and EARS2 is the newest member of this gene family. We have shown that in addition to central nervous system involvement, some patients develop a progressive and fatal course in infancy and therefore our case widens the clinical spectrum of EARS2 mutations.

752T

Novel phenotype associated with *OPA* mutations? *M. Tesarova[1], V. Stranecky[2], H. Kratochvilova[1], Z. Hajkova[1], J. Sladkova[1], J. Spacilova[1], H. Hansikova[1], T. Honzik[1], H. Hartmannova[2], L. Noskova[2], L. Piherova[2], E. Lalonde[3], J. Majewski[3], S. Kmoch[2], J. Zeman[1]*. 1) Department of Pediatrics and Adolescent Medicine, Charles Univesrity in Prague and General University Hospital in Prague, Prague 2, Czech Republic; 2) Institute of Inherited Metabolic Disorders, First Faculty of Medicine, Charles University in Prague and General University Hospital in Prague, Prague, Czech Republic; 3) Department of Human Genetics, McGill University Health Centre, Montreal, Quebec, Canada.

Mutations in *OPA1* gene, encoding a dynamin-like mitochondrial GTPase indispensable for mitochondrial network structure and morphology, were found to be responsible for autosomal dominant optic atrophy (ADOA). Moreover *OPA1* mutations may be responsible for a continuum of phenotypes ranging from mild disorders affecting only the optic nerve to severe and multi-systemic diseases. In Caucasian non-consanguineous family, 3 boys were born with lactic acidosis, a severe muscle hypotonia, breathing and breastfeeding difficulties. Patient 1 (P1) died on day 9 with early metabolic lactate acidosis. In his younger brothers (P2, P3), central hypotonic syndrome was present (2nd month of life). Severe vision and hearing impairment developed at 4 months followed by hypertrophic cardiomyopathy. Brain atrophy was found by MRI in P2 at 4 month of age. In P3, brain atrophy and Leigh syndrome were observed at autopsy. P2 died at the age of 10 months, P3 died at the age of 12,5 months. Metabolic examination revealed lactic acidosis with increased excretion of Krebs cycle intermediates and 3-methylglutaconic acid (only P3). In isolated muscle mitochondria of P2 and P3, mild decrease in activity of complex II+III, complex IV and PDHc were observed. Mutation in mtDNA and *TMEM70, TAZ, SCO2, SCO1* were excluded by direct sequencing in P3. Genome-wide copy number analysis was unremarkable. Subsequently, whole-exome sequencing in P2 and P3 revealed heterozygous mutations: c.1062_1064delAGA (p.E354del, novel mutation) inherited form father and c.1768C>T (p.R590W) inherited from mother in *OPA1* as the most probable candidates. Steady-state levels of OPA1 isoforms were significantly decreased in muscle, heart and frontal core of patient 3 with profound reduction of short isoforms. In cultured skin fibroblasts of P2, P3, and their parents, OPA1 isoform levels were decreased in both patients' and father's sample. More pronounced reduction of short OPA1 isoforms (S3, S4, S5) was observed in patients compared to parents. Moreover, mitochondrial network was more fragmented in patients' fibroblasts compare to parental samples. Nevertheless, RCC activities were in reference range. To conclude, whole-exome sequencing revealed possibly pathogenic mutations in *OPA1* gene. Despite ongoing analyses to confirm *OPA1* mutation as cause of the disease, symptoms described in our patients may enlarge phenotypic variability associated with OPA1. *Supported by RVO-VFN64165/2012 and IGA NT 13114-4/2012.*

753T

Evaluating pain, negative mood, and resilience in patients affected by Fabry disease. *A. M. Lelis[1], R. Duran[1], L. Beckman[1], J. Garbanati[1], W. Wilcox[2], J. Mirocha[2]*. 1) Alliant International University, 1000 S. Fremont, Alhambra, CA; 2) Cedars-Sinai Medical Center, 8700 Beverly Blvd, Los Angeles, CA 90048.

The purpose of the current study was to evaluate pain, negative mood, and resilience in individuals diagnosed with Fabry disease. Fabry disease (FD) is a rare, hereditary condition in which psychosocial correlates of pain experience have been severely understudied. A total of 75 men and women with FD were recruited from Cedars-Sinai Fabry Center in Los Angeles, California. This study utilized a variety of instruments, including: a demographic questionnaire, Brief Pain Inventory, Resilience Scale, Beck Depression Inventory-II, and Beck Anxiety Inventory. The results indicate that pain intensity is positively correlated with depression and anxiety, while pain intensity is negatively correlated with resilience. The results also show that pain frequency acts as a moderator between resilience and depression. While there were no statistically significant differences between men and women with FD, there were clinically important differences in the depression scale between the sexes. The clinical implications of this study are discussed. Specifically, resilience appears to be a key factor in successful adaptation to FD. Clinicians should focus their attention on resilience skills and training rather than solely trying to assess and medicate depression, anxiety, and pain. Future studies in the FD population are needed to investigate the effect of resilience training on the ability to manage stress and decrease experience of pain, depression, and anxiety.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

754W

Comparison of the differentiation potential of human mesenchymal stem cells and several animal species. *F. Piryaei[1], M. Ramezani[2], M. Kadivar[3], F. Piryaei[4]*. 1) Dept. of Biology, Faculty of Science, Payame Noor University, Center of Tehran,, Tehran, Iran; 2) Dept. of Biology, Faculty of Science, Islamic Azad University, Ashtian Branch, Ashtian, Iran; 3) Dept. of Biochemistry, Pasteur Institute of Iran, Tehran, Iran; 4) Dept. of Medical Genetics, National Institute of Genetic Engineering and Biotechnology (NIGEB), Tehran, Iran.

Considering the increasing use of mesenchymal stem cells isolated from various sources in the clinic, the basic studies needed to evaluate proliferation, differentiation and other biological characteristics of them done. In the present study, efficiency and power differentiation in five different categories mesenchymal stem cells are compared with each other. Bone marrow mesenchymal stem cells from rabbits, rats, C57 mice, chicken and mesenchymal stem cells derived from human knee synovium tissue were cultured with a density of 5000 cells/cm2 in 12-chamber dishes. After the cells reached the appropriate level of growth, specific differentiation inducers were added. 21 days later, differentiation of all stem cells to adipocyte, osteocyte and chondrocyte were evaluated and compared using specific staining methods. Our results indicated that all five mesenchymal stem cells were able to differentiate into osteocyte, adipocyte and chondrocyte. Moreover, the highest potentials for differentiation to adipocyte and osteocyte and chondrocyte were seen in mesenchymal stem cells derived from chicken bone marrow and human synovium, respectively. In this study, differences in the differentiation potential of mesenchymal stem cells isolated from different tissues and species were observed that could be caused by small differences in their environment in vivo. Understanding these differences can lead to more effective use of these cells in the clinic.

755W

Bone marrow derived cells as a stable source of sialic acid for mice with GNE myopathy. *MC. Malicdan[1,3], K. Momma[2], F. Funato[3], YK. Hayashi[3], I. Nonaka[1], M. Huizing[1], W. Gahl[1,4], CF. Boerkoel[1], I. Nishino[3], S. Noguchi[3]*. 1) MGB, NHGRI, National Institutes of Health, Bethesda, MD, USA; 2) Department of Neuromuscular Research, Natiional Institute of Neuroscience, NCNP, Tokyo, Japan; 3) Department of Neurology, National Defense Medical College, Saitama, Japan; 4) Office of the Clinical Director, NHGRI, NIH, Bethesda, MD, USA.

GNE myopathy, previously known as distal myopathy with rimmed vacuoles (DMRV), or hereditary inclusion body myopathy (hIBM), is an autosomal recessive myopathy characterized by progressive weakness and atrophy involving the distal muscles, and myofiber vacuolation and degeneration. GNE myopathy is secondary to mutations in the GNE gene, which encodes the bifunctional enzyme which catalyzes the first two crucial steps in sialic acid biosynthesis. We and others have shown that hyposialylation is one of the key players in disease pathomechanism. This notion is supported by our recent publications showing that the myopathic phenotype in the existing DMRV/hIBM mouse was prevented by giving exogenous oral sialic acid metabolites or modified ManNAc conjugates. Despite these promising results, designing an oral therapeutic regimen is often challenged by the the extremely rapid excretion of orally administered sialic acids that potentially requires frequent or almost continuous dosing. In this study, we established allogenic bone marrow transplantation (BMT) as a choice for therapy.Bone marrow cells were isolated from CAG-GFP expressing mice and injected intravenously to 30 weeks old DMRV mice after a single dose of sublethal irradiation. Chimerism and cell surface sialylation of peripheral leukocytes after RBC lysis was assessed four weeks after cell transplantation and monthly thereafter. Results show that there was a clear increment in cell surface sialylation and correlated with chimerism. GFP-positive donor cells were seen to engraft into skeletal muscles and other organs of DMRV/hIBM mice. DMRV/hIBM mice subjected to BMT had a marked improvement in lifespan and motor performance. Analysis of muscle contractile properties and pathology revealed an improvement in muscle phenotype. Evaluation of specific glycoproteins in the muscle demonstrated a recovery of cellular sialylation. GNE activity and sialic acid levels in several organs revealed a remarkable increase as compared to non-treated mutant mice. Our results provide a proof of concept for the utility of cell based therapies in DMRV/hIBM, and that hematopoetic cells can be a good source of sialic acid production. We envision that with this strategy, GNE myopathy patients can have a lasting supply of sialic acid that may benefit them towards disease recovery.

756W

Lysine Restricted Diet as Novel Therapy for Pyridoxine Dependent Epilepsy: First Observational Study. *C. van Karnebeek[1,5], H. Hartmann[2], S. Jaggawantru[3,5], J.P. Collet[3,5], B. Plecko[4], S. Stockler[1,5], Pyridoxine Dependent Epilepsy Consortium.* 1) Biochemical Diseases, BC Children's Hospital, University of British Columbia, Vancouver CA; 2) Clinic for Liver, Kidney & Metabolic Diseases, Hannover Medical School, Hannover, DE; 3) Dept of Pediatrics, BC Children's Hospital, University of British Columbia, Vancouver CA; 4) Neurology Dept, Zurich University Hospital, Zurich, CH; 5) TIDE-BC (Treatable Intellectual Disabiliy Endeavour in British Columbia), BC Children's Hospital, Vancouver CA.

Background: Pyridoxine Dependent Epilepsy classically presents as a neonatal onset epileptic encephalopathy. This cerebral organic aciduria is caused by deficiency of antiquitin (ATQ) (alpha-aminoadipic semialdehyde dehydrogenase, ALDH7A1), an enzyme facilitating lysine catabolism. Despite adequate seizure control with pyridoxine in most cases, neurodevelopmental outcome is poor with 75% of patients suffering developmental / cognitive deficits. **Objective:** To evaluate effects of dietary lysine restriction as add-on therapy on biochemical parameters, seizure control and neurodevelopmental outcomes in patients with molecularly confirmed ATQ deficiency. **Methods:** An observational was performed; 5 patients with confirmed ATQ deficiency were started at age 11mos-11yrs on lysine restriction 50–60mg/kg/day and regular monitoring and outcome assessments during a 12–24mos period. **Results:** The diet was well tolerated; good compliance, no reports of adverse-events. Observed reduction of plasma pipecolic acid on diet was 44–81percent; of pre-treatment level; for urine AASA 57–87percent. (Subjective) Improvements of neurodevelopment, behavior and seizure control are reported. **Conclusions**: Our observational study suggests a positive effect of lysine restriction on relevant outcomes in ATQ deficiency (evidence level 4). The PDE consortium (www.pde-online.org)has initiated further research using novel trial methodologies to generate natural history data and more solid therapeutic evidence, before burdensome dietary therapy becomes mainstay treatment.

757W

High-fat diet rescues lethality of homozygous knock-in R155H mice with VCP-Associated Disease. *K.J. Llewellyn, A. Nalbandian, C. Nguyen, V.E. Kimonis.* Department of Pediatrics, Division of Genetics and Metabolism, 2501 Hewitt Hall, University of California-Irvine, Irvine, CA 92697, USA.

Hereditary inclusion body myopathy, Paget's disease of bone and frontotemporal dementia (IBMPFD), is a rare, progressive and fatal disease caused by mutations in the VCP gene. The most common feature of this disease is the myopathy (IBM), which is present in 90% of individuals with an onset of 30–40 years of age. The progressive muscle weakness can often progress to respiratory failure or cardiomyopathy and cardiac failure that can result in death at 40–60 years of age. The Paget's disease of bone (PDB) aspect is observed in 50% of individuals with an onset in their 30–40s. Frontotemporal dementia (FTD) is observed in 30% of patients with a later onset of 50–60 years of age. ALS is seen in 10% of patients and recently VCP was shown to be the cause of 2% of isolated familial ALS. Studies of patients' myoblasts and the heterozygous knock-in mouse carrying the common VCP R155H/+ mutation, developed in our lab, have placed VCP at the intersection of the ubiquitin-proteasome signaling and autophagy pathways, both mechanisms being considered responsible for the intracellular protein degradation and abnormal pathology seen in muscle and brain. We have recently created a double dominant homozygote mouse (VCPR155H/R155H) that has a more rapid and severe muscle, brain and spinal cord pathology, with most surviving no more than 21 days. Feeding pregnant heterozygous dams a diet with 2.6% increased fat, high in polyunsaturated fatty acids, especially linoleic acid (18:2) and oleic acid (18:1) however results in a dramatic improvement in the survival of their homozygous offspring, who on average can now live up to 10 months. These mice still weigh less than their wild-type litter mates and are significantly weaker as demonstrated by grip strength analysis however, Preliminary histological studies have shown improvement in the muscle pathogenesis in these homozygotes at 3 weeks compared to homozygote's on a normal diet as observed by IHC, western blot and CT-scans. Our hypothesis is that linoleic acid in the fat diet is critical for the embryological neurodevelopment of the mouse. We are currently investigating the optimum diet that completely rescues the lethal phenotype. This discovery provides the opportunity to develop a promising therapeutic strategy for patients with VCP and related diseases.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

758W

Serin diet relieves symptoms of Hereditary Sensory and Autonomous Neuropathy type 1A caused by a c.992 C>T, p.(Ser331Phe), mutation in SPTLC1. *B. Rautenstrauss[1,2], E. Wilichowski[3], E. Holinski-Feder[1], T. Hornemann[4]*. 1) Medizinisch Genetisches Zentrum, Munich, Germany; 2) Ludwig-Maximilians-Unverisity, Friedrich-Baur-Institute, Munich, Germany; 3) Georg-August-University, Goettingen, Germany; 4) University Zurich, Zurich, Switzerland.

Hereditary sensory and autonomic neuropathies (HSAN) are a genetically and clinically heterogeneous group of disorders associated with sensory dysfunction. HSAN1 is a dominantly inherited sensorimotor axonal neuropathy. The patient has 3 healthy siblings and healthy parents. Symptoms: Coinciding with the start of ambulation motor disabilities; generalized hypotrophy; reduced walking distance; unstable gait; multiple falls; no mental retardation; MRI studies of the brain were normal; reduced pain sensitivity: recurrent traumatic and thermal injuries of feet, knees and hands; disturbed wound healing; ulcerated fingertips; unnoticed fracture of a metatarsal bone; at the age of 9 years: bilateral lensectomy due to juvenile cataract; cataract surgery: repetitive ulcerations of the cornea, poor healing tendency; complete retinal detachment (right eye); nerve biopsy: marked wasting of myelinated fibres, axonal damage; axonal motor and sensory neuropathy. Finally a de novo c.992 C>T, p.(Ser331Phe), mutation in the SPTLC1 gene was identified. This mutation turns the sphingolipid synthesis to neurotoxic lipids in this patient as well as in cell cultures. A serin diet (400mg/kg bodyweight) developed in an animal model (SPTLC1 mutation p. (Cys133Trp)) resulted after 3 months in an overall improvement: ameliorated growth of nails and hair; wounds of the skin show faster healing; gain of weight; coughassist could be reduced to only one treatment during the winter period; sweating is reconstituted; ability to an extended period of upright standing. This HSAN patient is an excellent example for effective translational medicine and the first individual in Germany receiving a causal treatment for a Hereditary Sensory and Autonomous Neuropathy.

759W

Pilot trial of high-dose carnitine supplementation in young, non-dysmorphic males with autism spectrum disorders (ASDs). *R. Goin-Kochel[1], F. Scaglia[2], C. Schaaf[2], D. Dang[2], A. Laakman[1], K. Nowell[3], A. Beaudet[2]*. 1) Pediatrics, Baylor Col Medicine, Houston, TX; 2) Molecular & Human Genetics, Baylor Col Medicine, Houston, TX; 3) Educational Psychology, Univ of Houston, Houston, TX.

Mutations in the trimethyllysine hydroxylase epsilon (*TMLHE*) gene cause an inborn error of carnitine metabolism that may be relevant in ASDs (Celestino-Soper, et al., 2012). This discovery led us to hypothesize that neuronal carnitine deficiency associated with low dietary intake, impaired transport, or defective synthesis of carnitine is a risk factor for ASDs, especially in non-dysmorphic males, and that carnitine supplementation may prevent or reverse symptoms of ASDs in this subgroup. We are conducting an open-label pilot trial of the behavioral effects of high-dose carnitine supplementation in such boys. Five males with ASDs (μ age = 6.1y, *SD* = 1.4) were recruited through the Simons Simplex Collection and Autism Treatment Network at Baylor College of Medicine. One child had a *TMLHE* deletion, and all had confirmed ASDs per scores on the *Autism Diagnostic Interview—Revised* (ADI-R) and the *Autism Diagnostic Observation Schedule* (ADOS). Levels of free carnitine in plasma were in the normal range pre-treatment, except for a slightly low level of 23.4 μM in one boy with a history of low plasma carnitine; levels of plasma carnitine rose in all patients on therapy by an average of 50%. Parent-reported behavioral data were gathered on children at 3 time points during an 8-week supplementation period; study clinicians also performed pre- and post-treatment behavioral assessments. Participants began oral supplementation with carnitine at 200mg/kg divided in 3 daily doses. If well tolerated after 1 week, the dose was increased to 400mg/kg per day up to a maximum daily dose of 6g per day. Dosage was reduced in cases where side effects (e.g., fishy odor, diarrhea) became difficult for the family to tolerate. Results indicated average improvements on: the *Aberrant Behavior Checklist* (39%), *Autism Impact Measure—Frequency Scale 1* (20%) and *Frequency Scale 2* (26%), and the *Social Communication Questionnaire—Current* (28%). For the 4 participants with complete post-test data, clinicians observed (a) few-to-no improvements on the ADOS and (b) mild improvements on the *Global Clinical Impression* scale. Evidence for benefit was greatest in the 7 yo with *TMLHE* deficiency and the 4 yo with a history of low plasma carnitine. These data suggest that some children with ASDs may show mild to modest behavioral improvements with carnitine supplementation. However, a larger, double-blinded, placebo-controlled study is necessary to eliminate possible response biases.

760W

RESULTS OF THE PHASE II MULTICENTER, SINGLE ARM, PHASE II CONTROLLED TRIAL `IEDAT-01'. *L. Chessa[1], V. Leuzzi[2], A. Plebani[3], R. Micheli[3], A.R. Soresina[3], I. Quinti[2], D. D'Agnano[3], C. Bozzao[1], P. Lulli[1], M. Magnani[4]*. 1) U.O.D. Medical Genetics, Hosp S Andrea, Via di Grottarossa 1035, I-00166 Roma, Italy; 2) Department of Child Neurology and Psychiatry, Sapienza University, Via dei Sabelli, Roma, Italy; 3) Department of Pediatrics, University of Brescia, Brescia, Italy; 4) Department of Biomolecular Sciences, University of Urbino "Carlo Bo", Via Saffi 2, I-61029 Urbino (PU), Italy.

A six months multicenter single arm phase II controlled trial entitled IEDAT-01 (Evaluation of Effects of Intra-Erythrocyte Dexamethasone Sodium Phosphate on Neurological Symptoms in Ataxia-Teleangiectasia Patients) has been recently performed in Italy. Primary objective of the trial was to evaluate the improvement in central nervous system symptoms measured by International Co-operative Ataxia Rating Scale (ICARS) in patients with Ataxia Teleangiectasia (AT) during a 6 months period of ERY-DEX treatment (dexamethasone sodium phosphate ex vivo encapsulated into human autologous erythrocytes). The ICARS parameters: posture, kinetics, speech and oculomotion were evaluated at time points 1 (enrollment), 4 (after three months of ERY-DEX treatment) and 7 (at the end of ERY-DEX therapy). Secondary objectives were to evaluate the effect of ERY-DEX on: 1) treatment-Emergent Adverse Events, 2) the global health status of the patient, 3) adaptive behavior measured by the Vineland Adaptive Behavior Scale (VABS), and 4) special laboratory parameters. A total number of 22 AT patients (11 females and 11 males, age range 4–18 yrs) were enrolled, following strict inclusion/exclusion criteria; four of them did not complete the trial, one for no compliance and three for low level CD4+ lymphocytes. The trial was completed without notification of major adverse events. The patients perceived the treatment as very positive and after the end of the trial most of them requested to the local Ethical Committees the compassionate use and were authorized to continue the ERY-DEX therapy. The patients who completed the trial can be subdivided in responders (11 of 18; 7 females and 4 males) and non-responders (1female and 6 males). The responders show a diminution of ICARS scores ranging from 6 and 17 points, the non-responders a variation of ICARS scores ranging from 3 and -6. The best neurologic results of the therapy were obtained on kinetic scores; on the contrary, evaluation of posture and oculomotion seems to provide lower improvements. In conclusion, the condition of responders seems to correlate with the loading dosage of ERY-DEX and with the sex of patients, while it is independent on the type of mutation, age and clinical status. New incoming trials would help to ameliorate the treatment and to elucidate the biological bases of responder/non responder condition.

761W

Generation and characterisation of FRDA mouse model cell lines. *C. Sandi, S. Al-Mahdawi, V. Ezzatizadeh, S. Virmouni, M. Sandi, M. Pook.* School of Health Sciences and Social Care, Brunel University, Middlesex, UB8 3PH United Kingdom.

Friedreich ataxia (FRDA) is an inherited autosomal recessive neurodegenerative disorder caused by a GAA trinucleotide repeat expansion mutation within the first intron of the FXN gene. Normal individuals have 5 to 30 GAA repeats, whereas affected individuals have from approximately 70 to more than 1,000 GAA triplets. In addition to progressive neurological disability, FRDA is associated with cardiomyopathy and an increased risk of diabetes mellitus. Currently there is no effective therapy for FRDA, which is perhaps due to the lack of an effective system to test potential drugs. To understand the pathological consequences and to find a potential therapy for FRDA, in vitro and/or in vivo model systems are considered to be essential. We have generated GAA repeat expansion mutation-based FRDA mouse models that exhibit GAA repeat instability, epigenetic changes and progressive mild pathology, representative of FRDA. Using these mouse models, we have established novel cell culture systems, including FRDA mouse primary fibroblasts and neural stem cells (NSCs) for the investigation of FRDA molecular disease mechanisms and in vitro therapeutic testing. Experiments are also under way to differentiate the NSCs to obtain cultures of differentiated neural cells. The FRDA mouse primary fibroblasts have shown less tolerance to both hydrogen peroxide and FAC/BSO-induced oxidative stress compared to control mouse fibroblasts. The frataxin expression levels in FRDA mouse derived fibroblasts are also significantly decreased compared to control fibroblasts. Furthermore, we have found reduced gene expression levels of antioxidant genes, such as Cat, Sod2, Gpx1, and mismatch repair (MMR) genes, such as Msh2, Msh3, Msh6 and Pms2. We are currently investigating the level of expression of these genes in NSCs and differentiated neural cells. In summary, we have generated FRDA mouse fibroblasts and neural cell lines that will be a valuable resource to further understand FRDA pathogenesis and therapy.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

762W

Non-penetrance in RP17? An investigation into the molecular mechanisms underlying retinitis pigmentosa 17 with the view to developing novel gene-based therapies. *A. Pandor, S. Prince, R. Ramesar.* Human Genetics, University of Cape Town, Observatory, Western Cape, South Africa.

Retinitis pigmentosa (RP) is a highly heterogeneous form of inherited blindness that affects more than 1.3 million individuals worldwide. The RP17 form of the disease is caused by an arginine to tryptophan (R14W) mutation in the signal sequence of carbonic anhydrase IV (CAIV). In an effort to elucidate the molecular mechanisms underlying RP17, three cell types were transfected with the wild type (WT) and the R14W mutant form of the protein. We show using immunocytochemistry that unlike transfected WT CAIV which is transported to the plasma membrane of transfected COS-7 and HT-1080 cells, R14W mutant CAIV is retained in the endoplasmic reticulum when transfected into the same cell type. Further analyses of these cells by western blotting reveal that whereas the WT CAIV is processed to its mature form in both these cell lines, significant levels of the R14W mutant protein remain in its immature form. Importantly, flow cytometry experiments demonstrate that compared to WT CAIV protein, expression of specifically the R14W CAIV results in an S and G2/M cell cycle block, followed by apoptosis. Interestingly, when the above experiments were repeated in the human embryonic kidney cell line, HEK-293, strikingly different results were obtained. These cells were unaffected by the expression of the R14W mutant CAIV and were able to process the mutant and WT protein equally effectively. These findings regarding cell type specificity were used as a basis to explore methods of therapy for RP17. In particular, allele-specific small hairpin RNA was used to silence expression of R14W mutant CAIV, and to rescue cells from undergoing cell cycle arrest and apoptosis. A study of specific chaperones involved in protein folding, as well as gene and protein expression studies (microarray and mass spectrometry analysis), were also carried out to determine which proteins that were expressed in HEK-293 cells play a part in the ability to fold, process and transport R14W mutant CAIV. The results of this study have important implications for our understanding of the RP17 phenotype, and in investigating gene and protein therapy for the prevention and treatment of RP17.

763W

Pharmacological Evaluation of a CNP Analogue for the Treatment of Achondroplasia. *F. Lorget[1], N. Kaci[2], J. Peng[1], C. Benoist-Lasselin[2], E. Mugniery[2], T. Oppeneer[1], D. Wendt[1], S. Bullens[1], S. Bunting[1], L. Tsuruda[1], C. O'Neill[1], F. Di Rocco[2], A. Munnich[2], L. Legeai-Mallet[2].* 1) BioMarin, Novato, CA; 2) Inserm U781, Paris Descartes Unversity, Necker hospital.

Achondroplasia (ACH), the most common form of dwarfism, is caused by a gain of function mutation in the fibroblast growth factor receptor 3 gene (FGFR3). The C-type natriuretic peptide (CNP) antagonizes FGFR3 downstream signaling by inhibiting the mitogen-activated protein kinase (MAPK) pathway. The use of CNP as a therapeutic agent for ACH is impracticable due to its short half-life. Here, we are presenting an update on the pharmacological activity of BMN 111, a novel CNP analogue with an extended plasma half-life allowing once daily subcutaneous (SC) administration. In ACH growth plate chondrocytes, we demonstrated that BMN 111 decreased MAPK pathway activation but did not affect STAT pathway activation. In Fgfr3Y367C/+ mice mimicking ACH, daily SC administration of BMN 111 led to the attenuation of the dwarfism phenotype. With 20 days of daily SC dosing of BMN 111 (800 μg/kg) beginning at seven days of age, we observed a significant increase in the axial and appendicular skeleton length with improvements in ACH related clinical features including flattening of the skull, reduced inverted crossbite and straightening of the tibias and femurs. The growth plate defect (reduced height of the pre-hypertrophic/hypertrophic zones and lack of columnar arrangement) was corrected along with the normalization of the size and shape of the proliferative and hypertrophic chondrocytes. A rescue of the height and architecture of the different zones of the growth plate was observed. The height of the proliferative zone was increased; the replicating chondrocytes were organized into columns parallel to the long axis of the bone. In the hypertrophic zone, the terminally differentiated chondrocytes recovered a columnar alignment. These data support further development of BMN 111 for the treatment for ACH and hypochondroplasia (HCH).

764W

Role of vitamin D in osteoarthritis knee: A six month double blind, randomized, placebo control trial. *S. R. Nath, D. Sanghi, S. Raj, K. Baghel, A. Mishra.* Orthopaedic Surgery, KG Medical College, CSM Medical University, Lucknow, UP, India.

Background: Inadequate sunlight exposure and lower serum levels of 25(OH)D appears to be associated with an increased risk for progression of OA knee.This study was planned to assess the effect of vitamin D on the progression of disease. Methods: A six month, double blind, randomized and placebo controlled trial of vitamin D, in vitamin D insufficient OA knee subjects (serum 25(OH)D levels <50 nmol/L) was conducted. 150 subjects of primary OA knee were subjected to personal interview to determine their vitamin D intake and the daily sunlight exposure in hours. Serum levels of calcium, phosphorus, alkaline phosphatase and 25(OH) D were measured to determine their vitamin D profile. We identified 64 vitamin D insufficient subjects out of 150 primary OA knee patients. These were randomized by random allocation table for intervention. In cases a bolus dose of calciferol in 60,000 IU/day for 10 days followed by 6,000 IU/month for six months was administered and in controls a placebo in the same schedules and durations was given. Primary outcome measures were clinical WOMAC scores and VAS for knee pain. Secondary outcome measures were radiological features and KL grades. The serum levels of vitamin D were assessed by using Enzyme Linked Immunosorbent Assay (ELISA) and calcium/phosphorus/alkaline phosphatase by UV end point method. Results: There was no significant baseline difference of age, sex, analgesic frequency, dietary vitamin D intake, serum levels of vitamin D, calcium, phosphorus and alkaline phosphatase and in clinical and radiological scores between the cases and controls. BMI (25.96 vs 25.65, p=0.75) and pain (10.94 vs 10.64, p=0.66) was higher in the placebo group although difference was not statistically significant. There was no significant difference in radiological features and KL grades from baseline and at 6 months in both the groups. At six month, both the groups had an improvement in WOMAC and VAS pain scores but vitamin D showed benefit over placebo from baseline (p<0.01); for WOMAC physical function vitamin D group showed significant improvement over placebo which remained same as their baseline levels. Conclusions: Although a long term study is being recommended to establish radiological progression, if any, this short term randomized placebo control trial yields a beneficial effect of vitamin D in pain and physical function in KOA. Vitamin D intake was beneficial in symptomatic improvement of KOA.

765W

Use of the amniotic fluid transcriptome to identify novel antenatal treatments for fetuses affected with Down syndrome. *F. Guedj[1], L. Massingham[1], K. Johnson[1], U. Tantravahi[2], D. Bianchi[1].* 1) Mother Inftant Research Institute, Tufts Medical Center, Boston, Massachusetts, USA; 2) Women and Infants Hospital Department of Pathology Providence, Rhode Island, USA.

Background: Using cell-free fetal RNA obtained from amniotic fluid supernatant, we previously identified statistically significant gene expression differences in second trimester fetuses with Down syndrome (DS) compared to euploid gestational age matched controls (PNAS 2009;106: 9425). Oxidative stress was identified as a major functional abnormality in DS fetuses. Here we applied Connectivity Map data (www.broadinstitute.org/cmap) to suggest small molecules that might “treat” the transcriptome alterations displayed by aneuploid fetuses. **Methods:** We evaluated the protective effect of the top three antioxidant molecules identified by the Connectivity Map: apigenin, celastrol and dimethyloxalylglycine (DMOG) on amniocytes obtained from euploid fetuses (n=5) and fetuses with a known diagnosis of DS (n=4). Cells were cultured and incubated in the absence or presence of different concentrations of antioxidants for three consecutive days before experimental assessment. Preliminary experiments focused on identifying optimal non-toxic concentrations of each molecule. The COMET assay, a fluorescent single cell electrophoresis test, was used as an objective measure of oxidative stress through the measurement of DNA in the comet tail. **Results:** We initially determined the most suitable antioxidant non-toxic concentrations for apigenin, celastrol and DMOG to be 1–4 μM, 50–200 nM and 50–100 μM, respectively. Analysis of 200–300 cells/cell line/condition revealed that: 1) Amniocytes from DS fetuses have higher oxidative stress (% DNA in tail = 14.12 ± 5.17 %) compared to euploid (% DNA in tail = 5.13 ± 2.50 %) (p<0.001). 2) Treatment of DS amniocytes with increasing concentrations of apigenin (1, 2 and 4 μM) significantly reduced the % of DNA in the comet tail by 52.35 % (p<0.05), 64.2 % (p<0.001) and 80.6 % (p<0.0001) respectively. **Conclusions:** These in vitro studies provide proof of principle that amniotic fluid transcriptome data can be used to suggest molecules for fetal treatment in aneuploidy. The results show that apigenin significantly reduced the oxidative stress observed in amniocytes from fetuses with DS. This antioxidant effect is now being evaluated for celastrol and DMOG using the same experimental approach. These molecules will be administered in vivo in wild-type pregnant mice carrying Ts1Cje pups to test the hypothesis that antenatal treatment of oxidative stress improves neurocognition in DS.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

766W

Treatment of basal cell carcinomas with a hedgehog pathway inhibitor vismodegib in Basal Cell Nevus Syndrome. *A. Treherne[1,2], S. Bale[3], W. Chen[4], J.R. Toro[5]*. 1) Dermatology Department, Veterans Affairs Medical Center, Washington, DC; 2) Dermatology Department, Howard University College of Medicine, Washington, D.C; 3) GeneDx Gaithersburg, M.D; 4) Pathology Department, Veterans Affairs Medical Center, Washington, D.C; 5) Genetic Epidemiology Branch, Division of Cancer Epidemiology and Genetics, NCI, Bethesda, M.D.

A 52-year-old caucasian man with a history of basal cell nevus syndrome (BCNS) presented for evaluation and treatment of new skin lesions. He has a history of multiple basal cell carcinomas (BCC) with the first one appearing on the corner of his right eye at age 19 followed by 100s of BCCs distributed over his trunk and extremities. Previous treatments (excision, Mohs surgery, currettage and electrodessication, cryotherapy, 5-aminolevulinic acid photodynamic therapy, imiquimod and 5-fluorouracil) only provided temporary relief followed with the development of numerous new BCCs. Patient reported a history of a seizure four years ago with transient loss of consciousness and bladder control. He denied history of visual changes, jaw pain, swelling or cysts. Dermatologic examination revealed frontal bossing, milia like-papules on his eyelids, 2–3 mm erythematous palmo-plantar pits, more than 100 erythematous to pearly papules and plaques scattered over the face, neck, trunk and upper extremities, and a 1 cm firm dermal nodule on the right vertex of the scalp. Dermoid cysts at the web space of fingers were not noted. MRI without and with IV contrast of the brain showed, no evidence of a primary neoplasm, calcification of the falx cerebri or agenesis of the corpus callosum. Microscopic examination of lesional skin from the scalp and various skin sites revealed BCCs. Genetic counseling was performed. Genetic testing using targeted CGH array analysis with exon-level resolution showed a duplication of exon 3 of the PTCH 1 gene confirming the diagnosis of BCNS. The patient was started on vismodegib (erivedge) 150 mg orally once daily. Three months after initiation of therapy with vismodegib, no new BCCs and decrease in size of BCCs by 51–59% percent on his left bicep and left upper jaw were noted. He denied common side effects reported associated with vismodegib including muscle spasms, alopecia, dysgeusia, weight loss, fatigue, nausea, diarrhea, decreased appetite, constipation, arthralgias, vomiting, and ageusia. He was advised to use condoms with spermicide, during sexual intercourse with female partner while taking vismodegib and for at least 2 months after the last dose. In addition, patient was recommended not to donate blood or blood products while taking vismodegib and for at least 7 months after the last dose. To date follow-up response to treatment will be presented.

767W

TGFβ signaling exerts dimorphic effects on aortic aneurysm formation and progression in Marfan syndrome. *L. Carta[1], J.R. Cook[1], M. del Solar[1], E. Chiu[1], N. Clayton[2], C. Nelson[2], B. Wentworth[2], F. Ramirez[1]*. 1) Pharmacology and Systems Therapeutics, Mount Sinai School of Medicine, New York, NY; 2) Genzyme Corporation, Framingham, MA.

TGFβ antagonism, using either the pan-TGFβ neutralizing antibody (TGFβNab) 1D11 or the angiotensin type I receptor (AT1r) blocker Losartan, has been shown to prevent formation of thoracic aortic aneurysm (TAA) in $Fbn1^{C1039G/+}$ mice, an animal model of MFS featuring non-lethal TAA progression. By contrast, Losartan treatment has been reported to delay TAA dissection and rupture in $Fbn1^{mgR/mgR}$ mice, which recapitulate the lethal phenotype of early onset, progressively severe MFS. Our ongoing investigations have confirmed this last finding by showing aortic growth normalization but significant medial degeneration in $Fbn1^{mgR/mgR}$ mice that are also deficient for the AT1 receptor. We furthermore found that neonatal administration of the TGFβNab 1D11 promotes TAA dissection and rupture in $Fbn1^{mgR/mgR}$ mice earlier than in untreated mutant mice, and that 1D11 administration after completion of postnatal vessel growth extends the average survival of $Fbn1^{mgR/mgR}$ mice to the same extent as Losartan treatment. These findings are therefore consistent with the notion that TGFβ signaling plays a protective role during early TAA formation and that factors other than promiscuous TGFβ activity trigger the onset of aortic disease in MFS. This view also implies that anti-TGFβ therapy, albeit effective in blunting aortic growth during adulthood, may be counterproductive in treating young MFS patients.

768W

Safety of Velaglucerase Alafa in Type 1 Gaucher Disease Patients with Anti-Imiglucerase Antibodies. *O. Goker-Alpan[1], GM. Pastores[2], B. Rosenbloom[3], J. Ibrahim[4], GM. Cohn[5], D. Zahrieh[5], PK. Mistry[6]*. 1) Lysosomal Disorders Research and Treatment Unit, Center for Clinical Trials, O&O Alpan, Fairfax, VA; 2) New York University School of Medicine, New York, NY; 3) Tower Hematology Oncology, Beverly Hills, CA; 4) St Joseph's Children's Hospital, Paterson, NJ; 5) Shire HGT, Lexington, MA; 6) Yale University School of Medicine, New Haven, CT.

PURPOSE To assess the safety of velaglucerase alfa in type 1 Gaucher disease (GD1) patients with anti-imiglucerase antibodies at baseline (BL) in the multicenter treatment protocol HGT-GCB-058 (which gave patients the option of velaglucerase alfa treatment during imiglucerase supply disruptions). METHODS Treatment-naïve patients received 60 U/kg velaglucerase alfa every other week (EOW); patients transitioning from imiglucerase (switch) received the same dose of velaglucerase alfa as their prior imiglucerase dose (15–60 U/kg EOW). Study participation ended upon transition to commercial therapy. Patients were tested for anti-imiglucerase and anti-velaglucerase alfa antibodies at BL and anti-velaglucerase alfa antibodies at 13, 25, 37, 51 weeks. Treatment-emergent adverse events (AE) were monitored. RESULTS 37/211 patients enrolled in HGT-GCB-058 were positive for anti-imiglucerase antibodies at BL (all switch patients; median age/range: 59/9–79 years; 17 females; 6 splenectomized). The median (lower-upper quartile) exposure time to velaglucerase alfa was 22.1 weeks (14.1–28.1). 13/37 experienced ≥1 AE (most mild or moderate in severity), 8 experienced ≥1 drug-related AE (mild-moderate), 6 experienced ≥1 infusion-related AE (mild), and 1 discontinued due to blood pressure increase (mild). 12/37 anti-imiglucerase antibody-positive patients were also anti-velaglucerase alfa antibody positive (cross-reactive) at BL: 1 subsequently tested negative for anti-velaglucerase alfa antibodies during the study (no reported AE), 2 discontinued before the first in-study antibody assessment (no reported AE), and 9 remained positive for both antibodies during the study. Only 1 patient developed treatment-emergent anti-velaglucerase alfa antibodies; he was anti-imiglucerase antibody positive at BL (no reported AE). Of the 10/37 patients who had positive in-study anti-velaglucerase alfa antibody tests, 4 experienced ≥1 AE (mild-moderate), 3 experienced ≥1 drug-related AE (mild-moderate) and 2 experienced ≥1 infusion-related AE (mild); none discontinued due to an AE. No antibody-positive patient experienced a serious AE. 1 patient positive for both antibodies at BL and who remained anti-velaglucerase alfa antibody positive throughout was also IgE positive for both antibodies. No hypersensitivity reactions were reported. CONCLUSIONS Velaglucerase alfa was generally well tolerated in GD1 patients with BL anti-imiglucerase antibodies who transitioned from imiglucerase.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

769W

Safety and Tolerability of Velaglucerase Alfa Enzyme Replacement Therapy in Patients Aaged ≥65 Years with Type 1 Gaucher Disease: A Multicenter, Open Label Trial. *J. Ibrahim[1], GM. Pastores[2], B. Rosenbloom[3], O. Goker-Alpan[4], GM. Cohn[5], D. Zahrieh[5], PK. Mistry[6].* 1) Genetics Division, St Joseph's Regional Med Ctr, Paterson, NJ; 2) New York University School of Medicine, New York, NY; 3) Tower Hematology Oncology, Beverly Hills, CA; 4) LSD Research and Treatment Unit, Springfield, VA; 5) Shire HGT, Lexington, MA; 6) Yale University School of Medicine, New Haven, CT.

PURPOSE The most common lysosomal storage disorder is Type 1 Gaucher disease (GD1), an inherited deficiency of lysosomal glucocerebrosidase. The FDA approved a treatment protocol in 2009 to offer velaglucerase alfa to GD1 patients affected by the imiglucerase supply disruption. The objective of the treatment protocol was to observe the safety of velaglucerase alfa in GD1 patients ≥ 2 years of age, treatment-naive or switched from imiglucerase, with no prior anaphylactic reaction to enzyme replacement therapy. METHODS This was a subset analysis of safety data in an elderly population being switched from prior imiglucerase therapy from the US multicenter treatment protocol. Switch patients received the same dose as their prior imiglucerase regimen (15–60 U/kg every other week [EOW]). RESULTS 52 of 211 patients who initiated velaglucerase alfa were ≥65 years of age. In this older subgroup, the median age was 68.5 (range, 65–89) years, 26 (50%) were women, 21 (40.4%) were splenectomized, and all had been switched from imiglucerase treatment. At baseline, the median hemoglobin concentration was 13.2 g/dL (range, 10.4–16.6 g/dL; n=51) and the median platelet count was 161.0 × 109/L (range, 47.0–474.0 × 109/L; n=49). Patients could discontinue the protocol when commercial therapy was available to them; median treatment duration was 197 (range, 1–475) days. 23 (44.4%) patients experienced ≥1 adverse events (AEs). Drug-related AEs were observed in 6 (11.5%) patients. A total of 3 serious AEs (cerebrovascular accident, diverticulitis, and bacterial peritonitis) were reported in 2 (3.8%) patients; all were considered unrelated to study drug. 5 severe AEs were reported in 3 (5.8%) patients (cerebrovascular accident; upper abdominal pain and leukopenia; diverticulitis and bacterial peritonitis); all were considered unrelated to study drug and all affected patients continued treatment with no change in dose. A mild infusion-related AE (blood pressure increase) caused study discontinuation in 1 (1.9%) patient. Mean hemoglobin and platelet levels remained stable throughout treatment. CONCLUSIONS These previously treated GD1 patients ≥65 years old were successfully transitioned to velaglucerase alfa, which was generally well tolerated and had a safety profile consistent with that reported previously in children and younger adults.

770W

A Double Blind, Placebo Controlled Trial of Minocycline in Children with Fragile X Syndrome. *R. Hagerman[1], M.J. Leigh[1], D.V. Nguyen[2], T.I. Winarni[3], S. Rivera[4], a. schneider[1], t. chechi[5], d. hessl[6].* 1) Dept Pediatrics, MIND Inst, Univ California Davis Med Ctr, Sacramento, CA; 2) Department of Biostatistics UC Davis, Davis, California; 3) Diponogoro University, Samarang, Java, Indonesia; 4) Department of Psychology, Center for MIND and Brain, UC Davis, Davis California; 5) MIND Institute, UC Davis Medical Center, Sacramento California; 6) Department of Psychiatry, MIND institute, UC Davis Medical Center, Sacramento California.

Background: Fragile X syndrome (FXS) is the most common inherited cause of intellectual disability and the most common known single gene cause of autism. Minocycline has been found to decrease levels of matrix metalloproteinase 9 and normalized synaptic connections and behavior in the knockout mouse model of FXS. Objective/Hypothesis: To determine the efficacy and tolerability of minocycline as a targeted treatment for children and adolescents with FXS. Design/Methods: Randomized, double-blind, placebo-controlled, crossover trial lasting for 6 months. Patients (n=55, 87.3% male, 78% white) were recruited from a sample of individuals with FXS, ages 3.5 to 16 years of age, mean age 9.0+/-3.5 years. Outcome measures including the Clinical Global Impressions-Improvement (CGI-I) scale, Visual Analogue Scale (VAS) for behaviors, and the Aberrant Behavior Checklist (ABC) were administered at baseline, 3 months and 6 months. Results: Sixty-nine subjects were screened and 66 were randomized into two treatment arms. Fifty-five subjects (83.3%) completed at least the first period and 48 (72.7%) completed the full 6 month trial. Intention-to-treat analysis demonstrated statistically greater improvements in CGI-I after treatment with minocycline compared to placebo (2.49 ±0.13, 2.97 ±0.13, respectively, p 0.0173) and greater improvement in anxiety and mood-related behaviors on the VAS (minocycline 5.26 cm ±0.46 cm, placebo 4.05 cm±0.46cm; p 0.0488). No significant treatment effects were seen in the secondary outcome measures. Side effects were not significantly different during the minocycline and placebo treatment periods, with the most common side effect being gastrointestinal problems such as loose stools. No serious adverse events occurred during minocycline treatment. Five out of 26 participants (23.8%) had positive ANA levels post treatment, but baseline values were not available and patients were asymptomatic. Conclusions: Treatment with minocycline for three months in children and adolescents with FXS resulted in greater global improvement. Treatment for three months appears safe; however, longer trials are indicated to further assess benefits, side effects, and factors associated with a clinical response to minocycline.

771W

Enhancing suppression of nonsense mutations in Duchenne muscular dystrophy. *P.S. Lai, R. Elangovan, S.K.H. Tay, P.S. Low.* Dept Pediatrics, National Univ Singapore, Singapore, Singapore.

Translational readthrough of premature stop codons (PTCs) arising from pathogenic nonsense mutations represents a potential approach for therapy in recessive genetic disorders. Compounds such as aminoglycosides have been shown to restore full length proteins through suppression of PTCs but they are commonly associated with severe side effects of toxicity or if used at subtoxic doses, result in reduced suppression of PTCs. We report a study demonstrating that co-administration of an adjunct compound can increase aminoglycoside-induced readthroughs of PTCs using a mammalian HEK293T cell model transfected with reporter constructs carrying three types of stop codons cloned from Duchenne muscular dystrophy patients. Readthroughs were assayed using the Dual Luciferase Reporter Assay System. Statistical analysis was carried out using independent group T-test and two-tailed ANOVA T-test. Comparative studies using five different aminoglycoside compounds (gentamycin, paramomycin, tobramycin, amikacin, ribostamycin) at concentrations ranging from 0.5–10 μg/ml show highest readthroughs induced by ribostamcyin (41%). Interestingly, the latter is a 3-ring compound in contrast to the rest and could have a structural advantage in more effective binding to the 30s ribosomal sub-unit during readthrough induction. Paramomycin and tobramycin was not as effective as gentamicin and amikacin, with highest readthroughs between 6.18 to 13.38%. As expected, there was a difference in readthrough susceptibility among the three codons: UAG>UAA>UGA. Co-treatment with a polyanion, poly-L-aspartic acid (PAA) led to 25% increase in readthrough levels after 24 hours. PAA was observed to be effective when used at a ratio of 1:10 of the concentration of each aminoglycoside. Our data also shows a time-lag for optimal effect of PAA with the increase in readthrough being even more significant on Day 5 (120 hours) after treatment. The maximum increase achieved was 49% for ribostamycin. The results suggest that PAA could be used to reduce the dose of aminoglycosides required to achieve clinically therapeutic threshold levels of readthoughs and attenuate the effects of toxicity commonly associated with long-term use of these compounds. The ability of PAA to enhance suppression of the PTCs is likely due to its binding with cationic aminglycosides to form intracellular complexes resulting in higher cytoplasmic levels of aminoglcosides accessible for ribosome binding and mediating readthroughs.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

772W

Long-term safety of migalastat HCl in patients with Fabry disease. *D.P. Germain[1], R. Giugliani[2], G.M. Pastores[3], K. Nicholls[4], S. Shankar[5], R. Schiffmann[6], D. Hughes[7], A.B. Mehta[7], S. Waldek[8], A. Jovanovic[8], K. Benistan[1], J.K. Simosky[9], V. Sniukiene[9], R. Winkler[9], P.F. Boudes[9]*. 1) Division of Medical Genetics, University of Versailles, Garches, France; 2) Med Genet Serv/HCPA, Dep Genetics/UFRGS and INAGEMP, Porto Alegre, RS, Brazil; 3) New York University, New York, NY; 4) Royal Melbourne Hospital, Parkville, Australia; 5) Emory University, Atlanta, GA; 6) Baylor Research Institute, Dallas, TX; 7) Royal Free Hospital, London, UK; 8) Salford Royal Hospital, Manchester, UK; 9) Amicus Therapeutics, Cranbury, NJ, USA.

Background: Fabry disease (FD) results from an inborn error of glycosphingolipid metabolism in which α-galactosidase A (α Gal A) deficiency leads to accumulation of its substrates, primarily globotriaosylceramide (GL-3). While FD is an X-linked lysosomal storage disorder, both males and heterozygous females are affected. If untreated, many patients will develop serious end-organ complications of kidney failure, cardiac arrhythmias, and/or stroke. Methods: The safety of oral migalastat (AT1001/GR181413A) has been evaluated in Study FAB-CL-205 (NCT00526071), an ongoing, open-label, uncontrolled study in 23 patients with FD who completed a preceding phase 2 study. The study primary objective was to evaluate the long-term safety and tolerability of migalastat. Initial patients (15/23) received 150 mg every other day (QOD). All patients participated in a dose escalation period (DEP), in which they received migalastat 250 mg for 3 days on/4 days off for 2 months before escalating to a dose of 500 mg for 3 days on/4 days off. As the 500 mg dose regimen had no clear additional benefit in increasing α Gal A activity or decreasing urine GL-3, the dose was returned to 150 mg QOD. Data is presented up to a cut-off date of April 2012. Results: The enrolled subjects (14 males, 9 females) had a median age of 42 years. Median treatment duration, including the preceding study, was 5 years (medians: 1.4 years at 500 mg, 1.7 years at 150 mg QOD post-DEP). There were no deaths. Seven patients reported 22 serious adverse events (SAEs), all unrelated to migalastat. Treatment was prematurely discontinued by 6 patients: 1 male for SAEs unrelated to treatment (ventricular fibrillation and stroke); 1 female for withdrawal of consent, and 4 males for lack of response (3/4 with a non-amenable mutation, as defined retrospectively by an in vitro cell-based assay). Two patients had a dose reduction from 500 mg migalastat due to AEs (Patient 1 reported muscle twitching; Patient 2 reported confusion, dizzy episode, intention tremor, arthralgia, and myalgia). All 23 patients reported at least one treatment emergent adverse event (TEAE) and 12 patients reported 35 TEAEs considered related to treatment. The most common TEAEs were arthralgia (9 subjects), fatigue (8), back pain (7), and pain in extremity (6). Conclusion: Migalastat doses of 250 mg 3 days on/4 days off and 150 mg QOD were well tolerated. The dose of 150 mg QOD is being used in phase 3 studies.

773W

Eliglustat, an Investigational Oral Therapy for Gaucher Disease Type 1 (GD1): Phase 2 Results After 4 Years of Treatment. *M.J. Peterschmitt[1], E. Lukina[2], N. Watman[3], M. Dragosky[4], G.M. Pastores[5], E. Avila Arreguin[6], H. Rosenbaum[7], A. Zimran[8], E. Sysoeva[2], R. Aguzzi[1], L.H. Ross[1], A.C. Puga[1]*. 1) Genzyme, a Sanofi company, Cambridge, MA; 2) Hematology Research Center, Moscow, Russia; 3) Hospital Ramos Mejia, Buenos Aires, Argentina; 4) IMAI-Research, Buenos Aires, Argentina; 5) New York University, New York, NY; 6) Instituto Mexicano del Seguro Social Hospital de Especialidades, Col. La Raza, Mexico; 7) Rambam Medical Center, Haifa, Israel; 8) Sha'are Zedek Medical Center, Jerusalem, Israel.

Background: In Gaucher disease type 1 (GD1), deficient acid β-glucosidase activity causes glucosylceramide accumulation primarily in tissue macrophages (Gaucher cells), resulting in multisystemic manifestations (thrombocytopenia, anemia, hepatosplenomegaly, bone disease). A variable clinical course and progression led to the development of therapeutic goals for disease manifestations (Pastores, 2004). Eliglustat, a potent and specific inhibitor of glucosylceramide synthase, is under development as an oral substrate reduction therapy for GD1. **Aim**: Report long-term efficacy and safety results. **Methods**: This ongoing, open-label, uncontrolled, multicenter Phase 2 clinical trial enrolled 26 adults with GD1 who had splenomegaly with thrombocytopenia and/or anemia and were not on treatment for the previous 12 months. Efficacy outcomes, assessed periodically, included changes from baseline in hemoglobin, platelets, spleen and liver volumes, skeletal manifestations, biomarker levels and achievement of therapeutic goals. **Results**: Nineteen patients completed 4 years of eliglustat treatment; no patient discontinued in the last 2 years. After 4 years, mean hemoglobin level and platelet count increased 2.3±1.5 g/dL (11.3±1.5 g/dL to 13.6±1.2 g/dL) and 95% (68,700±21,200/mm^3 to 125,400±51,100/mm^3), respectively; mean spleen and liver volumes (multiples of normal, MN) decreased 63% (17.3±9.5 to 6.1±3.4 MN) and 28% (1.7±0.4 MN to 1.2±0.3 MN), respectively. All patients met ≥3 of 4 long-term therapeutic goals (spleen, 100% of patients; liver, 94%; hemoglobin, 100%; platelets, 50%). Median chitotriosidase and CCL-18 each decreased 82%; plasma GL-1 and GM3 normalized. Mean lumbar spine bone mineral density increased 0.8 T-score (-1.6±1.1 to -0.88±1.3), with the greatest increases in patients with osteoporosis at baseline. Femur dark marrow, believed to reflect Gaucher cell infiltration into bone marrow, was reduced or stable in 17/18 patients. No bone crises were reported. Eliglustat was well-tolerated. Most adverse events (AEs) were mild and unrelated to treatment. Ten drug-related AEs, all mild, occurred in 8 patients. No new serious AEs were reported in any patient between 3 and 4 years of treatment. **Conclusions**: Eliglustat continues to show promising efficacy and safety, with clinically meaningful improvements across disease parameters. Results from 2 controlled Phase 3 trials in untreated and enzyme replacement therapy maintenance patients will be available in 2013.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

774W

Ammonia (NH3) Control in Children Ages 2 Months through 5 Years with Urea Cycle Disorders (UCDs); Comparison of Sodium Phenylbutyrate and Glycerol Phenylbutyrate. *W. Smith[1], G. Diaz[2], U. Lichter-Konecki[3], S.A. Berry[4], C.O. Harding[5], S.E. McCandless[6], C. LeMons[7], K. K Dickinson[8], D. Coakley[8], T. Moors[8], M. Mokhtarani[8], B.F. Scharschmidt[8], B. Lee[9].* 1) Dept Pediatrics, Div Gen, Maine Med Ctr, Portland, ME; 2) Mount Sinai School of Medicine, Department of Genetics and Genomic Sciences, Department of Pediatrics, 1428 Madison Avenue, New York, NY 10029; 3) Children's National Medical Center, Division of Genetics and Metabolism, 111 Michigan Avenue NW #1950, Washington, DC 20010; 4) University of Minnesota, 420 Delaware St. SE, MMC 75, Minneapolis, MN 55455; 5) Oregon Health Sciences University, 3181 SW Sam Jackson Park Road, A036/B198, Mailcode L103, Portland, OR 97239; 6) Case Western Reserve University, 11100 Euclid Avenue, Cleveland, OH 44106; 7) National Urea Cycle Disorders Foundation, 75 South Grand Avenue, Pasadena, CA 91105; 8) Hyperion Therapeutics, 601 Gateway Blvd, Suite 200, South San Francisco, CA 94080; 9) Baylor Baylor College of Medicine, One Baylor Plaza, Room R814, Houston, TX 77030.

Ammonia and pharmacokinetic (PK) profiles were examined in a study of glycerol phenylbutyrate (GPB or HPN-100), an investigational drug being developed for UCDs, vs. sodium phenylbutyrate (NaPBA) in children with UCDs. Design This was an open label switch-over study of UCD patients ages 29 days to under 6 yrs on a stable dose of NaPBA powder. On Day 1 patients underwent a 24-hr (0, 8, 12 and 24 hr) sample collection on NaPBA, were switched to a PBA-equimolar dose of GPB, underwent another 24-hr sampling on Day 10 and were offered enrollment in a 12-month extension study with GPB treatment. Plasma and urine samples were assayed for ammonia and drug metabolites. Patients/parents completed questionnaires regarding signs and symptoms associated with NaPBA use or UCD at the time of enrollment and again after treatment with GPB. Results. 15 patients (8 ASL, 3 ASS, 3 OTC, 1 ARG) ages 2 mo to 5 years were enrolled. Patients received a mean NaPBA dose of 5.27g/day (353 mg/kg/d; range 55.6, 583.9 mg/kg/d) and switched to a mean GPB dose of 5.16 g/day or 346.26 g/kg/d; range 61.1, 578.1 mg/kg/d). Plasma ammonia values were lower on GPB treatment at all time points and assessed as AUC0-24 (ratio of GPB/NaPBA means 0.79; 95% CI 0.593–1.055; p 0.03 Wilcoxon signed-rank test; 0.07 t-test) or peak ammonia values (39.39 vs. 52.74 μmol/L; p 0.14). The upper 95% confidence interval of 1.055 was less than the predefined margin of 1.25 achieving non-inferiority. Six patients experienced AEs on GPB; all mild with no SAEs, discontinuations, or significant lab changes. Eleven of 15 patients enrolled in the switch over reported 35 symptoms associated with NaPBA use or UCD on Day 1 (e.g. body odor, protein intolerance, lethargy, vomiting, refusal to eat, abdominal pain). Twenty-three of these 35 symptoms either had improved or resolved on treatment with GPB by Day 10, particularly body odor and recurrent vomiting. Mean systemic exposure (AUC0-24) of PBA, PAA and PAGN following GPB administration were similar to NaPBA. Urinary PAGN concentration and PAGN/creatinine ratio varied less over 24 hours and were greater with GPB treatment. All 15 patients enrolled in the 12-month safety extension study, along with an additional 8 patients who did not participate in the switch over part. Conclusions. GPB results in more evenly distributed urinary output of PAGN over 24 hours and offers ammonia control which is at least as good as that observed with NaPBA in young children with UCD.

775W

A comprehensive genetic and functional analysis of a large cohort of patients with Kabuki syndrome: suggestions for therapy. *L. Micale[1], C. Fusco[1], B. Augello[1], M.N. Loviglio[1,2], C. Maffeo[1], M.T. Pellico[1], L. Zelante[1], G. Merla[1].* 1) IRCCS, Casa Sollievo della Sofferenza, San Giovanni Rotondo, Foggia, Italy; 2) Center for Integrative Genomics, University of Lausanne, Lausanne, Switzerland.

Kabuki syndrome (KS) is a multiple congenital anomalies syndrome characterized by a peculiar face and varying degrees of mental retardation with mutations in the histone methyl transferase MLL2 as the underlying cause. MLL2 is involved in embryogenesis and development, acting as transcriptional activators of multiple genes. Recently three KS individuals with partial and complete deletion of KDM6A gene, encoding a histone demethylase interacting with MLL2, were reported. We performed sequence analysis of all 54 exons of the MLL2 gene in 250 individuals with KS, identifying causative MLL2 mutations in approximately 75 % of them. In line with other reports, the majority of them carry protein truncating mutations (nonsense or frameshift). Splice site mutations were also detected and some of them characterized. MLL2 mutation-negative KS patients were subsequently tested for exon (s) deletions/duplications in KDM6A and MLL2 genes, but no causative variants were found. These data indicate that MLL2 mutation is the main cause of KS and that genomic rearrangements of MLL2 and KDM6A are extremely rare in KS, substantiating its genetic heterogeneity and the search for additional causative gene(s). Moreover we investigated the functional consequences of some MLL2 mutations. We found that a number of MLL2 nonsense mutations result in mRNA degradation through the nonsense-mediated mRNA decay, leading to protein haploinsufficiency in KS-lymphoblastoid cell lines (KS-LCLs). Further we measured by qPCR the mRNA levels of MLL2 targeted HOX genes in KS-LCL, detecting a significant reduction of their expression level compared to control cells. These data support the hypothesis that the KS may benefit from a read-through therapy aimed to restore the physiologically levels of MLL2 protein. Therefore we are performing a proof-of-principle study on a selected number of KS-LCLs with MLL2 nonsense mutations by using specific compounds that mediate translational read-through of nonsense mutations and thereby stimulate the expression of full-length functional proteins. An overview of our recent insights in molecular and functional features of KS will be presented.

776W

Differential effects of ramipril and losartan therapy on phenotype in mice model for Marfan syndrome. *B.L. Lima, G.P. Cardozo, G.R. Fernandes, L.V. Pereira.* Dept Gen, Inst de Biociencias - USP, São Paulo, Brazil.

The Marfan syndrome (MFS) is an autosomal dominant disease of connective tissue, which affects 1 in 5,000 individuals. The main clinical manifestations include aneurysms and aortic disruption, excessive growth of bones, scoliosis and thoracic deformities. Mutations in the FBN1 gene, which encodes the fibrillin-1 protein, were genetically linked to the MFS, classifying this disease in the fibrilinopathies group. The complex pathogenesis of MFS also involves changes in TGFβ signaling, increased matrix metalloproteinase (MMP) expression, tissue fragmentation, and abnormal matrix-cell interactions. During the last years, therapeutic strategies, based on the use of drugs, have been studied in animal models and patients with MFS. Among these drugs we find Losartan, an antagonist of AT1, which may restore the vascular and pulmonary phenotype in animal models. Clinical trials involving patients with MFS revealed that, even with the patients presenting improvements in vascular alterations, the results are far from those seen in animal models. The hypothesis is that the dose given to patients is not sufficient to reproduce the effects observed in animals, which received very high drug dose. During the year of 2011 our group compared the effect of Ramipril, an ACE inhibitor, with the effects of Losartan, on the phenotypic manifestations of SMF in the murine model mg$\Delta^{loxPneo}$. The treatment with Ramipril was realized with doses equivalent to those already used in humans (5 mg/kg). At 6 months, animals treated with Ramipril, showed an increase of approximately 35% in the Fbn1 gene transcription, and a similar improvement in pulmonary and vascular manifestations, when compared with animals treated with Losartan. However, these animals also presented an impressive improvement on bone manifestations compared to Losartan treated mice (P<0,01), which has never been reported. Here we describe a most detailed effect of Ramipril and Losartan over the expression of genes involved with SMF manifestations, and the effect of both drugs over canonical (Smad) and non-canonical (ERK1/2 and JNK) TGF-β intracellular signaling pathways.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

777W

Curcumin facilitates a transitory cellular stress response in Trembler-J mice. *Y. Okamoto[1], K. Mehrdad[1], D. Pehlivan[1], W. Wiszniewski[1], G. Snipes[2], J. Lupski[1,3,4].* 1) Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX; 2) Department of Pathology, Baylor College of Medicine, One Baylor Plaza, Houston, TX; 3) Department of Pediatrics, Baylor College of Medicine, One Baylor Plaza, Houston, TX; 4) Texas Children's Hospital, 6621 Fannin Street Houston, Texas 77030.

Misfolded proteins can cause neurodegenerative and genetic diseases due to molecular stress and interference with the cell function. Many myelin gene mutants (e.g. MPZ and PMP22) cause severe neuropathy evidently by protein accumulation within the endoplasmic reticulum (ER), causing Schwann cell apoptosis, and subsequently peripheral neuropathy. We previously showed that oral administration of curcumin significantly decreases the percentage of apoptotic Schwann cells and partially mitigates the severe neuropathy phenotype of the Trembler-J (Tr-J) mouse model in a dose-dependent manner [Khajazvi, et al (2007) Am J Hum Genet 81:438–453]. To identify specific molecular and cellular pathways through which curcumin functions in Tr-J mice, we performed comparative expression profiling in sciatic nerves of curcumin treated and untreated Tr-J mice. Our results showed that UPR gene markers such as Bip, Ddit3 (chop), and other ER-stress sensors are up-regulated in Tr-J sciatic nerves suggesting a persistent ER stress in Tr-J mice.. We also found that administering curcumin induces the expression of multiple heat shock proteins including Hsp70 while reducing ER stress sensors in sciatic nerves. We further tested if Hsp70 levels could influence the severity of the Tr-J neuropathy. Notably, reduced dosage of the Hsp70, strongly potentiates the severity of the Tr-J neuropathy, though the absence of Hsp70 had little effect in wild type mice. Our findings provide corroborative evidence for the importance of ER associated degradation (ERAD) pathways in the pathogenesis of peripheral neuropathy disorders. In aggregate, these data provide further insights into the pathological disease mechanisms caused by myelin gene mutations and open new avenues for the induction of heat shock protein chaperones as a therapeutic approach for ER-retained mutants.

778W

Novel therapeutic assays for Epidermolysis Bullosa Simplex. *J. Lacroix[1], T. Farez[1], G. Leclerc[2], A. Dupérée[2], J. Powell[3], C. McCuaig[3], C. Morin[2], V. Legendre-Guillemin[1], M. Bchetnia[1], C. Laprise[1].* 1) Université du Québec à Chicoutimi, Chicoutimi, Québec, Canada; 2) Centre de santé et de services sociaux de Chicoutimi, Chicoutimi, Québec, Canada; 3) Hôpital Sainte-Justine, Montréal, Québec, Canada.

Background: Epidermolysis bullosa simplex (EBS) is a genodermatosis with a prevalence of 1/50,000 to 1/30,000 throughout the world. EBS is caused by mutations in the keratin (*KRT*) *5* and *KRT14*, leading to aggregations of keratin filaments, and is characterized by blistering of the skin. There is still no therapy available to cure this disorder but some molecules are known to reduce keratin aggregates by, among other ways, increasing the expression of other keratins. **Objective**: To assay the therapeutic potential of three molecules: trimethylamine N-oxide (TMAO), sodium 4-phenylbutyrate (4-PBA) and sulforaphane (SF). **Methods**: Two immortalized keratinocyte cell lines (1 control and 1 patient with 373C>A mutation on *KRT14*) were first pre-treated with the molecules and then subjected to heat shock. Afterwards, immunofluorescence for keratin staining and cells count were performed to determine the percentage of reduction of cells with aggregates. **Results**: The number of cells with aggregates is not significantly different after 4-PBA (p=0,13) pre-treatment. The number of cells with aggregates was significantly reduced after TMAO pre-treatment (p=0,03) but this molecule was found to be toxic for the cells. There were significantly less cells with aggregates after the SF and heat shock treatment (p=0,01). **Perspective**: We were able to establish *in vitro* the protection of SF against thermal damage in keratinocytes, which could translate *in vivo* to potential therapy in EBS with *KTR14* mutations.

779W

Sodium Butyrate and Valproic acid as a splicing restoring agents in erythroid cells of β-thalassaemic patients. *M. Taghizadeh[1], S. Mansoori Derakhsham[2], M. Shekari Khaniani[2], A. Hoseinpour Pheysi[3].* 1) Tarbiat Modares University, Tehran, Iran; 2) Tabriz University of Medical Sciences, Tabriz, Iran; 3) Hematology Oncology Research Center, Tabriz University of Medical Sciences, Tabriz, Iran.

Beta (β) Thallassemia is a common genetic disease which affects thousands of hundreds of people around the world. More than 400 mutations play role in Beta Thallassemia and the most common mutation in Thallassemia is RNA splicing mutation. The findings suggest that the butyrate and Valproic acid are good drug candidate for Thallassemia, but no research has been carried out on its effect on splicing mutations. In 2001 Chang et al reported the first compound butyrate drug which could considerably restore splicing defect in spinal mascular atrophy. In Cystic Fibrosis, Na-Butyrate (NaBu) increase expression of SR and SR-like proteins, which are splicing factors that increase exon splicing and restores aberrantly spliced mRNA. The purpose of this study is using pharmacologic elements to study the amount of β-globin transcript in Thallassemia patients with IVSI-110 mutation. Bone Marrow aspiration was carried out from two Thallasemic patients with homozygous form of IVSI-110 Mutation and Some 3X106 Erytroid progenitor cell of Bone marrow cultures derived from Thallassemic patients were transferred into dishes using medium composed of Dulbecco's modified Eagle's medium (DMEM), FBS, BSA, β -mercaptoethanol, glutamine, penicillin-streptomycin, EPO, SCF and dexamethosone. After tree Days NaBu, V.A and IsoBu was added dropwise for a final concentration of 0.5, 100, o.5 mM, respectively. Semi-quantitative RT-PCR analysis of β-Globin mRNA from human Erytroid progenitor cell of Bone marrow cultures treated with SBu, V.A and IsoBu was carried and PCR products were visualized on a 1% Agarose gel by ethidium bromide staining. Each experiment was repeated at least tree times with different passages of the respective cell line.The amount of Wild-type of β-Globin mRNAs in Erytroid progenitor cells cultured with NaBu, V.A and IsoBu was increased 1.7, 1.5, 1.4 fold respectively relative to normal β-Globin mRNAs in untreated Erytroid progenitor cells. This study suggests that restoring transcript in IVSI110 patients by increase of SR proteins and also the role of NaBu and V.A in increase of β-globin transcript by histone deacetylase inhibitors and increase of H3 and H4 histone acetylation which make histone proteins release from DNA, can be effective. Increase in β-globin transcript in patients of IVSI110, achieved in cells treated by NaBu and V.A can cause decrease in transforming of blood and advancement of the disease in the patients.

780W

Bisphosphonate therapy in Hajdu-Cheney syndrome: response, suggested protocol and future directions in treatment. *M.D. Irving[1,3], W.M. Drake[2], J. Pittaway[2], D. Dafou[3], M.A. Simpson[3], R.T. Trembath[3].* 1) Department of Clinical Genetics, Guy's and St Thomas' NHS Foundation Trust, London, London, United Kingdom; 2) Department of Endocrinology, Barts Health, London, United Kingdom; 3) Division of Genetics and Molecular Medicine, King's College London School of Medicine, London, United Kingdom.

Hajdu-Cheney syndrome (HCS) (OMIM 102500) comprises short stature, distinctive craniofacial abnormalities, childhood-onset of progressive bone loss, both generalised (osteoporosis) and focal (acro-osteolysis), irregular dentition, congenital anomalies, neurological complications and characteristic radiographic findings. Management is mainly conservative and focussed upon preserving bone mineralisation. Surgical intervention is employed to correct the secondary mechanical effects of reduced bone density, such as bathrocephaly with basilar invagination. The mainstay of managing clinically significant reduction in bone mineral density is through the use of bisphosphonates, a group of pharmacological agents with structural similarity to pyrophosphate and that preferentially binds to calcium, inhibiting activation of osteoclast enzymes and decreasing bone resorption. There is scant documentation of the use of bisphosphonate therapy, but it seems to have a useful clinical effect in improving bone mineral density to reduce fracture risk, either as monotherapy or in combination with other therapeutic and supplementary compounds, though acro-osteolysis appears unresponsive to therapy and continues to progress. We report a case series of 10 HCS patients and their response to bisphosphonate therapy, determined by markers of bone turnover and bone mineral density, suggest a standardised protocol of administration and discuss potential new therapeutic strategies for treating HCS using knowledge of the underlying molecular mechanism.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

781W

A small chemical, TG003, enhances skipping of mutated dystrophin exons: the third example revealing a decrease of exonic splicing enhancer density in common. *A. Nishida[1,2], Y. Takeshima[2], N. Kataoka[3], M. Yagi[2], H. Awano[2], T. Lee[2], K. Iijima[2], M. Hagiwara[4], M. Matsuo[1,2].* 1) Medical Rehabilitation, Kobegakuin University, Kobe, Japan; 2) Department of Pediatrics, Kobe University Graduate School of Medicine, Kobe, Japan; 3) Medical Innovation Center, Laboratory for Malignancy Control Research, Kyoto University Graduate School of Medicine, Kyoto, Japan; 4) Department of Anatomy and Developmental Biology, Kyoto University Graduate School of Medicine, Kyoto, Japan.

Duchenne muscular dystrophy (DMD) is a progressive and fatal muscle-wasting disease. DMD is characterized by a deficiency of dystrophin protein in the muscles, which is caused mostly by out-of-frame deletion or nonsense mutations in the dystrophin gene. Antisense oligonucleotide-mediated exon skipping, which produces in-frame dystrophin mRNA, is recognized as the most plausible way to express semi-functional dystrophin in DMD. In our previous report, we showed that TG003, a small chemical kinase inhibitor, enhanced the skipping of two mutated dystrophin exons and produced in-frame dystrophin mRNA (Nishida et al., Nature Commun 2011). Here, we examined TG003-mediated skipping of six different exons encoding a nonsense mutation and identified the third mutated exon that showed TG003-mediated enhancement of exon skipping. Twelve hybrid minigenes containing wild-type or mutated exons were subjected to in vitro splicing in HeLa cells. Although splicing of five mutated exons was not affected by the treatment with TG003, splicing of mutated exon 39 was modulated to produce exon 39 skipping. The ratio of exon 39-skipped product to the total spliced product increased depending on the dose of TG003. Furthermore, TG003 promoted the exon39 skipping in the endogenous dystrophin gene in a dose-dependent manner and increased production of dystrophin protein in the patient's cultured muscle cells. This is the third identified mutated exon that is sensitive to TG003-mediated enhancement of exon skipping. We analyzed characteristics of splicing regulatory mechanisms of three TG003 sensitive exons. RNA-binding protein motifs, as analyzed by SpliceAid, failed to reveal common protein motif changes among the three skipped exons. Decrease in exonic splicing enhancer density by acquisition of the mutation was the sole common change. In particular, a decrease in exonic splicing enhancer density obtained by the program PESX was unique to the three exons, and could discriminate these from the five TG003-insensitive exons.

782W

Bezafibrate as treatment option in patients with mitochondrial complex I deficiency. *B.M. Haberberger[1,2], P. Freisinger[3], V. Strecker[4], M. Steger[4], H. Heide[4], B.F. Müller[5], T. Beckhaus[5], K. Heim[2], U. Ahting[6], B. Rolinski[6], J. Mayr[7], A. Rötig[8], W. Sperl[7], M. Zeviani[9], I. Wittig[4], T. Meitinger[1,2], H. Prokisch[1,2].* 1) Institute of Human Genetics, Technical University Munich, Germany; 2) Institute of Human Genetics, Helmholtz Zentrum München, Neuherberg, Germany; 3) Department of Paediatrics, Klinikum Reutlingen, Germany; 4) Molekulare Bioenergetik, Zentrum der Biologischen Chemie, Goethe-University Frankfurt, Germany; 5) Institute for Pharmaceutical Chemistry, Cluster of Excellence "Macromolecular Complexes", Goethe-University Frankfurt, Germany; 6) Department of Clinical Chemistry, Städtisches Klinikum München GmbH, Munich, Germany; 7) Department of Pediatrics, Paracelsus Medical University Salzburg, Salzburg, Austria; 8) Institut National de la Santé et de la Recherche Médicale (INSERM), Paris, France; 9) Unit of Child Neurology, Neurological Institute `Carlo Besta'-IRCCS Foundation, Milan, Italy.

Faulty energy supply due to defective oxidative phosphorylation is the biochemical signature of the genetically heterogeneous group of mitochondrial disorders. A majority of them affect the respiratory chain complexes and the cellular ATP production. In spite of rapid progress in finding the molecular cause, in most instances no curative therapeutic options are available. Bastin et al. (2008) and Wenz at al. (2011) provided evidence from in vitro studies and mouse models that activation of the PPAR/PGC-1-alpha pathway with bezafibrate could be a new therapeutic approach. Djouadi et al. (2010) showed in a clinical study, that bezafibrate increased the resting activity of β-oxidation enzymes in patients with inherited β-oxidation disorders leading to a significant improvement in the condition of the patients. In order to verify this effect, we collected 32 fibroblast cell lines from patients with isolated complex I (CI) deficiency and a defined molecular diagnosis. The CI activity in fibroblasts ranged from 10% to 80% in comparison to control cell lines. After bezafibrate treatment we found no improvement in cell lines with mtDNA mutation (n=4); but in all cell lines with nuclear mutations an increase (15–140%) of CI activity was found. The largest effect was observed in a homogeneous group of missense mutations in the same gene, ACAD9. Analysis of genome-wide expression levels, confirmed increased expression of genes involved in lipid and fatty acid metabolism as well as transport, but provided no evidence of upregulation of genes coding for respiratory chain complex subunits (n=10). Raised CI activity was combined by a larger amount of CI assembled in supercomplexes visualized by 2D-BN/SDS-PAGE experiments (n=8). We further examined the consequences of bezafibrate treatment on protein level to get a clearer picture of the involved mechanisms. Enriched mitochondrial fractions from untreated and bezafibrate treated controls and patients with mutations in ACAD9 gene were analyzed by 2D-DIGE (n=5) and significant changes ($p<0.05$) were identified by MALDI-MS and ESI-MS. We detected reduced levels of proteins involved in metabolic pathways and stability of complexes compared to healthy controls. Proteomics data confirmed increased expression of proteins involved in lipid and fatty acid metabolism upon bezafibrate treatments. These results support bezafibrate as a promising treatment option for specific subgroups of patients with CI deficiency.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

783W

FMF patients homozygous for M694V have incomplete response to colchicine. *H. Yonath[1, 2, 6], M. Lidar[3,4,6], N. Shechter[6], F. Sikron[5], S. Sadetzki[5,6], P. Langevitz[3,4,6], A. Livneh[3,4,6], E. Pras[1,6]*. 1) Inst Human Gen, Sheba Med Ctr, Ramat Gan, Israel; 2) Department of Medicine A, Sheba Med Ctr; 3) Department of Medicine F, Sheba Med Ctr; 4) The Heller Institute of Medical Sciences, Sheba Med Ctr; 5) 5Cancer and Radiation Epidemiology Unit, Sheba Med Ctr and Gertner Institute for Epidemiology and Health Policy Research, Tel Hashomer; 6) Sackler Faculty of Medicine, Tel Aviv University, Israel.

Background: Prophylactic colchicine treatment is used in patients suffering from Familial Mediterranean fever (FMF). It is known, from previous studies, that 65% of the patients show a complete response, 30% a partial response and about 5% show minimal or no response, the cause of the difference was not known. We have observed many patients that are homozygotes for M694V in *MEFV* who do not respond well to colchicine despite being treated with maximal sustained doses. Genotyping enables us to study the response rates in different patients according to specific mutations. **Aim**: To assess the response rates to colchicine in M694V homozygous FMF patients in comparison to other prevalent genotypes. **Methods**: A telephonic survey was done, in 112 FMF patients with known gentotypes: 40 M694V homozygotes, and 2 comparison groups of 41 M694V/V726A compound heterozygotes and 31 V726A homozygotes. The questionnaire included demographic, social and clinical features, colchicine dose, response rates and reported side effects. **Results**: M694 homozygous patients had a more severe disease, and were treated with higher doses of colchicine (average dose 1.98±0.56 compared to 1.47±0.58, p=0.0001 and 1.13±0.41, p<0.001 in the M694V/V726A compound heterozygotes and the V726A homozygotes, respectively); Colchicine related side effects were noted in 40% of the M694V homozygotes. The average rate of attacks in treated M694V homozygotes (0.70±1.06) was higher compared to the two other groups (0.14±0.26, p=0.002 and 0.08±0.20, p=0.0009, respectively) and only 25% of them reported no attacks in the last year. None of the patients who took part in this study had amyloidosis. Side effects limiting the dose of colchicine were noted in 40% of the M694V homozygotes. **Conclusions**: Patients that are homozygous for M694Vhave only partial response to colchicine, even in higher doses than other genotypes. The results highlight the need for additional treatment modalities for these patients.

784W

CNS administration of recombinant human tripeptidyl peptidase-1 attenuates neurodegeneration, improves function, and increases lifespan in a canine model of classical late infantile neuronal ceroid lipofuscinosis. *B. Vuillemenot[1], M. Katz[2], D. Kennedy[1], J. Coates[2], C. Sibigtroth[2], R. Whiting[2], F. Wininger[2], R. Reed[3], E. Adams[3], M. Butt[4], R. Cahayag[1], S. Keve[1], L. Tsuruda[1], D. Musson[1], C. O'Neill[1]*. 1) BioMarin Pharmaceutical Inc., Novato, CA; 2) University of Missouri, Columbia, MO; 3) Northern Biomedical Research, Muskegon, MI; 4) Tox Path Specialists, Frederick, MD.

cLINCL is a lysosomal storage disorder caused by deficiency of the enzyme TPP1. Patients exhibit progressive neurodegeneration and loss of motor, cognitive, and visual function, leading to death during mid-childhood. TPP1-null dachshunds recapitulate the progression of the human disease. This study was conducted to characterize the pharmacologic effects of rhTPP1 administration to the CNS of TPP1-null Dachshunds. An additional objective was to determine the pharmacokinetic (PK) profile and CNS distribution. TPP1-null dogs and wild-type controls received 4 or 16 mg rhTPP1 or artificial cerebrospinal fluid (aCSF) vehicle (N=3/group). rh-TPP1 or vehicle was administered via an intracerebroventricular catheter, or if that catheter lost patency, via an intrathecal lumbar catheter every 2 weeks from 9 weeks of age to euthanasia. In some dogs both catheters lost patency during the course of treatment, in which cases enzyme replacement was continued via intrathecal injections into the cisterna magnum. Elevated CSF TPP1 concentrations were observed for 48 hours after administration and were approximately 1000-fold higher than plasma levels. Appearance of neurodegenerative clinical signs was delayed by 3.5 to 11 weeks in the 4 mg treated animals and 10 to 28 weeks in those treated with 16 mg rhTPP1. Pupillary light reflexes were restored to near wild-type levels in animals that received rhTPP1. Cognitive function, as assessed by reversal learning, was also normalized after rhTPP1 treatment. Quantitative magnetic resonance imaging demonstrated a decrease in ventricular enlargement, indicating preservation of brain morphology after rhTPP1 treatment. Mean survival was increased from 43 weeks in the vehicle treated controls to 54 and 62 weeks in animals administered 4 or 16 mg rhTPP1, respectively. Two of the three animals treated at the 16 mg level were euthanized prior to reaching end-stage disease due to complications unrelated to disease progression. This study justifies the administration of rhTPP1 enzyme replacement therapy directly to the CNS for cLINCL patients. Analyses of CNS tissues for rhTPP1 concentration, lysosomal storage material, and neuronal injury are ongoing.

785W

Intrathecal enzyme replacement therapy improves diffusion tensor imaging abnormalities and volume loss in the corpus callosum in canine mucopolysaccharidosis I. *P. Dickson[1], S. Chen[2], I. Nestrasil[3], V. Kovac[3], J. Jens[4], S.Q. Le[1], S.-h. Kan[1], M.A. Guzman[5], J. Provenzale[2,6], E.G. Shapiro[3], C.H. Vite[7], A. Mlikotic[1], N.M. Ellinwood[4]*. 1) LA BioMed Harbor-UCLA Med Ctr, Torrance, CA; 2) Duke University Department of Radiology, Durham, NC, USA; 3) University of Minnesota Department of Pediatrics, Minneapolis, MN, USA; 4) Iowa State University Department of Animal Science, Ames, IA, USA; 5) St. Louis University Department of Pathology, St. Louis, MO, USA; 6) Emory University Department of Radiology, Oncology and Biomedical Engineering, Atlanta, GA, USA; 7) University of Pennsylvania, School of Veterinary Medicine, Philadelphia, PA, USA.

{U}Introduction:{/U} Intrathecal enzyme replacement therapy (IT ERT) was previously shown to improve lysosomal storage in the brain in canine mucopolysaccharidosis I (MPS I). However, translation to the bedside will require clinically-relevant outcome measures. We performed diffusion tensor imaging (DTI) on 9 MPS I dogs, 4 unaffected heterozygous dogs, and 4 dogs treated with IT ERT from 4–21 months of age. Our goals were to determine 1) whether DTI studies in the brain are abnormal in the canine model, 2) whether these respond to IT ERT, and 3) what their underlying basis might be. {U}Methods:{/U} IT ERT-treated dogs received 0.058 mg/kg body weight recombinant human iduronidase diluted in artificial cerebrospinal fluid (total injection volume 2–5 ml). Enzyme was administered every 3 months into the cisterna magna. Diffusion-weighted images were acquired in the axial plane on a 3T GE Discovery MR750 using an 8-channel high-definition transmit/receive knee coil in 25 directions, number of excitations=4, b-value= 1000, matrix=128x128, thickness=1.5 mm. Eigenvalues were measured by tractography following b0 correction. Dogs were imaged twice and the results averaged. Volume was measured on T1-weighted images by manual tracing using BRAINS2. Brains were collected 48h after the final IT ERT dose and within 2 weeks of imaging. {U}Results:{/U} MPS I-affected dogs showed abnormal DTI in the corpus callosum. Fractional anisotropy was reduced by 20% in the genu (p=0.004) and 30% in the splenium (p=0.008) of affected dogs. Radial diffusivity was increased 29% in genu and 62% in splenium of MPS I dogs, suggesting abnormal myelin. Apparent diffusion coefficients and axial diffusivity were also increased. IT ERT-treated MPS I dogs showed improvements in all DTI parameters vs. untreated MPS I dogs, though these did not reach statistical significance in this small sample. Volume of the corpus callosum were diminished in MPS I dogs (0.182±0.0299 cm3, p= 0.02 vs. carriers 0.248±0.0331) and preserved with IT ERT (0.263±0.0493, p=0.006 vs. untreated MPS I). Myelin thickness appeared to be reduced in MPS I dogs and qualitatively improved with early IT ERT. Studies still underway include characterization of myelin lipid composition and levels of myelin-related transcripts and proteins. {U}Conclusion:{/U} MPS I dogs show abnormal DTI, reduced volume and abnormal myelin in the corpus callosum. IT ERT fully or partially prevented these abnormalities.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

786W

Enzyme replacement therapy using chemically modified enzyme markedly enhances therapeutic efficacies for bone lesions in mucopolysaccharidosis VII murine model. *S. Tomatsu[1], D. Rowan[2], J. Grubb[3], B. Haupt[4], A. Montaño[5]*. 1) Biomedical Research, Alfred I. duPont Hospital for Children, Wilmington, DE; 2) School of Medicine, Saint Louis University, St. Louis, Missouri, USA; 3) 3Edward A. Doisy Department of Biochemistry and Molecular Biology, Saint Louis University, St. Louis, Missouri, USA; 4) Department of Pathology, Saint Louis University, St. Louis, Missouri, USA; 5) Department of Pediatrics, Saint Louis University, St. Louis, Missouri, USA.

Mucopolysaccharidosis (MPS) type VII is a lysosomal storage disease caused by deficiency of the lysosomal enzyme β-glucuronidase (GUS), which is involved in degradation of glycosaminoglycans (GAGs). Glycosaminoglycan storage in the tissue of patients with MPS VII causes a wide spectrum of clinical manifestations that affect the central nervous system (CNS), bone, and viscera. Enzyme replacement therapy (ERT) effectively clears GAG storage in the viscera. However, clearance of storage material in CNS has been limited by the blood brain barrier. Nonetheless, recent studies showed that a chemically modified form of GUS (PerT-GUS), which escaped clearance by mannose 6-phosphate and mannose receptors, and showed markedly prolonged circulation, reduced CNS storage more effectively than native GUS. Although clearance of storage in bone is limited by the avascularity of the growth plate, prior work showed clinical improvement of skeletal disease in MPS VII mice treated from birth with weekly intravenous injections of native GUS [Sands et al., 1994]. To evaluate the effectiveness of long-circulating PerT-GUS in reducing the skeletal pathology, we treated MPS VII mice for 12 weeks beginning at 5 weeks of age with PerT-GUS or native GUS and used micro-CT, radiographs, and quantitative histopathological analysis for assessment of bones. Micro-CT findings showed PerT-GUS treated mice had a significantly lower BMD. Histopathological analysis also showed reduced storage material and a more organized growth plate in PerT-GUS treated mice compared with GUS treated mice. Long term treatment with PerT-GUS was also assessed by treating MPS VII mice for up to 57 weeks from birth. Micro-CT and radiographs demonstrated improvements in bone lesions of legs, ribs, and spine of mice treated with long-term PerT-GUS ERT. Quantitative histopathological assay also showed improvements in GAG storage and morphology of articular and epiphyseal chondrocytes.

787W

A phase 2 multicenter, open label, switch over trial to evaluate the safety and efficacy of Abcertin®(imiglucerase) in patients with type 1 Gaucher disease previously treated with imiglucerase. *H. W. Yoo[1], B. H. Lee[1], J. S. Lee[2], J. M. Ko[3], Y. B. Sohn[4]*. 1) Medical Genetics Center, Department of Pediatrics, Asan Medical Center Children's Hospital, Seoul, Korea; 2) Department of Clinical Genetics, Severance Children's Hospital, Seoul, Korea; 3) Department of Pediatrics, Seoul National University Children's Hospital, Seoul, Korea; 4) Genetics Clinic, Department of Medical Genetics, Ajou University Hospital, Suwon, Korea.

Introduction: The Gaucher disease type is the prototype of LSD where enzyme replacement therapy proved to be effective. This switch over clinical trial was undertaken to evaluate the safety of Abcertin® for patients receiving Imiglucerase therapy for type 1 Gaucher disease. Aim: The primary objective of this study aimed at evaluating the safety. The analyses of efficacy on changes in hemoglobin, platelet, liver, spleen, biomarkers, skeletal status and BMD were the secondary endpoints. Methods: Five Korean type I Gaucher patients previously treated with Imiglucerase have been enrolled. The previous Imiglucerase dose ranged from 15 to 60 U/kg with an every other week dosing regimen. Each patient was required to be on the same Imiglucerase dose for at least six months prior to enrollment. A patient's Abcertin® dose should be the same dose as the previous Imiglucerase dose administered for immediate past six months. Results: No clinically significant adverse event was observed. In addition, no changes in the hemoglobin concentration and platelet counts, liver and spleen volumes, and skeletal status and bone mineral density were found. No patient required a dose adjustment. All the patients did not generate antibodies to Abcertin®. As the efficacy and safety of Abcertin® were demonstrated as similar to those of Imiglucerase, study drug Abcertin® could be used as an alternative therapeutic agent in patients who are treated with Imiglucerase. Discussion: Even though small numbers of patients were recruited, this was the first trial to investigate the safety and efficacy of Abcertin®. Another clinical trial is under way to verify the efficacy and safety in naïve type I Gaucher patients and the data will be available soon.

788W

Results of 12 months use of taliglucerase alfa for Gaucher disease in an unselected cohort of adult patients on an Israeli compassionate use program. *A. Zimran, H. Maayan, G. Altaescu, M. Tiomkin, M. Phillips, D. Elstein.* Gaucher Clinic, Shaare Zedek Medical Center, Jerusalem, Israel.

Due to the world-wide reduced availability of imiglucerase (Cerezyme™, Genzyme Corporation, Cambridge, MA USA) since June 2009, the standard enzyme replacement therapy (ERT) for Gaucher disease at the time, regulatory authorities approved early access programs (EAP) for treatment-naïve patients with taliglucerase alfa (Protalix Biotherapeutics, Carmiel Israel), which was then undergoing Phase III clinical trials..Taliglucerase alfa is an ERT produced in a high-yield plant cell system that is readily up-scalable. The cells are grown in disposable bioreactors and are free from exposure to mammalian components. This report describes results in key Gaucher disease parameters of treatment-naïve patients receiving EAP taliglucerase alfa for 12 months. EAP was approved by institutional authorities. IRB approval was not needed for this retrospective review of de-identified material. All adult patients receiving at least one dose of taliglucerase alfa were included in the safety set; the efficacy set included only patients on 30units/kg/infusion with 12 months' data. Descriptive statistics and the Student's t-test were employed. Twenty-three treatment-naïve patients were dosed (one patient at 60units/kg/infusion was only included in the safety set). Two patients (8.7%) discontinued because of hypersensitivity reactions and no results were reported. Two female patients became pregnant during the study, continuously receiving taliglucerase alfa; both gave birth to healthy, full-term babies with no maternal or neonatal complications. Results: At 12 months, mean change from baseline values (n=19) were: increased hemoglobin level 8.9% and platelet count 38.3%; and mean changes (n=14) in organ index volumes measured by ultrasound: reduction in liver -0.3% and spleen -26% volumes. It should be noted that not all patients had anemia and/or hepatomegaly at baseline. The current cohort broadens the database of 30units/kg/every-other-week of taliglucerase alfa among unselected symptomatic adult patients. These results particularly demonstrate the robust response on the platelets counts also in the patients at the lower dose of 30units/ kg/every-other-week, Safety assessments among treatment-naïve patients are comparable to those reported in the Phase III clinical trial for treatment-naïve patients (Zimran et al, Blood 2011). In May 2012 taliglucerase alfa (Elelyso™) was approved by the FDA- for treating adults with type I Gaucher disease.

789W

Gaucher disease type 2: outcomes of enzyme replacement therapy and impact on quality of life. *C. Prada, R. Hopkin, L. Bailey, A. Lovell, N. Leslie, G. Grabowski.* Dept Pediatrics Genetics, Cincinnati Children's Hosp Med Ctr, Cincinnati, OH.

Background: Gaucher disease (GD) type 2 typically presents between 3 and 6 months of age with retroflexion of the neck, feeding difficulties, poor weight gain, laryngeal stridor, trismus, ocular paresis, hepatomegaly and hypersplenism. Untreated, these children do not speak, sit, or stand. The clinical course is characterized by rapid progression with a mean death age of 9 months. Clinical management consists of palliative and supportive care. The objective of this study was to characterize outcomes of enzyme replacement therapy (ERT) and impact on quality of life in GD type 2 patients. Methods: A retrospective analysis of clinical information was conducted for patients with GD type 2 at Cincinnati Children's Hospital from 1990 to 2010. Results: 8 patients were identified (4 untreated, 4 treated with ERT). The median age at symptom onset was 3 months (range=0 to 7). The median age at diagnosis was 3.5 months (range 1 to 7). The most common genotype was L444P/L444P (n=4), followed by L444P/G202R (n=2), L444P/R257Q (n=1), and L444P/Q414R (n=1). Untreated patients exhibited arching, gasping, and extreme irritability. Treated patients had less irritability, better interaction with caregivers, and more comfortable respirations as well as slower disease progression. No differences in brain imaging were noted between the two groups. Patients on ERT had minimal visceral disease. Untreated infants had poor weight gain and rapidly became cachectic. Patients on ERT had better growth parameters with weights between 3rd and 50thcentiles. All the treated patients survived beyond their first birthday; 3 lived longer than 20 months. Parents perceived an improvement of quality of life in the ERT group; patients appeared more comfortable and reached more milestones; one patient was able to walk. The mean ages at death were 8 months (range 5 to 10 months) and 23 months (range 15 to 35 months) for untreated and ERT groups (p=0.022), respectively. Conclusion: Motor and intellectual development was better in GD type 2 patients treated with ERT suggesting a temporizing effect on the central nervous system. The improvement in visceral disease burden and better growth parameters appear to improve cognitive development. Despite aggressive ERT GD type 2 was lethal. However, improvement in quality of life for the treated patients suggests that benefit of ERT outweighs burden of treatment and should be considered for these children.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

790W

Efficiency of various antisense oligonucleotides for correcting ATM pre-mRNA in a patient with ataxia-telangiectasia with a deep intronic splicing mutation. *S. Cavalieri[1], E. Pozzi[1], R. Cavalli[3], R.A. Gatti[2], A. Brusco[1].* 1) SCDU Medical Genetics, S.Giovanni Battista Hospital, Turin, Italy and Departrment of Genetics, Biology and Biochemistry, University of Torino, Italy; 2) Departments of Human Genetics and Pathology and Laboratory Medicine, UCLA School of Medicine, Los Angeles, CA, USA,; 3) Departments of Pharmaceutical Science and Thecnology, University of Torino, Italy.

Antisense oligonucleotides (AONs) correct the aberrant splicing of a defective ATM mutation by masking the aberrant splice site and restoring normal pre-mRNA splicing, without disturbing the surrounding splicing mechanisms. We found a deep intronic splicing (DIS) mutation (c.1236-405C>T) in an Italian AT patient. The mutation creates a novel donor (3'ss) splice site (Splice Site Predictor, score 0.22 wt > 1.00 mutated) at position -405 of intron 11 and hence activates an alternative acceptor (5'ss) splice site at position -619 of intron 11 (score 0.71). As a consequence, a 212 bp non-coding pseudoexon is inserted in the cDNA between exons 11 and 12, containing a premature stop codon that blocks protein translation. In both a lymphoblastoid cell line and primary fibroblasts from the patient we tested different AON chemistries to evaluate their potential for reversing the mutation. Using antisense morpholino oligonucleotides (AMO), designed to mask the donor splice-site created by the c.1236-405C>T mutation, we partially reverted the aberrant splicing product to a wild-type ATM transcript (26%). To improve the delivery and efficiency of the AON, we designed a structurally modified Vivo-AMO that almost completely abrogated the mutant mRNA and significantly enhanced ATM protein at nearly 50% of wild type levels. We also showed that treated cells had a restored ATM kinase activity, testing SMC1 phosphorylation after gamma-ray exposure. To circumvent the possible toxicity of Vivo-AMO, we tested a 2'-O-Methyl-phosphorotioate modification (2'-OMe), which is reported to be less toxic, with higher target affinity and longer half-life. Preliminary mRNA analysis showed a 50% rescue, even at a lower dose of 2'-OMe. Our data suggest that AONs may provide a promising approach to customized, mutation-based treatment for patients with DIS mutations. Experiments are in progress to test additional AON chemical modifications.

791W

In utero brain-directed AAV gene therapy results in rapid, robust, and specific transduction of mouse choroid plexus epithelia. *M. Haddad, A. Donsante, S.G. Kaler.* NICHD, NIH, Bethesda, MD.

In utero gene transfer represents a powerful tool for investigating the feasibility of therapeutic approaches in early neonatal lethal neurometabolic diseases (Rahim et al., Gene Ther, 2011) and assessing whether a mouse mutant with absence of a gene product may be rescued. Incomplete knowledge exists on the transduction efficiency of AAV pseudotypes administered prenatally. The capacity of prenatally administered AAV to transduce the developing choroid plexuses, specialized epithelial structures that project into the brain's ventricular cavities, is entirely unknown. We recently achieved long-term rescue of a lethal mouse model of Menkes disease via rAAV5-mediated ATP7A gene addition to the choroid plexus (Donsante et al., Molec Ther, 2011). To further characterize mammalian choroid plexus-specific transduction, we performed individual 5 μl brain lateral ventricle injections of 5x109 rAAV5-GFP and rAAV9-GFP in E15 embryos of C57Bl/6J and CD-1 pregnant mice and quantitated GFP expression in choroid plexus from lateral, third and fourth ventricles on E17, P2, and P22. Overall surgical survival was 61%. Sagittal brain sections at E17, P2, and P22 were stained with an anti-GFP antibody and the percentage of cells from the choroid plexus epithelia showing expression was quantitated by densitometry (Image J). GFP expression at E17 ranged from 2 to 7%, at P2 from 11 to 19%, and at P22 from 2 to 15%. By postnatal 22, rAAV9-GFP showed higher transduction in choroid plexus compared to rAAV5-GFP ($P<0.01$, two-tailed t-test). In addition, rAAV9 showed a broader biodistribution and pattern of transduction including neurons and glia in the brain cortex and cerebellum in addition to choroid plexus. Our findings indicate that rAAV5 and rAAV9 administered to the fetal brain at E15 results in rapid (within 48 hrs), robust, and specific transduction of choroid plexus epithelia. This is the first study to show rapid and specific rAAV5 and rAAV9 transduction in the developing choroid plexus, 48 hours post-injection, and in two different mouse strains. Based on the detection of the transgene expression within 48 hrs, rAAV9 may enable rescue of certain prenatal lethal alleles in mouse models of human disease, e.g., the Mottled-dappled mouse model of Menkes disease.

792W

Gene and stem cell treatment for alpha-1 antitrypsin deficiency. *B. Feinerman.* Research, STEMCELLREGENMED, Aventura, FL.

Alpha-1 antitrypsin deficiency is a genetic disorder caused by a defective production of alpha-1 antitrypsin(A1AT) leading to decreased activity in the blood and lungs and deposition of A1AT protein in the liver cells. Severe A1AT deficiency causes COPD or emphysema in adults and in 10% of cases liver disease. Treatment of alpha-1 antitrypsin deficiency consists of extraction of autologous stem cells from the patient's own peripheral blood or bone marrow. A knockdown of the abnormal mutation is done by taking a AAV-1 viral vector that will insert a shRNA Serpina 1 into the autologous stem cells prior to intravenous administration. The next step is done 48 hours later and the normal Serpina 1(Myc-DDK tagged ORF of Homo Sapiens chromosome 2 open reading frame 15(C3orf15)as transfection ready)is inserted with the AAV-1 vector into the autologous stem cells and given intravenously. Following this treatment the alpha-1 antitrypsin blood levels are checked at 1,4, 8 and 12 weeks later. An alternate symbol for Serpina1 is AAT1ALPHA.

793W

Identifying motor neuron transduction efficiencies that are efficacious in SMA mice and achievable by intrathecal delivery in a large animal model. *S. Cheng, M. Passini.* Rare Diseases Science, Genzyme, a Sanofi Company, Framingham, MA.

Spinal muscular atrophy (SMA) is a neuromuscular disease caused by a deficiency in SMN due to mutations in SMN1. A decrease in SMN levels results in motor neuron cell death in the spinal cord that leads to weakness and wasting of the muscles responsible for locomotion, swallowing, breathing and coughing. We previously reported that CNS-targeted AAV-hSMN therapy is highly efficacious in SMA mice. The goals of the current study are to determine the minimum percentage of spinal cord motor neurons that needs to be transduced for efficacy in SMA mice, and demonstrate if this transduction rate can be achieved by intrathecal (IT) delivery in large animal models. We hypothesize that AAV vector dose correlates with motor neuron transduction efficiencies to impact survival in SMA mice. Thus, doses of 5e10, 1e10, and 1e9 genome copies (gc) of scAAV9-hSMN were injected into the CNS of SMA mice. Treatment resulted in median lifespans of 153d (>800% increase compared to saline controls), 70d (>312%), and 18d (>6%), respectively. Grip strength and righting reflex tests in mice administered the two highest doses showed significant and sustained improvements in muscle strength and coordination at 14d and 175d. Analysis of the spinal cord at 14d in mice treated with 5e10, 1e10, or 1e9 gc resulted in motor neuron transduction efficiencies of 30–60%, 10–30%, and <5%, and SMN levels of 70–180%, 30–100%, and 10–20% of WT, respectively. Taken together, these data indicate that reconstituting at least 30% WT levels of SMN in 10–30% motor neurons is sufficient to improve function and survival in SMA mice, thus establishing a benchmark criteria for success in subsequent biodelivery experiments in large animal models. To this end, we performed an IT study in juvenile farm pigs to determine the feasibility of widespread gene delivery in a model whose spinal cord size and anatomy is similar to that of humans. Administration of 3e12 gc of scAAV9-eGFP showed robust vector expression in the ventral horn of the cervical, thoracic, and lumbar regions at 30 days post-injection. Importantly, double IHC staining for eGFP and ChAT showed a motor neuron transduction efficiency of 10–60% in the majority of spinal cord segments. This data confirms that a motor neuron transduction pattern predicted to be efficacious is achievable using a clinically viable delivery method, and justifies the continual development of gene therapy for SMA.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

794W

Differential response to exogenous MeCP2 in mouse models of Rett syndrome. *J. Young, V. Camarena, M. Saez, A. Abrams, K. Bayetti, A. Castle, K. Walz.* Hussman Inst Human Genomics, Univ Miami, Miami, FL.

Rett syndrome (RTT, MIM 312750) is a severe neurodevelopmental disorder that affects patient's ability to communicate, move and behave. It is one of the leading causes of mental retardation and autistic features in females, with a prevalence of 1:10–20,000 females. Clinical features of RTT include psychomotor regression, mental retardation, communication dysfunction, seizures, postural hypotonia, stereotypic hand movements, tremors, autonomic dysfunction and growth failure. The vast majority of RTT cases are caused by de novo mutations in the X-linked gene MECP2, encoding for methyl-CpG binding protein 2 (MeCP2), subject to X chromosome inactivation (XCI). MeCP2 is a nuclear protein that binds to methylated CpGs and regulates gene expression through a diverse set of mechanisms and interacting partners. Over 200 different pathogenic MECP2 mutations have been identified (see RettBASE: IRSF MECP2 Variation Database), with very different levels of incidence. Among the most frequently observed mutations are the missense mutations T158M, R133C, R306C and R106W, accounting for ~23% of genetically identified cases, and the nonsense mutations R255X and R294X that account for 11% of cases. Notably, it has been shown that the RTT-like phenotype exhibited by a Mecp2 null mouse model of the disease could be fully reversed by post-symptomatic activation of MeCP2. The majority of RTT cases, however, is caused by missense and nonsense mutations of MeCP2 rather than by deletions or other type of null mutations. Since only one allele of MeCP2 is expressed in a given cell (due to X chromosome inactivation or hemizygosity), these non-null mutant versions of MeCP2 cannot have autologous dominant negative activities. However, the possibility of mutant versions of MeCP2 exerting dominant negative effects upon purposefully expressed wild type MeCP2 cannot be excluded so far. We here report the use of a non-null mouse model of RTT to determine the phenotypic response of transgenic restoration of MeCP2. We compared the ability of exogenous MeCP2 to prevent the development of RTT-like phenotypes in Mecp2 null versus a MeCP2 truncated mouse models of RTT. Contrary to the transgenic prevention observed in Mecp2 null mice, mice carrying a truncated version of MeCP2 develop RTT-like phenotypic features regardless of the presence of a functional copy of MeCP2. These results suggest that the mutated MeCP2 interferes with the activities of the rescuing MeCP2 protein.

795W

Preclinical Trial of Trauma-induced Osteoarthritis Therapy by Overexpressing Matrix Protein and Reducing Inflammation by Gene-transfer. *M.ZC. Ruan, K. Guse, A. Erez, B.C. Dawson, Y. Chen, B. Lee.* Molecular and Human Genetics, Baylor College of Medicine, Houston, TX.

Osteoarthritis (OA) is a chronic debilitating disease characterized by degeneration of articular cartilage, subchondral bone, and intraatricular inflammation. To date, different treatment approaches have resulted in limited success due to the complex genetics and environmental interactions in the development of the disease. Here, we report novel effective gene transfer as treatment for mice with trauma-induced OA by cartilage matrix preservation and decreased inflammation. *Proteoglycan 4*, (*PRG4*), is one of the few genes known to be associated with OA and its loss of function mutation results in early onset OA in humans presented in Camptodactyly-Arthropathy-Coxa Vara-Pericarditis syndrome (CACP). PRG4 knockout mice also develop early OA. In addition, interleukin 1 receptor antagonist (*Il1ra*) has been shown in several studies to alleviate OA in different models. Hence, we hypothesized that gene transfer of *Prg4* and *Il1ra* to the intra-articular space would be beneficial in OA. To test our hypothesis, we generated a transgenic mouse over-expressing *Prg4* under the cartilage-specific promoter, *collagen 2a1* (*Col2a1*). Using a modified surgical OA model, involving cruciate ligament transection, we recapitulated the post-traumatic human OA in a mouse model. Our results show that PRG4 transgenic mice are protected against the development of OA as compared to WT mice. Moreover, to evaluate the potential synergistic therapeutic effects of PRG4 and IL1Ra, we injected helper-dependent adenoviral vectors expressing PRG4 (HDAd-PRG4) and IL1Ra (HDAd-IL1Ra) separately or in conjunction to wild type mice after induction of post-traumatic OA. Injections of either HDAd-PRG4 or HDAd-IL1Ra, resulted in a protective effect against OA while co-injection of the both had a positive synergistic protective effect. These data suggest that this combinatorial treatment may be effective in prevention and/or treatment of OA. In summary, we identified *PRG4* as a novel target for OA chondroprotection. The successful gene delivery of *PRG4* and *IL1Ra* using HDAd could potentially have therapeutic implications for OA.

796W

Towards lentiviral gene therapy for the treatment of MNGIE. *R. De Coo[1], M. Stok[2], R. Yadak[2], H. Smeets[3], N. Van Til[2], G. Wagemaker[2].* 1) Dept Child Neurology & Neurology, Erasmus MC Rotterdam, Rotterdam, The Netherlands; 2) Dept of Hematology, Erasmus MC, Rotterdam, The Netherlands; 3) Dept of Clinical Genomics, Maastricht University Medical Center, Maastricht, The Netherlands.

Mitochondrial neurogastrointestinal encephalomyopathy (MNGIE) is caused by mutations in the TYMP gene, encoding the enzyme thymidine phosphorylase (TYMP) resulting in elevated levels of the substrates thymidine (Thd) and deoxyuridine (dUrd). Due to the imbalance in the nucleotide pool, the mitochondrial DNA (mtDNA) pool becomes instable and leads to mitochondrial dysfunction. The disease is characterized by progressive gastrointestinal dysmotility, cachexia, ptosis, peripheral neuropathy and leukencephalopathy and may cause death between early to mid-adulthood with an unpredictable start of initial symptoms, but it usually starts between the first and fifth decade. Several attempts have been made to reduce the circulating nucleosides, including hemodialysis, and platelet infusions. Recently, an allogeneic hematopoietic stem cell transplantation (HSCT) trial showed elevated TYMP activity with reduction or complete disappearance of the plasma substrates. The longterm clinical benefit of this treatment needs further evaluation. The risks of transplant related morbidity and mortality due to the conditioning regimen and adverse immune reactions are observed with allogeneic HSCT. The use of autologous HSC, genetically modified to produce normal TYMP, would eliminate those risks, and has, in addition, the intrinsic advantage to induce immune tolerance to the recombinant transgene product. The study entails development of ex vivo gene therapy of the patients' own HSC, transplanted following nonmyeloablative conditioning to achieve partial chimerism with the majority of the hematopoietic system untouched. This would result in perpetual production of TYMP by descendants of the transduced stem cells. Proof-of-principle has been obtained in the disease-specific mouse model by our collaborators. With the use of self-inactivating (SIN) lentiviral (LV) vectors we will deliver an human TYMP gene into HSCs. We will establish an optimized protocol to develop a safe and efficient therapy, with a minimum integrations per HSC. The method will be validated in human patient cells and in a MNGIE mouse model. First results will be demonstrated.

797W

Directing miRNA-regulatory PNAs to breast cancer cells with synthetic targeting peptides. *Y.-Y. Jin[1], C.-P. Chen[1], M.L. Thakur[2,3], E. Wickstrom[1,3].* 1) Biochemistry & Molecular Biology, Thomas Jefferson University, Philadelphia, PA; 2) Radiology, Thomas Jefferson University, Philadelphia, PA; 3) Kimmel Cancer Center, Thomas Jefferson University, Philadelphia, PA.

Breast cancer, one of the leading causes of cancer death worldwide, must be treated by drugs with multiple activities due to its complex nature. microRNAs (miRNAs) are small non-coding RNAs that regulate many cellular activities. Multiple lines of evidence have suggested miRNAs being disregulated in cancer. Overexpressed oncogenic miRNAs, such as miR-21 and miR-17, are being studied intensively as targets for breast cancer therapeutics. miR-21 inhibits the translation of PDCD4, TPM1, maspin and PTEN mRNAs. miR-17 suppresses the translation of anti-proliferative p21 as well as proapoptotic BIM and PTEN mRNAs. Knockdown of elevated microRNAs such as miR-21 and miR-17 by a systemic drug that targets breast cancer cells specifically would offer a novel therapy for disseminated drug-resistant disease. Most breast cancer cells overexpress insulin-like growth factor 1 (IGF1) receptor. We have validated 95±0.79% knockdown of endogenous miR-21 in MCF7 cells and 98±0.51% knockdown of endogenous miR-17 in MDA-MB-231 cells using a complementary locked nucleic acid (LNA) inhibitor. But LNA required lipofection, contraindicated for systemic delivery. We hypothesize that reducing miR-21 or miR-17 activity by delivering a sequence-specific peptide nucleic acid (PNA) knockdown agent conjugated to a C-terminal IGF1 analog to enable receptor-mediated endocytosis into breast cancer cells will reduce cell proliferation, invasion and metastasis. Uptake and functional studies of PNA-IGF1 analog vs. PNA and peptide mismatch controls are underway. The effect of PNA-IGF1 analogs on miRNA activity following treatment will be measured by luciferase reporter vectors containing putative 3'UTR target sites of miR-21 and miR-17. The expression levels of suppressor proteins targeted by miR-21 and miR-17 in breast cancer cells following treatment will be compared to those in cells treated with control sequences. Changes in breast cancer cell proliferation, migration, microfilament organization, and survival will be analyzed in cells treated with PNA-peptides, vs. mismatch controls. Supported by NIH 1 R44 CA136306 and CA148565.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

798W

Effective inhibition of human cytomegalovirus gene expression and lytic replication by DNA-based external guide sequence. *H. Li, X. Jia, L. Sun, T. Zhou, Z. Deng.* Dept. Biotechnology, Jinan University, Guangzhou, Guangdong, China.

External guide sequences (EGSs), which are RNA molecules derived from natural tRNAs, bind to a target mRNA and render the mRNA susceptible to hydrolysis by RNase P, a tRNA processing enzyme. Unlike other strategies, such as antisense oligonucleotides, ribozymes, and RNA interference, the RNase P-based technology is unique because a custom-designed EGS molecule can bind to any complementary mRNA sequence and recruit intracellular RNase P for specific degradation of the target mRNA. In this study, we demonstrate that the DNA-based EGS is effective in blocking gene expression and lytic replication of human cytomegalovirus(HCMV), a leading cause of retinitis-associated blindness and other debilitating conditions such as pneumonia and enteritis among AIDS patients, and mental and behavioral dysfunctions in children that were infected in utero. We constructed DNA-based EGS molecules that target the mRNA encoding HCMV DNA polymerase, and we administered them directly to human foreskin fibroblast cells infected with HCMV. A reduction of 79% in DNA polymerase expression and a reduction of approximately 77% in viral replication were observed in cells treated with a functional EGS. Our study provides direct evidence that DNA-based EGSs are highly effective in inhibiting HCMV gene expression and lytic replication. Exogenous administration of chemically synthesized DNA-based EGSs in inducing RNase P-mediated cleavage represents an approach for inhibiting specific gene expression and for treating human diseases, including HCMV-associated diseases.

799W

Oncogenic *EGFR* allele specific inhibition by RNA interference for cancer therapy lacking adverse effects. *M. Takahashi[1], T. Chiyo[2], T. Okada[2], H. Hohjoh[1].* 1) Dept of Molecular Pharmacology, Nat Inst of Neurosci, NCNP, Kodaira, Tokyo, Japan; 2) Dept of Molecular Therapy, Nat Inst of Neurosci, NCNP, Kodaira, Tokyo, Japan.

Anti-cancer medication without adverse effects is an ideal cancer therapy, and the development of inhibitors specific for oncogene products may lead to the achievement of such a treatment. In this study, we showed that oncogenic *epidermal growth factor receptor* (*EGFR*) allele specific inhibition by RNA interference (RNAi) allowed for suppression of human lung cancer cells carrying the oncogenic alleles, but did not affect normal cells, both *in vitro* and *in vivo*. We designed siRNAs against two recurrent oncogenic EGFR deletions found in non-small cell lung cancer (NSCLC), and selected competent siRNAs that conferred effective allele-specific silencing against the oncogenic alleles by using our original assay system. Oncogenic allele-specific RNAi (ASP-RNAi) using the siRNAs resulted in a significant inhibition of proliferation of cancer cells carrying the target oncogenic alleles and induction of their cell death, even though cancer cells were less sensitive to conventional anti-cancer medicines (gefitinib and erlotinib). We further proved anti-cancer activities of the siRNAs with two types of tumor models (subcutaneous and lung tumors) in mice: local and systemic administrations of the siRNAs clearly suppressed subcutaneous and lung tumors, respectively. We also checked several toxicological and immune parameters in the treated mice. In addition to the above observations, systemic administration of siRNA targeting wild-type *EGFR* to normal mice resulted in a significant elevation of plasma alkaline phosphatase (ALP) level and increase in intestinal apoptosis in the treated mice; these results indicated adverse effects caused by the suppression of wild-type (normal) alleles. Taken together, our current study indicated that ASP-RNAi treatment of mutant *EGFR* may become a promising anticancer treatment lacking adverse effects, and be also available for treatment of cancers resistant to conventional anti-cancer drugs.

800W

Remission induction of spinocerebellar ataxia 6 by systems therapy with medicinal herbs. *T. Okabe.* Dept Integrated Trad Med, Univ Tokoy, Tokyo, Japan.

Spinocerebellar ataxia 6 (SCA6) is an autosomal dominant cerebellar ataxia associated with small polyglutamine-dependent expansions in the alpha 1A-voltage calcium channel. At present, we have no effective therapeutic tools. We report here treatments of six cases of spinocerebellar ataxia 6(SCA6)with mixtures of traditional medicinal herbs. Genetic tests revealed expanded allele of 22~25 CAG repeats at the spinocerebellar ataxia type 6 locus. Head MRI revealed a typical atrophic image in cerebellum. For the systems therapy with medicinal herbs, the differential diagnosis by traditional herbal medicine was made according to the guideline. Mixtures of 18~26 medicinal herbs were given according to the differential diagnosis in each patient. The remedies used for the cases consist of several different ingredients, which have well-established histories of use for treatment of vertigo, tremor, or ataxia and are expected to exert their specific effects. The ingredients include sedative, anticonvulsive, neuroprotective, neuroregulatory and DNA binding herbs. The 100-point semiquantitative International Cooperative Ataxia Rating Scales (ICARS) scores were evaluated before and after the herbal therapy. In 5 of 6 cases, most of the symptoms were significantly improved in 30~ 60 days of herbal treatment. 34~85% reduction were obtained on the 100-point semiquantitative International Cooperative Ataxia Rating Scales (ICARS) scores. While slight improvements in ataxia or speech disturbance were observed, marked reduction in ICARS scores was not attained in a wheelchiared-bound patient. While discontinuation of the herbal therapy caused a relapse, marked symptomatic improvements were obtained again by repeating the herbal treatment. The results imply therapeutic potentials of the medicinal herbs for spinocerebellar ataxia 6. Further extensive investigations are required to clarify the mechanisms by which the remission induction of this genetic disease of CAG repeat expansion mutation, has been attained with the medicinal herbs.

801W

Disease Progression in MPS IIIA (Sanfilippo Syndrome Type A) is Associated with Genotype and Age at Diagnosis. *C.B. Whitley[1], I. Nestrasil[1], K. Delaney[1], R. Cooksley[1], P. Haslett[2], C. Richard[2], E. Shapiro[1].* 1) Gene Therapy Ctr, Univ Minnesota, MMC 446, Minneapolis, MN; 2) Shire Human Genetic Therapies, Lexington, MA.

Sanfilippo syndrome type A (MPS IIIA) is an autosomal recessive disorder resulting from deficiency of N-sulphoglucosamine sulphohydrolase (SGSH) which degrades heparan sulfate. SGSH is localized to 17q25.3 and more than 68 mutations have been described. We hypothesize that genotype, age at diagnosis (a proxy for severity), and age at evaluation are associated with disease progression, measured as decrement of neurocognitive function and grey matter volume over the course of one year. Methods: Longitudinal cognitive and quantitative MRI data were collected for 24 children with documented Sanfilippo A syndrome in an ongoing natural history study (NCT01047306). Volumetric analysis obtained by automated segmentation, and cognitive developmental quotient (DQ) were assessed. Phenotypic severity was categorized by age at diagnosis (before/after age six). SGSH mutations were determined in 24 affected children (including 6 sibling pairs). Results: Within the 38 alleles, 7 novel mutations were revealed for which there was no prior knowledge of the associated phenotypic severity. Patients diagnosed before age 6, with both alleles containing mutations known to be associated with severe phenotype, had larger decrements in both cognitive and grey matter measures over one year in comparison to those with either one known severe mutation and one novel mutation or two novel mutations. Those diagnosed after age six (with mutations known to be associated with attenuated disease or novel mutations) showed little decrement in cognition or grey matter. For the entire group, age at evaluation was also significantly associated with declines in cognition and grey matter volume ($p<.01$) such that younger patients showed larger decrements. Conclusions: We conclude that severe genotype, younger age at evaluation and earlier diagnosis are associated with more rapid disease progression.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

802W

Therapeutic efficacy of D-penicillamine encapsulated alginate/chitosan nanoparticles in rat model of copper toxicity with neurobehavioral impairments. *R. Prasad[1], A. Pal[1], S. Attri[2], B. Thappa[3], R. Vasishta[4].* 1) Biochemistry, PGIMER, Chandigarh, India; 2) APC, PGIMER, Chandigarh, India; 3) Gastroenterology, PGIMER, Chandigarh, India; 4) Histopathology, PGIMER, Chandigarh, India.

Animal models of Wilson's disease (WD) viz. Long evans cinnamon rats, rarely exhibit neurological symptoms impeding the development of novel therapeutic approaches to treat neurological manifestationsin WD patients. The aim of this study was to (1) especially on copper and zinc levels in liver, kidney & brain tissues; expression of hepatic metallothionein-I (MT-I) and Atp7b gene, and MT-III and acetylcholine esterase (AChE) gene in brain, biochemical parameters, and neurobehavioral functions of male Wistar rats, and(2) therapeutic evaluation of orally administered D-penicillamine encapsulated alginate/chitosan nanoparticles for 90 days on Cu intoxicated Wistar rats. Reverse transcription-PCR and Morris water maze test were used for expression and neurobehavioral studies. Copper intoxicated animals showed significantly increased ceruloplasmin, serum & urine copper levels and decreased serum acetyl choline esterase (AChE) activity, increased expression of hepatic MT-I gene with impaired neuromuscular coordination andspatial memory. However, no changes were observed on the expression levels of hepatic Atp7b gene, and MT-III and AChE gene in brain. Cu intoxicated rats revealed asignificantincrease in the liver, brain and kidney tissuescopper content (117, 3.7 and 4 times respectively) and interestingly, increased brain zinc content (4.4 times). Histopathological studies demonstrated copper deposition in the liver and brain tissues and, copper associated proteins in liver tissues of test rats by rhodanine and orcein stains respectively. Nanoparticles based therapy resulted in significant reduction of liver and brain Cu content, serum ceruloplasmin level & increase in serum AChE activity with improvement in neurological functions.In conclusion, chronic copper toxicity may lead to increased copper content in liver & brain, increased hepatic MT-I gene expression, ceruloplasmin levels and, neurobehavioral impairments may be by interfering in acetylcholine modulated neurotransmission; however, D-penicillamine encapsulatednanoparticles may reverseimpairments caused by chronic Cu intoxication to a significant extentin Wistar rats.

803W

High-efficiency gene correction of a Parkinson's disease associated LRRK2 gene variant in patient-derived iPS cells. *B. Schuele[1], J. Laganiere[2], S.K. Mak[1], B.J. Vu[2], M. Vangipuram[1], Y.A. Huang[1], D.E. Paschon[2], O. Macaranas[1], A. Flierl[1], R. Sundararajan[1], P.D. Gregory[2], F.D. Urnov[2], J.W. Langston[1], H.S. Zhang[2].* 1) Parkinson's Institute, Sunnyvale, CA; 2) Sangamo BioSciences, Inc., Richmond, CA.

Correction of disease-causing genetic variants from patient-derived iPS cells offers the opportunity to study such variants in an isogenic setting, to model complex diseases in a cell culture dish, and may also lead to potentially curative autologous cellular transplants. Our goal was to develop a highly efficient genome editing process based on zinc finger nucleases (ZFNs) that would require no enrichment of transfected cells (i.e. no cell sorting or drug-based selection), and that could be scaled to 96-well format without a significant increase in labor. As a model locus, we focused on the most common PD mutation in the LRRK2 gene, a single-base G->A transition that results in a missense substitution of the glycine at position 2019 for a serine. We co-delivered the ZFNs targeted to this location along with a corrective donor construct (aimed at restoring the wt LRRK2 sequence) to patient-derived iPSCs, immediately performed limiting dilution in the absence of selection, and then performed target locus genotyping using nested-PCR. Remarkably, in 64 out of 81 analyzed single-cell-derived clones, an unprecedented 77% of clones carried a ZFN edited allele at the endogenous LRRK2 locus. As expected, given the prevalence of non-homologous end-joining over the homology-directed repair pathways in human cells, a majority (67%) of the clones carried a small insertion or deletion at the target locus, either in homozygous or heterozygous form. Of significant note, 14 clones (19%) carried an endogenous allele of LRRK2 that had been corrected from the mutant to the wild-type form of the gene via homologous recombination with the corrective donor DNA, again without selection for this event. Importantly, all characterized ZFN-modified iPSC clones exhibited a normal karyotype, expressed pluripotency markers (including a demethylated Oct 4 promoter), and remained multipotent as evidenced by their differentiated into all three germ layers. Taken together our data demonstrate that the method described here permits the selection-free, sorting-free isolation of iPSCs carrying novel investigator-specified alleles at an unprecedented frequency.

804W

Inhibition of CTR1 by antisense oligonucleotides in mouse model for Wilson disease reduces copper accumulation and improves liver pathology. *T.R Grossman, P. Lincoln, M. McCaleb.* ISIS PHARMACEUTICALS, Carlsbad, CA.

Wilson disease (WD) is an autosomal recessive genetic disorder of copper accumulation in liver, brain, kidney and cornea. Wilson disease gene ATP7B encodes a transmembrane protein ATPase which functions as a copper-dependent P-type ATPase responsible for incorporation of copper to ceruloplasmin and secreting access of copper to the bile. Clinical manifestations include hepatic disease ranging from mild hepatitis to acute liver failure or cirrhosis and neurological symptoms. WD is fatal if untreated and early recognition by means of clinical, biochemical or genetic examination and initiation of therapy with copper chelators, zinc salts or even liver transplantation in cases of acute and chronic liver failure are essential for favorable outcome. The current treatments are lifelong with severe side effects such as neurological deterioration, hypersensitivity syndrome and bone marrow depression. The high affinity copper transporter CTR1 plays a critical role in copper acquisition in the liver. Our therapeutic objective is to slow the progression of liver damage in WD by reducing the level of copper in the liver by antisense oligonucleotides (ASO) to CTR1. We developed a specific and highly efficacious ASO targeting mouse CTR1 and as a proof of concept we treated toxic milk mice with CTR1 ASO for 6 weeks. The toxic milk mouse is the commonly used mouse model for WD characterized by a gradual accumulation of hepatic copper. Analysis of liver samples from toxic milk mice treated with CTR1 ASO showed 53±1.3% reduction in CTR1 mRNA level of in liver, 45.6±2.1%; in the kidney and 54.1±2.9%; in the intestine which resulted in a dramatic reduction of liver copper level determined by MS ICP (1045±50.5 ppm in untreated compared to 327±68 ppm treated with CTR1 ASO). Moreover, animals treated with CTR1 ASO showed an improved liver pathology when compared with mice treated with control ASO or vehicle only. The hepatocyte karyomegaly, glycogenated nuclei and microvesicular changes were significantly improved with treatment. Based on our results we suggest that inhibition of hepatic CTR1 by ASO could be considered as a novel therapeutic strategy for Wilson disease patients.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

805F

Genome wide identification of human cardiac developmental genes located within regions associated with congenital heart disease. *K.D. Ajbro[1], T.A. Andersen[1], M. Bak[1], J.A. Rosenfeld[2], K. Møllgård[3], S.C. Eliasson[4], R.R. Jakobsen[1], E. Bendsen[5], N. Tommerup[1], L.G. Schaffer[2], K. Lage[4,6,7,8], I.A. Larsen[1]*. 1) Wilhelm Johannsen Centre for Functional Genome Research, Department of Cellular and Molecular Medicine, University of Copenhagen, Copenhagen, Denmark; 2) Signature Genomic Laboratories, Spokane, WA; 3) Developmental Biology Unit, Department of Cellular and Molecular Medicine, University of Copenhagen, Copenhagen, Denmark; 4) Center for Protein Research, University of Copenhagen, Copenhagen, Denmark; 5) Fertility Clinic, Department of Obstetrics and Gynaecology, University Hospital of Odense, Odense, Denmark; 6) Pediatric Surgical Research Laboratory, Massachusetts General Hospital, Boston, MA; 7) Harvard Medical School, Boston, MA; 8) Broad Institute, Boston, MA.

Congenital heart defects (CHD) are the most common congenital malformations in newborns and are a major cause of infant morbidity and mortality. Despite many genetic studies, only a fraction of the genetic mechanisms behind normal and abnormal heart development has been elucidated. To identify novel candidate loci and genes in CHD, we constructed a morbid map of CHD by determination of the overlap between copy number variants (CNVs) identified in patients with CHD. This map was compiled from 896 published and unpublished CNVs. To find connections between dosage sensitive genes in CHD and transcriptional programs during normal human heart development, we generated whole genome expression profiles from a time series of human embryonic and fetal hearts from day 41 to 68 post fertilization. Differentially expressed genes were identified using student's T-test. The set of differentially expressed genes was enriched for genes encoding sarcomeric proteins and genes known to cause heart defects when deleted in mice. The two datasets were analyzed for overlap and we identified 36 candidate genes, which are differentially expressed during heart development and positioned in a candidate locus determined by three or more overlapping CNVs. We examined the spatial and temporal protein expression pattern for selected candidate genes by immunohistochemical analysis of human embryonic hearts, and identified strong and regionally distinct protein expression patterns. In summary our data show, that CHD candidate disease genes can be identified, by combining large datasets of patient CNVs with expression profiles from human developing hearts.

806F

Voltage-gated sodium channels are required for proliferation of embryonic myocardium in zebrafish. *J. Bennett[1], D. Stroud[2], J. Becker[2], D. Roden[1,2,3]*. 1) Center for Human Genetics, Vanderbilt University School of Medicine, Nashville, TN; 2) Department of Medicine, Vanderbilt University School of Medicine, Nashville, TN; 3) Department of Pharmacology, Vanderbilt University School of Medicine, Nashville, TN.

In mice, homozygous deletion of the predominant cardiac sodium channel *Scn5a* results in embryonic lethality at day 10.5, before robust sodium current can be detected. These mice exhibit defects in cardiac morphology, including reduced ventricle size and diminished trabeculation. In zebrafish, we have previously shown that morpholino knockdown of cardiac sodium channel orthologs *scn5Laa* and *scn5Lab* perturbs specification of pre-cardiac mesoderm through reduced *nkx2.5* expression, impairs concentric thickening of the ventricular myocardium and inhibits looping and growth of the embryonic heart. However, it is not known which developmental processes are perturbed by sodium channel knockdown and whether sodium channel knockdown reduces cell number by altering migration of cardiac progenitors into the heart, myocyte proliferation, or both.

We assayed expression of cardiac-specific genes at early stages of development by in situ hybridization. To address secondary differentiation later in development, we conducted studies in Tg(*cmlc2:Kaede*) fish we generated. Primary cardiomyocytes were marked using photoconversion of Kaede protein, and late-differentiating cardiomyocyte (non-photoconverted) were counted. Embryos deficient in the ortholog *scn5Lab* showed defects in primary cardiogenesis specific to loss of *nkx2.5*. Despite this impaired early specification of pre-cardiac mesoderm, embryos treated with anti-*scn5Lab* morpholino showed normal secondary differentiation of cardiomyocytes at the venous and arterial poles between 30 and 48 hpf. However, these embryos had significantly reduced myocardial proliferation (3.6±0.7 vs 0.7±0.4 BrdU-positive cardiomyocytes, n= 6–11) at 48 hpf, resulting in reduced chamber size and thin ventricular walls. The results suggest that while voltage-gated sodium channels are required for primary cardiac specification and later cardiac growth, they are not required for differentiation of the mesenchymal field of cardiac progenitors between 24 and 48 hours post-fertilization. The channel encoded by *scn5Lab*, either through participation in a developmental pathway or as an integral component of the structure and physiology of the cardiomyocyte, is required for normal proliferation of myocardium and growth of the embryonic heart.

807F

Modeling Early Heart Development in Down Syndrome using Sibling hESC Lines. *A. Bosman[1], A. Letourneau[1], M. Zuchelli[2], V. Tohonen[2], O. Hovatta[2], S.E. Antonarakis[1], M.E. Jaconi[1]*. 1) University of Geneva, Geneva, Geneva, Switzerland; 2) Karolinska Instituet, Sweden.

Human embryonic stem cells (hESC) carrying known diseases provide excellent models for examining the cellular consequences of a disease from the earliest time in development. Due to differing genetic backgrounds, however, hESC lines are known to display intrinsic differences in their differentiation capacities and epigenetic patterns. These genetic differences result in significant differences upon differentiation, thus making it difficult to detect small variations between diseased and control lines. Using a hESC sibling model of disease provides a more sensitive approach to detecting small variations due to greater genetic similarity. In this study, we have isolated and characterized a trisomy 21 (T21) hESC model of Down syndrome (DS) using sibling hESC lines as controls. As congenital heart defects (CHD) are the leading cause of morbidity DS, we examined the genetic pathways associated with cardiogenesis to ascertain perturbations in development which may lead to CHD. Upon differentiation, T21-hESC show many significant differences in expression of genes associated with both mesodermal and cardiac development, which is particularly evident with genes associated to the secondary heart field (SHF). Additionally, genes of the T-box transcription factor family were found to be significantly over-expressed in T21-hESC. Some of these perturbations also coincide with known causative genes for CHD observed in the general population. Furthermore, we identified at least one gene located on chromosome 21 which may account for some of these perturbations. Therefore, our work shows for the first time, that T21-hESC and their sibling control lines are a useful model facilitating the identification of differentially expressed genes associated with early cardiogenesis, which may underlie the cause of CHD observed in DS.

808F

***Ofd1* controls dorso-ventral patterning and axoneme elongation during embryonic brain development.** *A. de Angelis[1], A. D'Angelo[1], B. Avallone[2], I. Piscopo[1], R. Tammaro[1], M. Studer[1], B. Franco[1,3]*. 1) Telethon Institute of Genetics and Medicine, Italy; 2) Department of Biological Science, University of Naples “Federico II”, Naples (Italy); 3) Medical Genetics, Department of Pediatrics, Federico II University, Naples, Italy.

Oral-facial-digital type I syndrome (OFDI) is a human X-linked dominant-male-lethal developmental disorder caused by mutations in the OFD1 gene. Similar to other inherited disorders associated to ciliary dysfunction OFD type I patients display neurological abnormalities. We characterized the neuronal phenotype that results from *Ofd1* inactivation in early phases of mouse embryonic development and at post-natal stages. We determined that *Ofd1* plays a crucial role in forebrain development, and, in particular in the control of dorso-ventral patterning and early corticogenesis. We observed abnormal activation of Sonic hedgehog (Shh), a major pathway modulating brain development. Ultrastructural studies demonstrated that early *Ofd1* inactivation results in the absence of ciliary axonemes despite the presence of mature basal bodies that are correctly orientated and docked. *Ofd1* inducible-mediated inactivation at birth does not affect ciliogenesis in the cortex suggesting a developmental stage-dependent role for a basal body protein in ciliogenesis. Moreover, we showed defects in cytoskeletal organization and apical-basal polarity in *Ofd1* mutant embryos, most likely due to lack of ciliary axonemes. Thus, the present study identifies Ofd1 as a developmental disease gene that is critical for forebrain development and ciliogenesis in embryonic life, and indicates that Ofd1 functions after docking and before elaboration of the axoneme *in vivo*.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

809F

Exome sequencing identifies a single mutation in the *CCDC114* gene as the cause for Primary Ciliary Dyskinesia in the Volendam population. *A. Onoufriadis[1], T. Paff[2], D. Anthony[1], A. Shoemark[3], D. Micha[2], B. Kuit[2], M. Schmidts[1], S. Petridi[1], R. Wilson[4], C. Hogg[3], R. Emes[5],. UK10K[6], EMK. Chung[7], G. Pals[2], HM. Mitchison[1].* 1) Molecular Medicine Unit and Birth Defects Research Centre, University College London (UCL) Institute of Child Health, London, United Kingdom; 2) Department of Clinical Genetics, VU University Medical Center, Amsterdam, The Netherlands; 3) Department of Paediatric Respiratory Medicine, Royal Brompton and Harefield NHS Trust, London, UK; 4) Respiratory Medicine, Royal Brompton and Harefield NHS Trust, London, UK; 5) School of Veterinary Medicine and Science, University of Nottingham, Leicestershire, UK; 6) uk10k.org.uk; 7) General and Adolescent Paediatric Unit, University College London (UCL) Institute of Child Health, London, UK.

Primary ciliary dyskinesia (PCD) is heterogeneous genetic disorder arising from ultrastructural defects that cause abnormal function of motile cilia and flagella. The disease manifests with an autosomal recessive mode of inheritance, affecting 1 in every 15,000–30,000 births. Abnormal motility leads to a number of symptoms including chronic respiratory disease, otitis media, subfertility and situs abnormalities. So far, genetic defects in 14 genes have been associated with PCD including *DNAH5*, *DNAH11*, *DNAI1*, *DNAI2*, *DNAL1* and *TXNDC3* that encode subunits of the axonemal outer dynein arm motor structures responsible for ciliary beating. To identify the genetic defect underlying PCD in the Volendam population we whole exome-sequenced two distantly related affected individuals of an extended pedigree as part of the UK10K project. Since the Volendam population is isolated and one affected case was the offspring of a consanguineous marriage, we followed a model of rare autosomal recessive inheritance. Only protein altering variants with an estimated population frequency of <0.01 were considered as candidate pathogenic mutations. This analysis revealed *CCDC114* as the only gene harbouring low-frequency variants that were compatible with recessive inheritance. *CCDC114* is the human homolog of the biflagellate *Chlamydomonas DC2* gene, which encodes an outer dynein arm docking complex subunit. Both affected individuals were homozygous for a c.742G>A substitution (p.A248fs50*), which is predicted to result in the loss of a donor splice site leading to premature termination of the CCDC114 protein product. The same variant was observed segregating in seven additional pedigrees from the Volendam population, affecting a total of 16 patients. Transmission electron microscopy and immunofluorescence of respiratory cilia cross-sections from an affected individual showed loss of outer dynein arms. In summary, our data suggest that a single mutation in the *CCDC114* gene underlies all PCD cases in the Volendam population, providing a high priority novel target for future therapeutic intervention.

810F

ALDH1A3 loss of function causes bilateral anophthalmia and hypoplasia of the optic nerve and optic chiasm. *M. Yahyavi[1], A.S. De Preux Charles[2], T. Xiao[2], T. Bardakjian[3], A. Schneider[3], H. Baier[2], A. Slavotinek[1].* 1) Dept. of Pediatrics, UCSF, San Francisco, CA; 2) Dept. Physiology, UCSF, San Francisco, CA; 3) Division of Genetics, Albert Einstein Medical Center, Philadelphia, PA.

Cellular levels of the major active retinoid, all-trans retinoic acid (atRA), are controlled by a balance between enzymes involved in synthesis, (including the retinaldehyde dehydrogenases ALDH1A1, ALDH1A2 and ALDH1A3), and the enzymes involved in degradation. atRA has long been recognized as critical for the development of several organs, including the eye. Mutations in STRA6, the receptor for vitamin A, in patients with Matthew-Wood syndrome and anophthalmia/microphthalmia (A/M), have previously demonstrated the importance of retinoid metabolism in human eye disease. We used whole genome sequencing (Complete Genomics, Mountain View) to investigate patients with A/M. We studied a female born to first cousin parents who had bilateral anophthalmia with hypoplasia of the optic nerve and optic chiasm. A BAC array had shown no diagnostic chromosome aberrations and mutations in SOX2, OTX2, BMP4, CHX10, PITX2 and FOXE3 had been excluded. Analysis of more than 175 novel sequence variants revealed homozygosity for a nonsense mutation in ALDH1A3 (p.Lys389X) that was verified by Sanger sequencing and that was not present in public databases (Exome Variant Server) or in 120 control chromosomes of diverse ethnicity. Both parents were heterozygotes for the same stop mutation. Aldh1a3 is expressed in the surface ectoderm of murine eye field at E8.75 and is present in the ventral retina and optic nerve at E10.5. Mice that are homozygous null for Aldh1a3 show shortening of the ventral retina, lens rotation and persistence of the primary vitreous body. Inhibition of the Zebrafish orthologue of Aldh1a3 with citral resulted in fish with absence of the ventral retina that could be rescued by RA administration (Marsh-Armstrong 1994). Conversely, provision of atRA to murine embryos early in development can cause A/M (Sulik 1995). We performed antisense morpholino studies in zebrafish to characterize the developmental effects of transcription knock down of Aldh1a3. The treatment group showed a significant reduction in eye size and retinal ganglion cell population and aberrant axonal projections to the tectum were noted. We conclude that loss of function for Aldh1a3 can cause eye defects and disordered optic nerve development in humans and animal model systems.

811F

Differential expression of lung function genes by *in utero* smoke exposure during human lung development. *S. Sharma[1,2], A. Kho[1], K. Haley[2], C. Vyhlidal[3], R. Gaedigk[3], L. Kobzik[4], J.S. Leeder[3], K.G. Tantisira[1,2], S.T. Weiss[1].* 1) Department of Medicine, Channing Laboratory, Brigham and Women's Hospital, Boston, MA; 2) Division of Pulmonary and Critical Care Medicine, Brigham and Women's Hospital, Boston, MA; 3) Children's Mercy Hospital and Clinics, Kansas City, MO; 4) Department of Pathology, Brigham and Women's Hospital, Boston, MA.

Introduction: Lung function is an independent predictor of population morbidity and mortality. In utero smoke exposure (IUS) is associated with reduced lung function resulting in increased susceptibility to chronic respiratory disease. Genome-wide association studies (GWAS) have identified multiple replicated SNPs in 39 genes associated with lung function. However, the effect of *in utero* smoke exposure on the expression of these genes is not yet known. We hypothesized that genes associated with lung function in published GWAS will be differentially expressed due to IUS during human lung development. **Methods**: Genes associated with lung function in previously published GWAS were identified through a literature search. Genome-wide gene expression profiles on 370 human fetal lung tissue samples (post-conceptional age:53–137 days) were generated on the Affymetrix Human Gene 1.0 ST Array. IUS exposure was detected using placental cotinine values, measured with competitive ELISA (Calbiotech) from placenta from the same 370 individuals. Cotinine levels that were undetectable or below the limit of quantitation of the assay (2 ng total cotinine/ml) were classified as unexposed. Detectable cotinine levels >30 ng/mg total protein were classified as having had IUS exposure. Differential gene expression between IUS exposed and unexposed fetal lung tissue was performed using linear regression models adjusted for age. **Results**: In total, 39 genes have been associated and replicated with at least one measure of lung function (FEV1,FVC, or FEV1/FVC) in previously published GWAS studies. Of the 39 genes, 22 (56%) were differentially expressed by IUS exposure in the human developing lung at a genome-wide level of significance (FDR=0.05), which was more than expected by chance alone (p=0.006). Genes that were differentially expressed by IUS were enriched for genes in the regulation of smoothened signaling and dorsal/ventral pattern formation pathways. **Conclusions**: Our results demonstrate that genes associated with lung function in the general population are differentially regulated within the developing lung by smoke exposure *in utero*. Our results suggest that the changes in gene expression that result from *in utero* smoke exposure during lung development may be a biologic basis for their association with abnormal lung function. Further investigation of these biologic mechanisms is being pursued.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

812F

The *C7orf58* locus is strongly associated with both bone mineral content and density at the pediatric distal radius. *B.S. Zemel[1,2], M. Li[3], S. Deliard[4], C.E. Kim[5], L. Qu[3], R.M. Chiavacci[5], J.M. Lappe[6], H.J. Kalkwarf[7], V. Gilsanz[8], H. Hakonarson[2,4,5], S.E. Oberfield[9], J.A. Shepherd[10], S.F.A. Grant[2,4,5]*. 1) Division of Gastroenterology, Hepatology and Nutrition, Children's Hospital of Philadelphia, Philadelphia, PA; 2) Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA; 3) Department of Biostatistics and Epidemiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA; 4) Division of Genetics, Children's Hospital of Philadelphia, Philadelphia, PA; 5) Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, PA; 6) Osteoporosis Research Center, Creighton University, Omaha, NE; 7) Division of General and Community Pediatrics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH; 8) Department of Radiology, Children's Hospital Los Angeles, Los Angeles, CA; 9) Division of Pediatric Endocrinology, Morgan Stanley Children's Hospital of New York, Columbia University Medical Center, New York, NY; 10) Department of Radiology and Bioimaging, University of California San Francisco, San Francisco, CA.

Osteoporosis has its origins in childhood, where bone mineral accretion during growth and development is a critical determinant of future bone health. Failure to achieve optimal bone mineral content and density (BMC and BMD respectively) during the critical period of growth results in suboptimal peak bone mass, contributing to low bone mass and osteoporosis later in life. A recently reported meta-analysis of genome wide association studies of adult BMD revealed 64 single nucleotide polymorphisms (SNPs) at 56 loci were significantly associated with the trait. We elected to leverage data from an ongoing genome wide association study of pediatric bone accretion (NIH R01 HD058886) to investigate the association of these previously reported adult bone density loci with pediatric BMC and BMD of the spine, hip and radius. Our cohort is derived from the multi-center National Institute of Child Health and Development (NICHD) Bone Mineral Density in Childhood (BMDC) Study, which was initiated in 2001 to establish national reference standards for BMD and bone accrual for children ages 6 years and older. The initial cohort of 1554 subjects (ages 6 to 16) was later enriched with an additional 452 subjects 5 and 19 years of age to further strengthen the reference data. To date, 604 of the 945 subjects from the Caucasian portion of the cohort have been both genotyped on the Illumina OMNI Express BeadChip platform and imputed for ~2.5 million SNPs, enabling us to query 59 of the 64 adult BMD SNPs in our pediatric cohort. Following Bonferroni correction for both the number of SNPs tested and the six traits tested (BMC and BMD at spine, hip and radius), there was strong evidence for association between allele A of rs13245690 at the *C7orf58* locus and reduced BMC and BMD of the distal 1/3 radius ($P = 6.96 \times 10^{-5}$; $\beta = -0.217$ and $P = 1.51 \times 10^{-4}$; $\beta = -0.224$). There was also multi-site evidence of nominal association for rs1053051 (*C12orf23*), rs17482952 (*WLS*), rs2062377 (*TNFRSF11B*), rs3755955 (*IDUA*), rs7851693 (*FUBP3*) and rs9533090 (*AKAP11*). In summary, as a consequence of analyzing a unique genome-wide genotyped dataset derived from an in-depth phenotyped pediatric cohort, we are well positioned to ascertain which loci contributing to adult BMD are operating in early life and at which skeletal site(s). As our dataset continues to grow, we will follow up these observations and also aim to uncover novel loci specific to bone accrual during childhood and adolescence.

813F

A novel *GNAI3* gain-of-function mutation associated with auriculo-condylar syndrome in a Brazilian family. *V.L.R. Tavares[1], T.T. Torres[1], H. Buermans[3], C. Masotti[1], D.F. Bueno[1], J.A. Horst[4], R.M. Zechi-Ceide[2], M.L. Cunningham[5], M.L. Guion-Almeida[2], M.R. Passos-Bueno[1]*. 1) Human Genome Research Center, University of Sao Paulo, Sao Paulo, Brazil; 2) Department of Clinical Genetics, Hospital of Rehabilitation of Craniofacial Anomalies, University of São Paulo (HRCA-USP), Bauru, SP, Brazil; 3) Leiden University Medical Center, Leiden Genome Technology Center, Leiden, Holland; 4) Division of Pediatric Dentistry, Department of Orofacial Sciences, UCSF, San Francisco, CA, USA; 5) Craniofacial Center, Seattle Children's Research Institute, Seattle, WA, USA.

Purpose: Auriculo condylar syndrome (ACS), characterized by micrognathia, microstomia, microtia and agenesis/hypoplasia of the mandibular condyle, is an autosomal dominant disorder. We have mapped ACS1 to a large candidate region at 1p21.1-q23.3. Recently, mutations in two genes, *GNAI3* and *PLCB4* were found to be mutated in ACS. Therefore, the present study has been conducted in order to elucidate the mutational mechanism in the Brazilian ACS family linked to 1p21.1-q23.3 (ACS1). Methods: We performed exome sequencing (ES) in two ACS1 affected individuals using the Illumina HiSeq2000 platform. Sanger sequencing (SS) was used to validate our findings in 3 affected and 4 non-affected individuals belonging to our previously published family (Masotti *et al*., 2008). Given that expression of *DLX5* and *DLX6* was decreased in mandibular osteoblasts of ACS patients (Rieder *et al*., 2012), we tested their expression in RNA extracted from mesenchymal stem cell (MSC) cultures. As cartilage is compromised in the ACS1 patients, we also compared the in vitro MSC chondrogenic differentiation between an ACS1 patient and a control. Results: No pathogenic mutations were found in *PLCB4*, but we found 4 variants in *GNAI3*. Of these, only one missense mutation, p.Asn269Tyr, segregates with the disease in the family, which was found to be a conserved amino acid by multiple sequence alignment, using MUSCLE software. GNAI3 protein structure modeling of specific amino acid substitution suggests that this N269Y mutation increases the stability of a natural ligand in a specific pocket, decreasing the reliance on the molecule. No significant difference was observed for *DLX5* and *DLX6* in MSCs, and neither between patient and control during in vitro chondrogenesis. Conclusions: N269Y is a mutation found in *GNAI3*, different from that previously described. Our findings suggest that the mutation is the cause of ACS1 in this Brazilian family and it is a gain-of-function mutation. If taken together our results with those from the literature, it seems that the gain-of-function mutation in *GNAI3* does not primarily compromise cartilage differentiation. Financial Support: CEPID/FAPESP, CNPq, CAPES, MCT.

814F

Non-syndromic cleft lip/palate: a disease driven by deficiency in genomic repair? *G.S. Kobayashi[1], L.A. Cruz[1], D.Y. Sunaga[1], D.F. Bueno[1], P. Francis-West[3], A. Kutta[3], B.V.P. Almada[1], S.G. Ferreira[1], M. Aguena[1], L.C. Andrade-Lima[2], C.F. Menck[2], M.R. Passos-Bueno[1]*. 1) Human Genome Research Center, University of Sao Paulo, Brazil; 2) Institute of Biomedical Sciences, University of Sao Paulo, Brazil; 3) Dept of Craniofacial Development & Orthodontics, King's College London, United Kingdom.

Purpose: Non-syndromic cleft lip/palate (NSCL/P) is a multifactorial disease that arises from errors during embryonic development. Although much effort has been put into identifying genetic and environmental factors underlying disease susceptibility, the aetiology of this complex malformation remains obscure. Since a tight regulation of ontogenetic mechanisms is required to ensure appropriate orofacial morphogenesis, our objective was to identify dysregulated pathways involved in the pathogenesis of NSCL/P. **Methods:** We performed a global transcriptome profiling of 7 dental pulp mesenchymal stem cell cultures from NSCL/P patients by comparison to 6 controls, using Affymetrix HuGene 1.0 ST chips. Differentially expressed genes were obtained using SAM and RankProd algorithms. Functional annotation and gene network analysis were performed with Ingenuity Pathways Analysis, while gene clustering, transcription factor and functional enrichment procedures were carried out using EXPANDER. Quantification of H_2O_2-induced DNA damage was assessed with flow cytometry for anti-γH2AX in 6 NSCL/P, 2 van der Woude syndrome (VWS), and 7 control stem cell cultures. Finally, RNA *in situ* hybridisation studies were carried out to CD1 mouse embryos at E10.5, E11.5, E12.5 and E13.5, using riboprobes for *Brca1*, *Rad51*, and *E2f1*. **Results/Conclusion:** We identified, in NSCL/P cells, a dysregulated transcriptional network associated with cell cycle progression and response to DNA damage, comprising abnormal expression of important molecules for these processes (e.g. *BRCA1*, *BRIP1*, *MSH2*, *RAD51* and *BLM*). These genes, also validated with qRT-PCR, exhibited a marked pattern of co-expression, and suggest that E2F1 may be a putative upstream regulator. By quantifying γ-H2AX in H_2O_2-exposed cells, we were able to confirm that NSCL/P cells exhibit abnormal response to DNA damage, which was also observed in IRF6-haploinsufficient (VWS) cells. *In situ* hybridisation studies revealed co-localised expression of *Brca1*, *Rad51* and *E2f1* in the facial primordia and developing palatal shelves of mouse embryos. Our results suggest that impairment of DNA damage response may be involved in the aetiology of NSCL/P and indicate that IRF6 may be functionally relevant in this process. Financial support: CEPID/FAPESP, CNPq, MCT.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

815F

Genetic characterization of osteopenia and bone overgrowth in Marfan syndrome mice. *S. Smaldone, J.R. Cook, M. Del Solar, F. Ramirez.* Pharmacology & Systems Therapeutics, Mount Sinai School of Medicine, New York, NY.

Mice under-expressing fibrillin-1 ($Fbn1^{mgR/mgR}$ mice) replicate early onset, progressively severe Marfan syndrome (MFS). Like MFS patients, 3 month-old $Fbn1^{mgR/mgR}$ mice display elongated tubular bones and reduced bone mass (osteopenia). The last phenotype was previously correlated with abnormally elevated TGFβ and BMP activity stimulating osteoblast differentiation and osteoblast-directed osteoclast activity. To circumvent the lethality of MFS mice from aortic aneurysm, a conditional Fbn1null allele ($Fbn1^{Lox}$) was inactivated in skeletal progenitor cells or osteoblast precursors by crossing $Fbn1^{Lox/-}$ mice with Prx1-Cre or Osx-Cre transgenic mice respectively. Microcomputer tomography (μCT) showed reduced bone mass in 3 months old $Fbn1^{Prx-/-}$ and $Fbn1^{Osx-/-}$ mice, and additional bone loss in 6 month-old animals $Fbn1^{Prx-/-}$ but not $Fbn1^{Osx-/-}$ mice. Colony forming unit-fibroblast (CFU-F) efficiency assays performed on marrow-derived mesenchyme stem cells (MSCs) yielded comparable numbers of CFU-Fs in 3 and 6 month-old WT and $Fbn1^{Osx-/-}$ preparations, more and fewer CFU-Fs than normal in MSCs isolated from 3 and 6 month-old $Fbn1^{Prx-/-}$ mice respectively. These findings therefore suggest that premature loss of osteoprogenitor cells exacerbates osteoclast-driven osteopenia in MFS. Similar ongoing analyses of crosses between $Fbn1^{Lox/-}$ mice and Prx1-Cre, Osx-Cre or Col2-Cre transgenic mice also indicate that limb overgrowth in MFS is a postnatal abnormality that is mostly accounted for by perturbed function of the perichondrial matrix.

816F

Evaluating the genetic disruption of lung hypoplasia in a mouse model of CDH. *S.C. Schecter, J. Wu, D. Miniati.* University of California San Francisco, Division of Pediatric Surgery 513 Parnassus Avenue HSW-1601, Box 0570 San Francisco, CA 94143-0570.

Introduction: An exponential increase in lung branching occurs between gestational days (E) 11.5 and 13.5 in mice. Treatment of pregnant mice with the mutagen combination of nitrofen-bisdiamine (N/B) causes lung hypoplasia and diaphragm defects in the fetal mice similar to the human condition of congenital diaphragmatic hernia (CDH). We hypothesized that genetic evaluation of fetuses exposed to N/B would identify specific disrupted molecular pathways critical in CDH. Methods: Pregnant CD-1 mice were treated with N/B or vehicle at embryonic day 8.5 (E8.5). Lung buds were dissected at E11.5–13.5 for histological and RNA analysis (n=6/timepoint/group), and evaluated for presence of CDH. Mutagen treatment causes two phenotypes—with and without CDH—which were analyzed separately. RNA expression profiles were derived using the Agilent whole mouse genome 4x44K Ink-jet array and analyzed using iReport (Ingenuity Systems, Redwood City). Results: Lungs of E11.5 and E12.5 mutagen-treated pups hypoplastic mainstem bronchi, dysmorphic primary lung buds, and delayed secondary domain branching. At E13.5, 30% of N/B-treated fetuses demonstrated diaphragmatic hernias along with lung hypoplasia. Analysis of N/B-treated fetuses versus control showed 365 differentially expressed genes. Among these, 237 were up-regulated and 128 were down-regulated. 8 differentially expressed genes (DEG) associated with retinoic acid receptor (RAR) activation were identified—COUP-TFII, PRKAR2A, FOS, ADCY1, and EP300 were upregulated, and AKT1, CYP26A1, and IL3RA were down-regulated. 4 DEGs associated with the Wnt/B-catenin signaling—EP300, AKT1, WNT8B, and LRP5. Exploring differences between N/B-treated E13.5 fetuses with CDH versus without CDH revealed 53 DEGs. Ion transport channel component gene ABCA9 was down-regulated in CDH, while TGFB and Wnt/B-catenin signaling FOXH1 and APC2 were up-regulated. Conclusion: N/B-administration results in changes in gene expression. Genes associated with promotion of RAR/RXR heterodimerism are down-regulated, while COUP-TFII—which inhibits RAR/RXR function was up-regulated. Ion transport may be associated with lung hypoplasia.

817F

The role of FMRP during early steps of neurodevelopment. *B. Bardoni, L. Davidovic, B. Mari, O. Khalfallah.* IPMC, CNRS UMR 6097, Valbonne, France.

The Fragile X Syndrome (FXS) is the first cause of inherited intellectual disability (ID). Affected children often manifest hyperactive, autistic behaviour, epileptic seizures, attention deficit, sleep disorders. The disease is due to the inactivation of the *FMR1* gene leading to the loss of the corresponding protein FMRP, that is a selective RNA binding protein involved in translational control. Indeed, FMRP is associates with polyribosomes and is involved in synaptic plasticity and dendritic spines development, explaining abnormal dendritic spines that are observed in FXS patients and animal model. Recently, it has been shown that FMRP plays also a role during neurogenesis suggesting that it is involved in early steps of development. Indeed, additional abnormalities have been observed in the brain of FXS young children. Moreover, several studies showed that *Fmr1*-null embryonic neural progenitors generated an altered neuron/glia ratio. Thus, we analysed the function of FMRP during neurogenesis, starting from murine embryonic stem (ES) cells using an RNAi strategy. *Fmr1*-knock-down (KD) ES cells were characterized by a microarray analysis and a quantitative proteomic analysis. Our data suggest an early signature of Fmr1 repression and it is then remarkable that at this phase of development the effects of the reduction of FMRP expression already affect also some pathways that are deregulated in FXS mature neurons. Then, we analysed the consequence of the repression of FMRP during the differentiation of ES cells into neural progenitors (NP), then into cells of the central nervous system (neurons and glial cells). NP and neurons derived from *Fmr1*-KD cells show an increased level of cell death, compared to NP obtained from *Fmr1*-WT cells and we observed that the expression of nestin is strongly reduced in Fmr1-KD NP compared with *Fmr1*-WT NP suggesting a delay or a switch in the differentiation of ES Fmr1-KD cells. Thus, these results strongly support the fact that the molecular alterations observed in ES cells have an impact on their differentiation into NP and neurons. Finally, we were able to rescue the apoptotic phenotype by treating *Fmr1*-KD cells with several drugs, interestingly some of these bio-active molecules are known to be able to rescue hallmarks characterizing the phenotype of the *Fmr1*-null mouse.

818F

[cf4]rbm24a[cf1] and *rbm24b* are required for normal somitogenesis and craniofacial development and are essential for Notch signaling.[cf1] *S. Maragh[1,2], R.A. Miller[2], S.L. Bessling[2], W. Huang[2], M.J. Parsons[2], A.S. McCallion[2,3].* 1) Biochemical Science Division, National Institute of Standards and Technology, Gaithersburg, MD; 2) McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD; 3) Department of Molecular and Comparative Pathobiology, Johns Hopkins University School of Medicine, Baltimore MD.

We recently identified the *RNA binding motif protein 24* (*Rbm24*) as a novel gene expressed during mouse cardiogenesis, and evaluated its developmental role in zebrafish cardiac development (orthologs *rbm24a* and *rbm24b*). Morpholino-based reduction of either ortholog by morpholino antisense oligonucleotides resulted in cardiogenic defects, reduced circulation, as well as defects in vasculogenesis and early angiogenesis. We have now demonstrated that both *rbm24* genes are expressed in the forming somites and are also expressed later in developing craniofacial structures. Consistent with these observations these orthologs are also required for normal somite and craniofacial development. All morpholino-induced dysmorphic phenotypes are rescued with target mRNA. We then investigated the possible function played by Rbm in Sonic hedgehog (SHH), Bone morphongenesis protein (BMP) and Notch developmental pathways. Notch-signaling initiation and maintenance is strongly reduced, but we have not observed any evidence of SHH or BMP pathway disruption. Characterization of spatial and temporal expression of components of the Notch-pathway via whole mount *in situ* hybridization on *rbm24* morpholino injected embryos indicates there is reduced expression of genes encoding Notch-ligands; namely, *deltaC*, *deltaD* and *dll4*. Subsequently, the expression of Notch-signaling target genes *her1*, *her7* and *hey2* was also suppressed. Expression of *deltaC* and d*eltaD* is required for somite patterning, while *dll4* is required for proper cardiovascular development. Consistent with these observations *hey1*, which is expressed in somites during somitogenesis and later required for craniofacial cartilage patterning, is also reduced in expression. By contrast relative expression of the Notch-receptors is not altered. Taken together these data indicate a previously unknown requirement of *rbm24a* and *rbm24b* for normal cardiovascular, somite and craniofacial development and support the hypothesis these genes play a role in proper expression of Notch pathway ligands. The *RBM24* ortholog in *C. elegans* is known to be involved in splicing. Our ongoing studies will continue to on focus the potential mechanism of action that may connect this critical pathway (Notch) with RBM24.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

819F

Kabuki Syndrome: Functional analysis of *MLL2* and *KDM6A* in Vertebrate Development. *P. Van Laarhoven[1], L. Neitzel[1], E.A. Geiger[1], E.H. Zackai[2], K.B. Artinger[3,4], J.E. Ming[2], T.H. Shaikh[1,3]*. 1) Section of Clinical Genetics and Metabolism, Department of Pediatrics, University of Colorado School of Medicine, Aurora, CO; 2) Division of Human Genetics, The Children's Hospital of Philadelphia, Philadelphia, PA; 3) Department of Molecular Pharmacology and Therapeutics and Center for Biomedical Informatics, Loyola University Chicago Stritch School of Medicine, Maywood, IL; 4) University of Colorado School of Dental Medicine, Aurora, CO.

Kabuki syndrome (KS) is a rare multiple congenital anomaly syndrome characterized by distinctive facial features, skeletal and dental abnormalities and intellectual disability. Congenital heart defects, short stature and renal abnormalities are also seen in a significant number of KS patients. The application of high-throughput sequencing technologies has led to the identification of mutations in *MLL2* as the cause of the disease in 75% of KS patients. *MLL2* is a histone methyltransferase that is an important regulator of transcription. The recent discovery of deletions in *KDM6A* in three patients with KS suggested that additional genes are likely to be involved. Interestingly, the histone H3 lysine 27 demethylase *KDM6A* is a a component of a multi-protein complex that also contains MLL2. We analyzed 40 KS patients for gene mutations using a combination of Sanger sequencing, arrayCGH and whole exome sequencing (WES). We detected *MLL2* mutations in 27/40 and *KDM6A* mutations in 3/40 KS patients. The *KDM6A* mutations included two microdeletions and one point mutation. This data led us to hypothesize that KS may result from mutations in multiple genes within a complex of interacting proteins. The 10 mutation-negative KS patients are currently being analyzed by WES to detect novel gene mutations. To further demonstrate causality and roles of candidate genes in common developmental pathways for KS, we have used morpholino-based knockdowns of gene expression in zebrafish. We knocked down zebrafish *mll2* and the two paralogs *kdm6a* and *kdm6al*. Both *mll2* and *kdm6a* morphant fish exhibited abnormal craniofacial structures which included absence of the branchial arches and otoliths, as well as absence, clefting or inversion of the ceratohyal and abnormal patterning or clefting of Meckel's cartilages. In contrast, *kdm6al* morphants did not exhibit a craniofacial phenotype. Analysis of transgenic lines expressing GFP under neuronal (*isl1*) and cardiac (*cmlc2*) specific promoters revealed that *mll2* and *kdm6al*, but not *kdm6a* morphants, exhibited gross morphological defects in neuronal patterning and cardiac development. Furthermore, the two zebrafish paralogs, *kdm6a* and *kdm6al* have functionally diverged, yet together they regulate the same developmental pathways as human *KDM6A*. These observations lend further support to the hypothesis that perturbation of a shared regulatory pathway containing *MLL2* and *KDM6A* is responsible for the Kabuki syndrome phenotype.

820F

[cf4]Fibulin-4b[cf1] is required for cardiovascular and musculoskeletal development as an antagonist of transforming growth factor-beta.[cf1] *Z. Urban[1], S.M. Khatri[1], A.B. Maxfield[2]*. 1) Department of Human Genetics, University of Pittsburgh, Pittsburgh, PA; 2) Department of Pediatrics, Washington University, St. Louis, MO.

Human mutations in the *fibulin-4* (*FBLN4*, *EFEMP2*) gene cause a recessive disease with arterial tortuosity, aneurysms, developmental emphysema and cutis laxa. Our goal was to elucidate the molecular mechanisms of this disease using zebrafish as a model organism. The zebrafish genome has two paralogues of human *FBLN4*, *fbln4a* and *fbln4b*, with *fbln4b* showing greater homology and conserved synteny to human *FBLN4*. To determine the role of *flbn4b* in early development, we used antisense morpholinos to knockdown gene function. *Fbln4b* knockdown showed a severe phenotype at 2 days post fertilization including cardiac edema, decreased heart rate, muscle dysfunction leading to reduced tail twitching and circulation failure with pooling of the blood in the caudal region. The observed phenotypes caused by *fbln4b* depletion were similar to phenotypes observed in embryos with hyperactivated transforming growth factor beta (tgfb) signaling. To test if tgfb signaling was altered in *fbln4b*-depleted embryos, we treated them with a small molecule tgfb receptor 1 inhibitor, which rescued the *fbln4b*-deficient phenotype. The specific tgfb isoform responsible for altered signaling was identified by silencing *tgfb1b*, *tgfb2* and *tgfb3* in *fbln4b* morphants. *Tgfb2* and *tgfb3* but not *tgfb1b* knockdown rescued all of the defects caused by *fbln4b* depletion. Our results show that the main function of *fbln4b* is to negatively regulate the activity of *tgfb2* and *tgfb3* in cardiovascular and musculoskeletal development. The molecular mechanism of this function is unclear but likely involves stabilization of latent tgfb complexes in the extracellular matrix.

821F

Three-dimensional Study Of Ror2-/- and Wnta5-/- Mice Craniofacial Phenotype Using Two Imaging Techniques. *S. Beiraghi[1], V. Leon-Salazar[1], J. Zhang[2], D. Idiyatullin[2], M. Garwood[2], A.S.L. Fok[1], M. Maga[3], Y. Yang[4]*. 1) School of Dentistry, University of Minnesota, Minneapolis, MN; 2) Center for Magnetic Resonance Research, University of Minnesota, Minneapolis, MN; 3) Division of Craniofacial Medicine, Seattle Children's Hospital, Seattle, WA; 4) Genetic Disease Research Branch National Human Genome Research Institute Bethesda, MD.

Background: Animal studies have shown that skeletal and soft tissue phenotypes of Robinow syndrome are manifested in Ror2-/- and Wnta5-/- mice. The aims of this study were: 1) To compare the craniofacial phenotype of Ror2-/- and Wnta5-/- mice, and 2) to assess the performance of a new MRI method, known as SWIFT (Sweep Imaging with Fourier Transform), for the three-dimensional analysis of the skeletal craniofacial characteristics of Ror2-/- and Wnta5-/- mice as compared to micro-computed tomography (μCT). Methods: Eight E18.5 mouse heads (two Wnt5a-/-, two Ror2-/-, and four wild-type) were used in this study. High-resolution μCT of each head was obtained using a X-Tek XTV μCT scanner (Nikon metrology NV, Leuven, Belgium) at 90 kV, 91 μA, and 2000 projections. Qualitative and quantitative comparisons included: Craniofacial bone volume, maxillary and mandibular size and shape variation, and trabecular bone pattern. Proton MRI measurements were performed on a 31 cm bore 9.4 T scanner (Oxford Magnet/Agilent DirectDrive Console) using a home-made single loop surface coil (2 cm diameter). SWIFT images were acquired with bandwidth=125 kHz, TR=3.3 ms, matrix=256^3, number of spokes=128000, and FOV=2 cm x 2cm x 2cm. The acquisition time was 6 minutes for each SWIFT image. Results: Wnt5a-/- and Ror2-/- mice presented a significant reduction in the facial and cranial vault development when compared to the wild-type mice. Midfacial development was also reduced in both mutant mice; however the Wnt5a-/- mice presented a more severe maxillary deficiency. Mandibular length was also decreased in both Wnt5a-/- and Ror2-/- mice. The trabecular bone pattern of the maxillary and mandibular alveolar bone was dysplastic and thickened in the Wnt5a-/- mice. Interestingly, the mandible, maxilla, and teeth in the wnt5a-/- mice were more severely affected than in the Ror2-/- mice. Using SWIFT MRI it was possible to delineate craniofacial skeletal structures and the three-dimensional images were similar to those obtained by μCT. Differentiation between cortical and cancellous bone was also possible using SWIFT MRI. Conclusion: Our data confirms that mutations of the Ror2/Wnt5a pathway affect skeletal and dental growth and development and cause severe malformation of the maxillary and mandibular alveolar bone. SWIFT MRI allows visualization of the morphology of the craniofacial skeleton of fetal mice and is a promising imaging technique for phenotypic studies.

822F

Retinal disease gene identification and generation of photoreceptor gene regulatory networks by ChIP-Seq analysis of multiple transcription factorsPurpose: A network of photoreceptor transcription factors, including neural retina leucine zipper (NRL) and cone-rod homeobox (CRX), dictates the develop. *H. Hao[1], M. Brooks[1], M. Seifert[2], A. Swaroop[1]*. 1) Neurobiology Neurodegeneration & Repair Laboratory, National Eye Institute, Bethesda, MD; 2) Genomatix GmbH, 80335 Munich, Germany.

Purpose: A network of photoreceptor transcription factors, including neural retina leucine zipper (NRL) and cone-rod homeobox (CRX), dictates the development and survival of photoreceptors. Mutations in human NRL, CRX and several of their target genes are associated with retinal diseases. We aim to identify the global direct target genes of NRL and CRX during distinct developmental stages to facilitate retinal disease gene discovery and construction of gene regulatory networks. Methods: Genome-wide in vivo occupancies of transcription factors were identified by ChIP-seq analysis of P2 mouse retina and compared to published data from the adult retina. Changes in transcription profiling between wild type mouse retina and Nrl-knockout retina were used as a second filter to identify direct NRL target genes. Results: We identified direct NRL targets genes at both P2 and compared these with the previously identified targets at P28. The direct NRL target genes are enriched in retinal disease loci, suggesting that these targets genes will serve as retinal candidate disease genes. Conclusions: Genome-wide occupancy analysis of photoreceptor specific TF at distinct developmental stages allowed identification of genes that are of functional importance to photoreceptor development and maintenance. Our study will facilitate the identification of retinal disease genes and therapeutic discovery.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

823F

Genome-wide ChIP-seq analysis to identify mouse Foxl2 binding sites *in vivo*. *M. Marongiu[1], A. Sbardellati[2], L. Marcia[1], A. Meloni[1], R. Cusano[2], A. Angius[1], G. Fotia[2], F. Cucca[1,3], L. Crisponi[1]*. 1) IRGB - CNR, Monserrato (CA), Cagliari, Italy; 2) CRS4, Pula, Cagliari, Italy; 3) University of Sassari, Sassari, Italy.

The transcription factor FOXL2 belongs to the evolutionarily conserved family of winged helix forkhead proteins that regulate transcription of genes involved in cell growth, proliferation and differentiation and play a central role during development. Foxl2 is involved in female sex determination, ovarian development and follicle formation. In human, FOXL2 mutations are responsible of blepharophimosis, ptosis, and epicanthus inversus syndrome (BPES) type I and II, associated with premature ovarian failure (POF) in type I. Recently, a somatic FOXL2 mutation (p.Cys134Trp) has been found in adult ovarian granulosa cell tumors (OGCT). Foxl2 expression has been found starting at 10.5 dpc in the 1st branchial arch and it is expressed in the developing ovary starting at 13.5 dpc, when sex determination from bipotential gonad occurs, until adulthood. Foxl2 is also expressed in the developing pituitary gland starting at 11.5 dpc. To better characterize its function both in ovarian tissue and pituitary gland, we performed chromatin immunoprecipitation (ChIP) on mouse P7 ovaries, and on alpha T3-1 gonadotropic cell line. We used ChIP-seq, which combines ChIP with high-throughput massively parallel sequencing, to identify the genomic locations bound by Foxl2, and we found a high relative enrichment level of these regions in the promoters of known genes (Refseq) with respect to the genome background (5% ovary; 7% alpha T3-1). Using MEME-chip (http://meme.sdsc.edu/meme/) and CisFinder (http://lgsun.grc.nia.nih.gov/CisFinder/), specifically designed for finding over-representing short DNA motifs in ChIP-seq data, we were able to clearly identify a core consensus binding sequence for Foxl2 that is over-represented in the enriched ChIP regions from both experiments. The results of this study will help in finding new targets for Foxl2 as well as in better understanding the mechanisms underlying the role of Foxl2 in ovarian and pituitary development.

824F

Cc2d2a is required for cilia biogenesis. *S. Veleri[1], TJ. Foskett[1], SH. Manjunath[1], A. Longo[1], MA. English[1,4], P. Liu[2], J. Lei[2], C. Gao[3], RN. Fariss[3], R. Sood[4], RA. Rachel[1], P. Liu[4], L. Dong[2], A. Swaroop[1]*. 1) NNRL, NEI, NIH, Bethesda, MD; 2) Genetic Engineering Core, NEI, NIH, Bethesda, MD; 3) Biological Imaging Core, NEI, NIH, Bethesda, MD; 4) NHGRI, NIH, Bethesda, MD.

Purpose: Mutations in CC2D2A (Coiled-Coil & C2 Domain containing 2A) are associated with Meckel-Gruber syndrome (MKS) and Joubert syndrome (JS), the two ciliopathies characterized by retinal dystrophy, mental retardation, polydactyly, liver fibrosis, and polycystic kidney. CC2D2A protein is localized to basal body and transition zone of the cilia. The purpose of this study was to elucidate how CC2D2A mutations cause human disease by loss of function studies using model organisms. Methods: We used antisense splice blocking and translational blocking morpholinos (MOs) to knockdown zebrafish cc2d2a. The splice blocking MO mimicked a disease-associated splicing defect in human patients. The Cc2d2a-knockout (KO) mice were produced by targeted deletion of exons 6 to 8 using homologous recombination. Results: Knockdown of cc2d2a by both splice and translational blocking MOs in the zebrafish embryos demonstrated dose-dependent developmental defects, including brain atrophy and microphthalmia. The Cc2d2a-/- mice exhibit embryonic lethality, with extensive developmental defects that include situs inversus, heterotaxy, polydactyly, anophthalmia, hydrocephalus and liver fibrosis. Occasionally, Cc2d2a-/- animals survive for a month but display severe hydrocephalus and retinal dystrophy. The analysis of embryonic fibroblasts, embryonic node and kidney tubules showed that cilia biogenesis is disrupted by Cc2d2a-/- mutation. Conclusion: Our studies suggest that Cc2d2a function is critical for normal body plan and organ development that involves cilia-mediated signaling pathways.

825F

***Prdm16* is required during mouse craniofacial development.** *B.C. Bjork[1], F. Schnutgen[2], L. Furlan[1], B. Nelson[1], D.R. Beier[3]*. 1) Dept Biochemistry, Midwestern Univ, Downers Grove, IL; 2) Molecular Hematology Dept, Univ of Frankfurt, Frankfurt, Germany; 3) Div Genetics, Brigham & Women's Hospital, Harvard Med School, Boston, MA.

The autosomal recessive *csp1* mutation exhibits perinatal lethality with cleft palate accompanied by mandibular hypoplasia, incisor tooth defects, choroid plexus hypoplasia, retinal folds, and lung and cardiac defects. This mutant models Pierre Robin Sequence-type CP, a form of cleft palate exhibiting mandibular hypoplasia and ankyloglossia. PRDM16 is a regulator of TGFβ superfamily signaling, and we have shown that TGFβ signaling is perturbed in *Prdm16* loss of function mutants. *Prdm16* has been shown to be critical during brown adipose versus skeletal muscle cell fate determination and in the maintenance of stem cell populations; however, its role in craniofacial patterning is relatively uncharacterized. We aim to determine the molecular consequences of altered *Prdm16* expression during embryonic palate, mandible and tooth development.

To this end, we have generated a novel conditional gene trap null allele of *Prdm16* (*Prdm16^{cGT}*) using "targeted trapping" in mouse embryonic stem cells. We previously validated the conditional capability of this combination "*cre-loxP*" and "*Flp-Frt*"-based FLEx recombination system using ubiquitous *cre*- and *Flp*-expressing mouse strains. Using whole-mount *in situ* hybridization, we are interrogating gene trap null mutant versus wild type embryos with an array of genes of known importance during early mandible development and formation and patterning of the secondary palate to identify *Prdm16*-dependent gene expression differences. To study the mechanism by which cleft secondary palate occurs in these mutants (i.e. palate-intrinsic versus palate-extrinsic), we initiated conditional *Prdm16* ablation studies using palate mesenchyme-specific (*Osr2::cre*) and mandible-specific (*dHand::cre*) strains. These studies will allow us to dissect the genetic and developmental mechanisms that involve *Prdm16* during craniofacial development.

826F

Examination of lungs and kidneys in a mouse model of osteogenesis imperfecta. *E. Carter[1], S. Doty[2], N. Pleshko[3], C. Raggio[1]*. 1) Ctr Skeletal Dysplasias, Hosp Special Surgery, New York City, NY; 2) Tissue Microscopy, Hosp Special Surgery, New York City, NY; 3) Mechanical Engineering, Temple Univ., Philadelphia, PA.

Background: Osteogenesis Imperfecta (OI) is a genetic disorder of the connective tissues characterized by predisposition to fracture, tissue fragility, skeletal deformities, and in severe cases, death. It is primarily caused by heterogeneous dominant mutations in the genes encoding type I collagen, the major protein of bone, skin, and tendons. Treatment options include osteoclast inhibition, through bisphosphonate therapy and/or RANKL inhibitors. The *oim/oim* mouse model is a naturally-occurring mutation resulting in a deficiency of proalpha2(I) collagen (B6C3Fe*a/a*Col1a2oim/J (Jackson Laboratories)). The phenotype is similar to that of moderate to severe OI in humans, Sillence type 3. In addition to the bony phenotype, *oim/oim* mice have a type I collagen glomerulopathy, demonstrating the negative affect of alpha1(I) homotrimers in non-osseous tissues. As part of a larger ongoing study examining the effect of osteoclast inhibition in the *oim/oim*, we collected lungs and kidneys from *oim/oim* and wild type (WT) mice across various treatment groups for histopathology and fourier transform infrared imaging spectroscopy (FTIR-IS). Our hypotheses were: (1) *oim/oim* would have decreased type 1 collagen in lung tissue compared to WT; (2) osteoclast inhibitors would not affect type 1 collagen in the lung; (3) treatment would decrease deposition of type I collagen breakdown products in glomeruli of *oim/oim*. Methods: Between 3–5 animals from each of the ten treatment groups were sacrificed at 26 weeks of age. Lungs (whole lobe) were colleted in 4% paraformaldehyde for histopathological study by H&E, picosirius red (PSR), type 1 collagen immunostaining (rabbit Ab to purified fetal mouse skin). Kidneys were collected and halved into electron microscopy fixative, formalin, and 80% ethanol for histopathology (as listed above) and FTIR-IS. Results: Lungs from *oim/oim* animals were smaller than WT specimens. On light microscopy, both the *oim/oim* and WT lungs showed some tissue compression with decreased alveolar space. This was random with no trend towards genotype. There were no significant differences between H&E, type 1 collagen, or PSR in both genotypes and all treatment groups. Immunostaining, PSR, and FTIR-IS showed expected deposition of type I collagen breakdown products in the glomeruli of *oim/oim* (glomerulopathy). Treatment with osteoclast inhibitors did not improve the glomerulopathy in *oim/oim* mice.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

827F

Pathway analysis by quantitative profiling of gene expression patterns in developing mouse retina. *M.J. Brooks[1], I. Nasonkin[1], S. Perez[1], S. Veleri[1], J. Roger[1], M. Seifert[2], A. Swaroop[1]*. 1) NNRL, NEI, NIH, Bethesda, MD; 2) Genomatix GmbH, Munich, Germany.

The developing mouse retina provides a unique model for *in vivo* investigation of gene expression and regulatory pathway analysis during differentiation of a neuronal tissue. Ganglion cells, amacrine neurons, horizontal cells and cone photoreceptors are primarily generated prenatally in mouse retinal development; whereas bipolar, Muller glia and a majority of rod photoreceptors are born after birth. Some of the key transcription factors involved in retinal differentiation have been identified, yet the regulatory pathways underlying neuronal cell fate determination are poorly understood. This study employs the power of next-generation sequencing and pathway discovery tools to investigate the transcriptional regulatory networks during retinal development. RNA from invaginated neuroectoderm of the optic cup from embryonic day (E)11 and neural retina from E14 and 16 and postnatal days 0, 2, 4, 6, 9, 10 and 21 were used for the construction of RNA-seq libraries. Using a modified Illumina TruSeq protocol, we performed sequencing of directional RNA-seq libraries of 76 nucleotides on an Illumina GAIIx. We used Genomatix, JMP 10, Ingenuity Pathway Analysis and AltAnalyze for sequence alignment, pathway and alternate splicing and promoter analysis. A range of 28–36 million uniquely mapped, pass filter reads were aligned to the reference genome mm9. Of over 17,000 transcripts that show expression greater than 1.0 RPKM at any profiled time-point, approximately half show significant differential expression over the course of retinal development. As predicted, genes involved in DNA replication, cell proliferation and mitosis are among the most down-regulated genes as differentiation proceeds. Photoreceptor specific transcripts are the most up-regulated genes. Pattern recognition analysis of known key regulatory genes has revealed a more thorough assessment of genes that are likely involved in distinct developmental pathways. Alternative splice and promoter analysis has identified many genes that exhibit different isoforms during differentiation. In addition, the transcript structure of several genes was significantly different in the retina compared to the currently published annotation. Our studies provide the framework for identifying novel candidate genes for mutation screening of patients with retinal diseases and for elucidating the regulatory network that dictates differentiation of multiple cell types in the mammalian retina.

828F

Essential Roles of FGF Ligands in Endochondral and Intramembranous Bone Development. *I.H. Hung[1,3], D.M. Ornitz[2], G.C. Schoenwolf[1], M. Lewandoski[3]*. 1) Department of Pediatrics, University of Utah, Salt Lake City, UT; 2) Department of Developmental Biology, Washington University School of Medicine, St. Louis, MO; 3) Cancer and Developmental Biology Lab, Frederick National Laboratory for Cancer Research, Frederick, MD.

Activating mutations in fibroblast growth factor (FGF) receptors result in chondrodysplasia and craniosynostosis syndromes, highlighting the critical role for FGF signaling in skeletal development. Although the roles of the FGFRs in bone development have been relatively well-characterized, only two FGF ligands, FGF9 and FGF18, have been shown to regulate embryonic skeletogenesis. To further elucidate the roles of these FGF ligands during skeletal development, we have generated mice lacking both Fgf9 and Fgf18. Analyses of Fgf9/18 double knock-out (DKO) mice demonstrate severe defects in endochondral and intramembranous ossification. We provide evidence that FGF signaling interacts with key gene regulatory networks in osteogenesis. Together these findings provide novel insights into mechanisms of skeletal biology.

829F

Taking a different pathway: Notchless plays with p53 and Wnt during embryogenesis. *A.C. Lossie[1,2], C.-L. Lo[1,3], J.B. Sherrill[1,4]*. 1) Department of Animal Science, Purdue University, West Lafayette, IN; 2) Dept of Medicine, Indiana University School of Medicine, Indianapolis, IN; 3) PULSe Interdisciplinary Graduate Program, Purdue University, West Lafayette, IN; 4) Department of Biological Sciences, Purdue University, West Lafayette, IN.

Our interests lie in determining the genes and genetic pathways that are important for establishing and maintaining maternal-fetal interactions during pregnancy. Through a positional cloning strategy, we discovered that mutations in *Notchless* (*Nle1*) lead to embryonic lethality during peri-implantation in mice. Nle1 is a member of the WD40-repeat protein family, and is thought to signal via the canonical Notch pathway. In invertebrates and lower vertebrates, the Notch pathway is critical for directing cell fate prior to gastrulation. However, the role of Notch signaling during the earliest stages of mammalian development is unclear, as gene targeting studies of Notch family members and factors that are necessary for Notch signal transduction, demonstrate that Notch signaling is dispensable for gastrulation in mice. The phenotype of *Nle1* mutant embryos is much more severe than single *Notch* receptor mutations or even in animals where Notch signaling is blocked. To test the hypothesis that *Nle1* functions in multiple signaling pathways during pre-implantation development, we examined expression of multiple *Notch* downstream target genes, as well as select members of the *Wnt* and *Trp53* pathways in wild-type and mutant embryos. Surprisingly, we saw no indication that the *Notch* pathway is disrupted in mutant embryos; *Notch* receptors, ligands and downstream targets showed normal expression levels. Instead, we found that members of the *Wnt* pathway are downregulated in *Nle1* mutants, while *Trp53* and *Cdkn1a* were upregulated in a stage-specific manner. Induction of *Cdkn1a* was highest in morulae and full blastocysts, while *Trp53* overexpression was confined to hatched blastocysts. Our results refute the possibility that Nle1 is a negative regulator of Notch signaling during mammalian pre-implantation development, as mutation of *Nle1* does not lead to increased expression of downstream target genes. Instead, our data implicate Nle1 in Wnt signaling, cell cycle arrest via Cdkn1a, and p53-mediated apoptosis. Although Notch signaling is dispensable in mice prior to gastrulation, Wnt signaling is not. Deletion of *Wnt3* leads to failure prior to primitive streak formation, and multiple ligands and receptors are detected in blastocysts and the uterus during peri-implantation. As p53 inhibits self-renewal and promotes differentiation in ES cells, Nle1 and its co-opted Wnt and p53 pathways could provide novel targets for the design of therapeutic interventions for infertility.

830F

Developmental defects in the retina of Pias3 conditional knockout mice. *J.E. Roger[1], H. Breit[1], D. Cheng[1], L. Dong[2], A. Swaroop[1]*. 1) NNRL, National Eye Institute, Bethesda, MD; 2) Genetic Engineering Facility, National Eye Institute, Bethesda, MD.

A stringent control of gene expression is essential for development and homeostasis. Posttranslational modifications provide a critical level of control and additional complexity. Protein Inhibitor of Activated STAT3 (PIAS3) is a transcriptional modulator that directly binds to multiple factors and modulates their activity. PIAS3 can also function as a SUMO (Small Ubiquitin-like Modifier)-E3 ligase, and covalent linkage of SUMO proteins can modify the function of target proteins. Recently, Blackshaw and our laboratory demonstrated the importance of SUMOylation in modulating activity of NR2E3 and NRL, two key regulators of photoreceptor differentiation. In order to analyze the consequence of the lack of PIAS3 expression early in retinal development and identify new PIAS3 targets, we have generated and analyzed conditional Pias3 knockout mice (CKO). Pias3 floxed mice were produced by introducing LoxP sites flanking exon 2 to 5 of the Pias3 gene through homologous recombination. Pias3 was specifically deleted in the forebrain and the retina by crossing with Rx-Cre mice. Phenotype was analyzed by immunohistochemistry (IHC), qPCR and electroretinogramm (ERG). A specific deletion of Pias3 in the retina in Pias3f/f; Rx-Cre mice was confirmed by RT-PCR and PCR on retinal cDNA and genomic DNA respectively. Immunoblot analysis confirmed the complete absence of Pias3 protein in the retina of Pias3f/f; Rx-Cre mice. Immunohistochemical (IHC) analysis with anti-S-opsin antibody and peanut agglutinin on flat mount retina revealed an increased number of S-cones along the ventral-dorsal axis but with shorter outer segments. No obvious change was observed in other retinal cell types based on IHC, with the exception of Muller cells showing displaced cell bodies in the inner nuclear layer. In 6-month old Pias3 CKO, no signs of retinal degeneration were observed. Interestingly, ERG analysis showed a significant decrease of the photopic response (both S-cones and M-cones) compared to control animals, whereas scotopic ERG (rod-mediated) appeared to be normal. The sumoylation state of key factors involved in retinal diseases is being tested, and gene expression analysis will be performed in these mice. Our studies of CKO retina reveal the importance of PIAS3 in regulating cone photoreceptor differentiation and function.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

831F

Uncovering the function of TMED2 during. *T. Heba.* Human Genetics & Pediatrics, Research Inst Place Toulon, Montreal, PQ, Canada,-tAGHREED HEBA.

Uncovering the function of TMED2 during trophoblast differentiation Taghreed Heba, Abeer Zakariyah, Loydie A. Jerome-Majewska Transmembrane emp24 domain trafficking protein 2, (TMED2) is a member of the p24 family of proteins involved in vesicle transport between the ER and Golgi. During vesicular transport between the ER and Golgi p24 proteins function as receptors for both cargos and coat proteins. Our group showed that Tmed2 is required for normal embryo and placental development in mouse and that syncytiotrophoblast cells of the mouse labyrinth placenta failed to differentiate in homozygous mutant embryos. In human placenta, we showed expression of TMED2 between 5.5 and 40 weeks of gestation in all trophoblast cell types. We noted that early in gestation TMED2 was more highly expressed in cytotrophoblast cells versus syncytiotrophoblast. The choriocarcinoma cell lines BeWo and JEG-3 are widely used for the study of trophoblast differentiation. These cells share many properties with villous trophoblast in terms of their morphology, biochemical markers, and hormone secretion. We found that TMED2 was more highly expressed in a choriocarcinoma cell line, BeWo, which can be induced to differentiate and form syncytiotrophoblast when compared to the JEG-3 cell line, which does not fuse to form syncytiotrophoblast. We hypothesized that TMED2 is required for fusion of trophoblast cells during syncytiotrophoblast differentiation. To test this hypothesis we are examining the function of TMED2 during trophoblast differentiation of BeWo and Jeg-3 cell lines. We will show our plans to ectopically express TMED2 in Jeg-3 cells and to knockdown TMED2 expression in BeWo choriocarcinoma cells with shRNA. Our work suggests that TMED2 is required for trafficking cargoes that are essential for placental development.

832F

Mouse models of neurodevelopmental disorders: Overlap between FoxG1 and MeCP2. *S.A. Adams[1,2], R. Woods[1], S. Luu[1], J.M. LaSalle[1].* 1) Medical Microbiology and Immunology, University of California, Davis, Davis, CA; 2) Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge, UK.

Rett Syndrome is a neurodevelopmental disorder caused by mutations in the X-linked gene *MECP2*. The role of MeCP2 in early neuronal development has not been fully investigated. More recently, mutations in the autosomal gene *FOXG1* have been identified in children with a congenital variant of Rett Syndrome. Others have suggested that *FOXG1* mutations define and overlapping but distinct syndrome from classic Rett syndrome, defined by microcephaly, severe mental retardation, dyskinesia, and corpus collosum hypogenesis. *FOXG1* encodes a developmental transcription factor, and is a known regulator of embryonic cortical development. Given that mutations in *FOXG1* and *MECP2* produce a similar phenotype, we investigated whether they are involved in parallel or overlapping molecular pathways during specific developmental time frames. We show by immunofluorescence in embryonic brain sections and cultured cortical neurons that MeCP2 and FoxG1 are expressed at increasing levels in the developing mouse cortex from E13.5 to E17.5. Using chromatin immunoprecipitation (ChIP) and quantitative RT-PCR, we found that MeCP2 and FoxG1 bind to each other's promoters and regulate each other's expression. Finally, we used ELISA to identify that IGF-1 protein levels are differentially regulated in neurons isolated from Foxg1 haploinsufficient mice ($Foxg1^{+/-}$) and MeCP2 knockout mice ($Mecp2^{-/y}$). We treated $Foxg1^{+/-}$ and $Mecp2^{-/y}$ neurons with IGF-1 at different stages in early neuronal differentiation and demonstrate that aberrant IGF-1 signaling is responsible for the defects in early neurite outgrowth that may go on to impede aspects of postnatal synapse formation. Our research demonstrates that an early cortical phenotype is associated with Rett Syndrome, and may be beneficial in the future design of treatments for neurodevelopmental disorders.

833F

Comprehensive characterization of a Zebrafish model for pseudoxanthoma elasticum. *M.J. Hosen[1,2,4], O.M. Vanakker[1,4], A. Willaert[1], A. Huysseune[3], P.J. Coucke[1], A. De Paepe[1].* 1) Centre for Medical Genetics, Ghent University Hospital, Ghent, Flanders, Belgium; 2) Dept. of Genetic Engineering and Biotechnology, Shahjalal University of Science and Technology, Sylhet, Bangladesh; 3) Dept. of Biology, Ghent University, Ghent, Belgium; 4) These authors contributed equally to the work.

Aims: Pseudoxanthoma elasticum (PXE) is a connective tissue disease characterized by elastic fiber calcification and fragmentation, resulting from mutations in the ABCC6 gene. In a zebrafish (ZF) model, which has three abcc6 isoforms, abcc6a knockdown has been associated with abnormal longitudinal growth of the tail, which is of interest as the pro-osteogenic BMP2-Runx2 pathway - which was recently found upregulated in human and murine PXE - is important in ZF tail patterning. Consequently, we aimed to study mineralization and BMP2 signaling in PXE ZF. **Methods & Results:** Abcc6a was knocked down by injecting translational blocking or splice junction morpholinos (MO) in 1–4 cell stage embryos. At 3 days post fertilization, both groups showed curving and shortening of the tail, variable in severity, pericardial edema and decreased mobility. Moreover, decreased total body length and underdevelopment of head and eyes were observed. Calcein and Alizarin staining showed advanced skeletal mineralization compared to controls. QPCR analysis on isolated tail tissue revealed significant upregulation of BMP2a, RUNX2a and MSXc. **Discussion & Conclusion:** Our study expands the knowledge on PXE ZF and illustrates a distinct phenotype, affecting longitudinal growth and eye development. Interestingly, similar to humans, variability in severity is observed. Though we could not detect soft tissue calcification - possibly because of early demise due to pericardial edema - the mutants showed advanced skeletal mineralization, suggesting that abcc6a deficiency also influences calcium precipitation. Confirmation of BMP2-Runx2 and Msx2 involvement in the PXE pathogenesis leads us to conclude that the ZF is a useful model organism for future PXE cell signalling research.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

834F

Mouse model reveals the role of SOX7 in the development of congenital diaphragmatic hernia associated with recurrent deletions of 8p23.1. *D.A. Scott[1,5], M.J. Wat[1], T.F. Beck[1], A. Hernández-García[1,2], Z. Yu[1], D. Veenma[3,4], M. Garcia[5], A.M. Holder[6], J.J. Wat[7], Y. Chen[1,8], C.A. Mohila[9], K.P. Lally[10], M. Dickinson[5], D. Tibboel[3], A. de Klein[4], B. Lee[1,8]*. 1) Deptartment of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA; 2) Universidad Autónoma de Coahuila, Saltillo, Coahuila, México; 3) Department of Pediatric Surgery, Erasmus Medical Center, Rotterdam, the Netherlands; 4) Department of Clinical Genetics, Erasmus Medical Center, Rotterdam, the Netherlands; 5) Department of Molecular Physiology and Biophysics, Baylor College of Medicine, Houston, TX., USA; 6) Department of Surgical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA; 7) Department of Biochemistry and Cell Biology, Rice University, Houston, TX, USA; 8) Howard Hughes Medical Institute, Houston, TX, USA; 9) Department of Pathology and Immunology, Baylor College of Medicine, Houston, TX, USA; 10) Department of Pediatric Surgery, University of Texas Medical School, Houston, TX, USA.

Recurrent microdeletions of chromosome 8p23.1 confer a high risk of both congenital diaphragmatic hernia (CDH) and cardiac defects. Haploinsufficiency of *GATA4*—a transcription factor encoding gene located on 8p23.1—is sufficient to cause cardiac defects in humans, and GATA4-deficient mice have both CDH and cardiac defects. Although this suggests that deletion of *GATA4* contributes to the development of both CDH and cardiac defects, CDH has not been reported in individuals with cardiac defects attributed to *GATA4* mutations. We were also unable to identify deleterious *GATA4* sequence changes in a cohort of CDH patients. This suggested that haploinsufficiency of another gene on chromosome 8p23.1 may contribute, along with *GATA4*, to the development of CDH. Of the 21 other genes located in the CDH critical region on 8p23.1, *SOX7*—which encodes a retinoic acid responsive transcription factor—appeared to be the most likely to play a role in diaphragm development. To determine if haploinsufficiency of *SOX7* contributes to the development of CDH, we generated mice with a deletion of the second exon of *Sox7*. A portion of these $Sox7^{\Delta Ex2/+}$ mice developed retrosternal diaphragmatic hernias located in the anterior muscular portion of the diaphragm. Anterior CDH is also seen in $Gata4^{+/-}$ mice and has been described in association with 8p23.1 deletions in humans. Immunohistochemistry revealed that SOX7 is expressed in the vascular endothelial cells of the developing diaphragm and may be weakly expressed in some diaphragmatic muscle cells. This pattern of expression is different from that seen for GATA4 making it unlikely that SOX7 directly regulates GATA4 expression during diaphragm development. Similar to our experience screening *GATA4*, no clearly deleterious *SOX7* sequence changes were identified in our CDH cohort. We conclude that haploinsufficiency of *Sox7* or *Gata4* is sufficient to produce anterior CDH in mice and that haploinsufficiency of *SOX7* and *GATA4* may each contribute to the development of CDH in individuals with 8p23.1 deletions. These results underscore the importance of considering the possibility that two or more genes may be contributing to individual phenotypes associated with a recurrent genomic disorder. This is particularly true when mutation screens fail to identify deleterious changes in individual candidate genes.

835F

Generation and Characterization of Humanized DUF1220 Transgenic Mice. *J.G. Keeney[1], N. Anderson[1], D. Restrepo[2], J.M. Sikela[1]*. 1) Biochemistry and Molecular Genetics, University of Colorado, Denver, Aurora, CO; 2) Cell and Developmental Biology, University of Colorado, Denver, Aurora, CO.

DUF1220 protein domains have undergone a striking recent expansion in copy number that increases generally as a function of a species evolutionary proximity to human, where the highest copy number is found (~275 copies). The increase in human represents the largest human lineage-specific copy number increase of any protein coding region in the genome. DUF1220 copy number correlates strongly with brain size (R2 = 0.98; p = 1.8 x 10^-6) and cortical neuron number (R2 = 0.98; p = .0011) across primate lineages. Copy number variations in the human 1q21.1 region, where most DUF1220 domains map, have been linked to several cognitive diseases, including deletions associated with microcephaly and duplications associated with macrocephaly. For the above reasons, DUF1220 domains are plausible candidates to underlie brain size and brain structure-related pathologies. To investigate this possibility, transgenic mice have been developed in which human DUF1220 domains have been incorporated into the mouse genome. Mice lack the recently increased primate form of DUF1220 (they have only a single ancestral copy of DUF1220 in their genome), and as a result provide an ideal background upon which to study the effects of increasing DUF1220 copy number. “Humanized” DUF1220 mice were produced by pronuclear injection of a BAC clone that contained the NBPF15 gene which encodes 6 human DUF1220 domains. Resulting transgenic mice represent the first animal model for the in vivo study of DUF1220 function and are being used to investigate the effects of DUF1220 copy number on brain structure and development using neuro-anatomical investigations and behavioral testing.

836F

Overexpression of Sox9 in keratinocytes changes hair follicle cycling in mice. *Y. Zhang[1], Z.H. Zheng[2], Y.Q. Xu[1], S.L. Zhou[2], L. Shi[1], X. Zhang[1,3]*. 1) Research Center for Medical Genomics, China Medical University, Shenyang 110001, China; 2) The Laboratory Animal Center, China Medical University, Shenyang 110001, China; 3) McKusick-Zhang Center for Genetic Medicine and State Key Laboratory of Medical Molecular Biology, Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100005, China.

(ABSTRACT)OBJECTIVE: Congenital generalized hypertrichosis terminalis (CGHT) is characterized by universal excessive growth of pigmented terminal hairs and often accompanied with gingival hyperplasia. It is one of the least well understood of genetic hair disorders. Through the study of three Han Chinese families and a sporadic case with CGHT and gingival hyperplasia, we inferred that enhanced expression of Sox9 might lead to hair overgrowth. We then use transgenic mice model to explore the potential pathogenesis of CGHT. METHODS: We generated transgenic mice overexpressing Sox9 under control of keratin 14 promoter (K14-Sox9 mice), which drives expression of a transgene specifically in epidermal cells and outer root sheath (ORS) of hair follicle (HF). RESULTS: Our preliminary phenotypic analysis in the K14-Sox9 mice showed early entry into the anagen stage of the first hair follicle (HF) cycle, suggesting an accelerated HF morphogenesis. Meanwhile, these alterations are accompanied with elongation of the guard and awl hair shafts in the skin of the K14-Sox9 transgenic mice. Histological analysis shows thicker hypodermis in the transgenic skin by 16 weeks compared with that of wild-type, concomitant with irregular HF diameter. In addition, K14/hSOX9 transgenic mice also showed increased level of β-catenin, while some differentiation markers expressed properly. CONCLUSION: In summary, our primary results characterize the simultaneous, stimulatory effect of Sox9 delivery to keratinocytes on two types of regeneration processes with transgenic mouse model. The continuous expression of Sox9 in differentiated basal layer cells results in abnormalities in morphology of HF and hair cycling, which resemble only some characteristics of the human hypertrichosis phenotypes. Expression profiling experiments are being performed to search for the molecular basis for the same phenotypes. (KEY WORDS) hair follicle cycling; Sox9; transgenic mice (ACKNOWLEDGEMENT) This work was mainly supported by the National Natural Science Foundation of China (81071283).

837F

Integration of BAC transgenesis, evolutionary conservation, 3C, DNaseHS, and chIP on chip data identifies a candidate osteoblast enhancer of *Bmp2*. *E.M. Broeckelmann, S. Pregizer, D.P. Mortlock*. Vanderbilt University, Nashville, TN.

Bmp2, a member of the TGF-β superfamily of secreted signaling molecules, is an indispensable factor for vertebrate embryonic development. Furthermore, it is known to play critical roles during initial bone formation and osteoblast differentiation as well as the maintenance of bone health during all stages of life. Genetic association studies have also linked *Bmp2* with such chronic conditions as osteoporosis and osteoarthritis, and rhBMP2 has long been FDA approved for use in fracture repair and spinal fusion. In order to advance strategies to effectively treat and/or prevent various skeletal pathologies, it is imperative that we try and elucidate the factors and regulatory mechanisms that control *Bmp2* expression in osteoblasts. Our previous BAC transgenic approaches in mice not only provided evidence for the great number and complexity of cis-regulatory elements harbored within the large gene desert surrounding *Bmp2* itself, but also precisely identified a 656bp evolutionarily conserved element (ECR1) 156kb downstream of the promoter, that exhibits osteoblast-specific enhancer function. A thorough analysis of transgene expression driven by ECR1, however, revealed that it in fact cannot recapitulate the whole range of endogenous *Bmp2* expression in osteoblasts, as it fails to exhibit any activity in cranial intramembranous bones. This implies the existence of (an) additional osteoblast cis-acting enhancer(s). Subsequent transient transgenic analysis of a *Bmp2*-containing BAC transgene from which ECR1 has been deleted further supports this notion. We have generated new Chromosome Conformation Capture (3C) results for the *Bmp2* 3' gene desert that actually indicate the strongest physical interaction of the promoter to be not with the ECR1 locus itself, but somewhat further downstream. We then performed bioinformatic analysis to correlate these results with other data that are frequently suggestive of regulatory elements. This highlighted an evolutionarily conserved element, located ~10kb 3' of ECR1, that coincides not only with a DNaseHS site in human osteoblasts, but also with histone modifications (meH3K4) and remodeling signatures (H2AZ) in murine osteoblasts and embryonic calvaria. In order to investigate its function in murine osteoblasts, this putative enhancer is now being tested in luciferase reporter and DNaseHS assays.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

838F

The role of Filamin b in skeletogenesis and BMP signaling in mice. *J. Zieba, A. Sarukhanov, M. Ivanova, A.S. Kim, A.E. Merrill, D. Krakow.* Orthopaedic Surgery, University of California at Los Angeles, Los Angeles, CA.

Filamins are ubiquitously expressed cytoplasmic actin-binding proteins that affect the cytoskeletal network and intracellular signaling pathways. In 2004, the Krakow lab discovered that mutations in the FLNB gene were an underlying cause of spondylocarpotarsal syndrome (SCT), a recessively-inherited disorder due to absence of FLNB. Individuals affected by these distinct FLNB mutations suffer from abnormally fused vertebral bodies and carpal bones, as well as short stature. While the phenotypic consequences of FLNB mutations have been identified, the molecular basis for this defect in skeletal morphogenesis remains unclear. Our lab generated Flnb knockout mice that mimic the human disease and we have now shown that FLNB is expressed during chondrogenesis and joint formation. The enhanced cartilage formation observed in knockout mice led to the hypothesis that the absence of FLNB leads to an increase in the activity of bone morphogenetic proteins (BMPs). In the chondrocytes of Flnb-/- mice, we have demonstrated an increase in the levels of phosphorylated Smad 1,5,8, as compared to control, 30 and 60 minutes after BMP2 stimulation. These levels persisted in mutant chondrocytes whereas activity in control chondrocytes had returned to pre-stimulus levels. Similar elevated levels were discovered when observing phosphorylated Erk1,2 levels in mutant chondrocytes. We have also demonstrated, following BMP2 stimulation, a prolonged p-Smad 1,5,8 translocation response with persistent levels of nuclear p-Smad1,5,8 in mutants as compared to control. Furthermore, mutant chondrocytes exhibited nuclear p-Smad 1,5,8 in unstimulated cells whereas control chondrocytes did not. This suggests that the absence of FLNB leads to constitutive activation of the BMP pathway. FLNB has a juxtamembrane/subcortical location without BMP2 stimulation; however, 30 minutes after stimulation, FLNB shuttles to the Golgi but not to the endoplasmic reticulum. This demonstrates that FLNB can traffic to the Golgi upon BMP stimulation, supporting its role in endosomal trafficking. These data show FLNB modulates and acts as an adaptor molecule for BMP signaling during skeletogenesis. Our studies identify a previously unappreciated but vital role for FLNB in the development of the human skeleton and enhance fundamental knowledge regarding mechanisms involved in progressive vertebral fusions, which occur in many skeletal dysplasias.

839F

The Role of SOX7 in Cardiovascular Development. *A. Hernandez-Garcia[1,2], M. Wat[1], M. Garcia[3], T.F. Beck[1], Z. Yu[1], R. Schwartz[5], M. Dickinson[3], B. Lee[1,4], D.A. Scott[1,3].* 1) Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA; 2) Universidad Autónoma de Coahuila, Saltillo, Coahuila, México; 3) Department of Molecular Physiology and Biophysics, Baylor College of Medicine, Houston, Texas, USA; 4) Howard Hughes Medical Institute, Houston, TX, USA; 5) Departement of Biology and Biochemistry, University of Houston, TX, USA.

SOX7 is a transcription factor encoding gene located on chromosome 8p23.1. Recurrent microdeletions of chromosome 8p23.1 that include *SOX7* area are associated with a high risk of developing cardiovascular malformations. We have shown that SOX7 is highly expressed in the endocardium of the heart and in the endothelial cells of the developing vasculature but its role in these cells remains unclear. To explore SOX7's role in mammalian development, we targeted the second exon of *Sox7* to create both standard and conditional *Sox7* knockout mice. In contrast to $Sox7^{+/-}$ mice which are viable and fertile, $Sox7^{-/-}$ embryos die around E10.5 with pericardial edema and failure of yolk sac remodeling—signs indicative of cardiac failure. A similar phenotype was seen in $Sox7^{+/-}$; $Sox17^{+/-}$ double heterozygote mice. To confirm that these abnormalities were caused by a defect in endothelial derived cells, we generated endothelial-specific *Sox7* knockout mice using a Tie2 cre transgene. $Sox7^{-/flox}$;Tie2cre and $Sox7^{+/-}$;$Sox17^{flox/+}$;Tie2 cre mice were not recovered in Mendelian ratios at weaning. Further analysis revealed that the majority of these embryos died between E10.5 and E15.5. Between these time points, mutant embryos exhibited pericardial edema, abnormal yolk sac vasculature morphology, vascular hemorrhage, atypical cardiovascular development, aberrant liver histology, and developmental arrest/retardation, recapitulating the phenotype seen in *Sox7*-null and *Sox7/Sox17* double heterozygous mice. By comparing the gene expression patterns of wild-type and *Sox7*-null embryos at E9.5 we identified a number of genes with SOX7-dependent expression including genes that play a critical role in cardiovascular development and genes expressed in the developing liver. We conclude that SOX7 expression in endothelial cells is required for normal cardiovascular development and that *Sox17* interacts genetically with *Sox7* in endothelial derived cells. These studies suggest that deletion of *SOX7* may also contribute to the development of cardiovascular malformations associated with recurrent 8p23.1 microdeletions.

840F

Novel Mouse Models Reveal the Role of RERE in the Development of Inner Ear Phenotypes Seen in 1p36 Deletions. *H. Zaveri[1], B.J. Kim[1], O.A. Shchelochkov[2], F.A. Pereira[3], A.K. Groves[1], J.S. Oghalai[4], M. Justice[1], B. Lee[1], D.A. Scott[1].* 1) Molec & Human Gen, Baylor College Med, Houston, TX; 2) Dept of Pediatrics, University of Iowa, Iowa City, IA; 3) Cell & Molecular Biology, Baylor College Med, Houston, TX; 4) Otolaryngology, Stanford School of Medicine, Palo Alto, CA.

Terminal and interstitial deletions of 1p36 are a common cause of cognitive and developmental delay with an incidence of 1 in 5000 newborns. Although 28% of individuals with 1p36 deletions also have sensorineural hearing loss, the gene(s) responsible for this phenotype have yet to be identified. *RERE* is located in the proximal region of 1p36, encodes a highly conserved nuclear receptor coregulator that is required for normal embryogenesis. In zebrafish, reduced expression of the *RERE* homolog *Rerea* leads to inner ear anomalies—fused otoliths and abnormal semicircular canals—and diminished microphonic potentials. We hypothesized that *RERE* plays a similar role in the development of the inner ear in mammals. To test this hypothesis, we used immunohistochemistry to show that RERE is expressed in critical regions of the mouse inner ear including the inner and outer hair cells and the marginal and basal cell layers of the stria vascularis. It is also present in the saccule and the semi-circular canals of the vestibular system. However, further studies of the function of RERE in the inner ear were hampered by the early lethality seen in Rere null mice ($Rere^{-/-}$), which die in utero at E9.5 due to cardiac failure. To overcome this, we generated an allelic series of RERE-deficient mice bearing different combinations of a Rere null allele and a hypomorphic allele (*eye3*) we identified in an ENU mutagenesis screen. $Rere^{-/eye3}$ mice have diminished startle responses to a 108 dB burst at 19.9 kHz emitted from a click box—a common preliminary test for severe hearing loss. DPOAE and ABR testing at P21 revealed that $Rere^{-/eye3}$ and $Rere^{eye3/eye3}$ mice have early onset hearing loss. $Rere^{-/eye3}$ mice also fared poorly in the dowel test which requires normal balance and coordination. Histological analyses and 3D micro-CT reconstructions of the inner ear showed that $Rere^{-/eye3}$ mice have normal cochlear morphology at P21 but have increased cross-sectional diameter of their semi-circular canals when compared to their wild type littermates. In preparation for future experiments, we have recently generated RERE conditional knockout mice which can be used to control the spatiotemporal pattern of RERE-depletion within the inner ear. We conclude that RERE is required for normal hearing and vestibular function and that deletion of *RERE* may contribute to the development of hearing loss and gross motor delay in children with 1p36 deletions.

841F

Xenopus laevis, a model for studying the involvement of HDAC6 in a novel form of chondrodysplasia. *K. Massé[1], C. Blanchard[2], D. Simon[2], M. Barillot[2], P. Helwi[2], C. Durand[2], P. Babin[2], A. Tingaud[2], P. Thiebaud[1], N. Thézé[1], D. Lacombe[2,3], B. Arveiler[2,3].* 1) CNRS-UMR 5164, Univ Bordeaux, Bordeaux, France; 2) Univ. Bordeaux Segalen, Maladies Rares : Génétique et Métabolisme (MRGM), EA 4576, Bordeaux; 3) CHU Bordeaux, Service de Génétique Médicale, Bordeaux.

HDAC6 belongs to the superfamily of histone deacetylases, which by altering chromosome structure have important effects on transcription and therefore on cell cycle, differentiation and developmental events. However, HDAC6 is mostly located in the cytoplasm. This enzyme deacetylates tubulin or cortactin, forms complexes with other partner proteins and is involved in a variety of biological processes, in a non-histone linked manner. We previously showed that a variant of the human HDAC6 gene segregates with a form of dominant X-linked chondrosplasia in one family (1), resulting in the overexpression of HDAC6 in the patients, due to abolition of the post-transcriptional regulation of the gene's expression by micro-RNA hs-mir-433 (2). However, the physiopathological mechanisms of the disease remain unknown. We used Xenopus laevis as a vertebrate model to decipher the functions of HDAC6. We report the expression pattern of HDAC6 during development and in the adult frog and compare it to HDAC6 potential partners. HDAC6 overexpression leads to eye and neural crest formation defects. These data bring insights into the physiopathology of the disease and emphasize Xenopus laevis as a model for human pathologies. (1)Chassaing et al. (2005) Am J Med Genet A 136:307–312. (2)Simon et al. (2010) Hum Mol Genet 19:2015–2027.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

842F

Genetic interaction between Sox10 and β1 integrins during enteric nervous system development. *Y. Watanabe[1, 2], F. Bondon-Broders[3], P. Paul-Gilloteaux[3], B. Prosper[1, 2], V. Baral[1, 2], S. Dufour[3], N. Bondurand[1, 2].* 1) genetics department, INSERM U955, IMRB, Creteil, France; 2) Université Paris-Est, Creteil, France; 3) Institut Curie/CNRS UMR144, Paris, France.

The involvement of SOX10 in Waardenburg-Hirschsprung disease (hypopigmentation, deafness, and intestinal aganglionosis) in human and various animal models highlighted the importance of this transcription factor during neural crest (NC) and enteric nervous system (ENS) development in particular. Its role is now well described and numerous interactions between Sox10 and other genes involved in ENS development have been described over the years. In particular, we showed that Sox10 interactions with Zeb2 and the endothelin-3/EDNRB signaling pathway are required for proper vagal and enteric neural crest cells survival and differentiation. However, a possible function of this transcription factor during migration and cell adhesion remained elusive. The importance of Itgb1, encoding β1 integrin, in enteric progenitor adhesion and migration and its co-expression with Sox10 during ENS development prompted us to test the possibility of an interaction between this major extracellular matrix adhesion receptor and SOX10. To this end, we crossed the $Sox10^{LacZ/+}$ mice with the conditional HtPA-Cre; $\beta1^{neo/+}$ and the $\beta1^{fl/fl}$ mice and compared the phenotype of embryos of different genotypes during NC development. The $Sox10^{LacZ/+}$;HtPA-Cre; $\beta1^{neo/fl}$ double-mutant embryos presented a more severe ENS phenotype compared to single mutants which is clearly visible as early as E11.5 and that is not compensated after birth. Using time lapse imaging we showed that defects observed are due to reduced speed of migration and altered directionality of double mutant cells, resulting in severe neuronal network disorganization. Experiments are ongoing to determine the molecular origin of these observations. Together with previous studies, our results strongly support the implication of Sox10 in enteric neural crest cells adhesion and migration.

843F

Mowat-Wilson syndrome associated *ZEB2* zinc-finger mutations lead to phenotype-correlated impairment of the transcription factor function. *J. Ghoumid[1, 2, 3], L. Drevillon[1, 2], SM. Alavi-Naini[4], M. Rio[5, 6], A. Briand-Suleau[1, 2], L. Goodwin[7], P. Raymond[1], N. Bondurand[2], C. Yanicostas[4], M. Goossens[1, 2, 3], S. Lyonnet[5, 6], D. Mowat[8], N. Soussi-Yanicostas[4], J. Amiel[5, 6], I. Giurgea[1, 2, 3].* 1) AP-HP, Henri Mondor Hospital, Creteil, France; 2) Inserm, U955 equipe 11, Creteil, France; 3) Paris Est University, Faculty of Medecine, Creteil, France; 4) Inserm, U676, Robert Debre Hospital, Paris, France; 5) Department of Clinical Genetics, Necker Enfants-Malades Hospital, Paris, France; 6) INSERM U-781, Paris, France; 7) Department of Clinical Genetics, Nepean Hospital, Sydney, Australia; 8) Department of Clinical Genetics, Sydney Children's Hospital, Sydney, Australia.

Mowat-Wilson syndrome (MWS, MIM#235730) is an intellectual disability-multiple congenital anomalies syndrome. Patients harbor distinctive facial features, severe intellectual disability, microcephaly, and Hirschsprung disease. The prevalence of MWS is estimated to be 1 per 50,000–70,000 live births. Patients carry *de novo* heterozygous *ZEB2* molecular abnormalities, that consist in point mutations or small indels (85%), and deletions (15%). So far, more than 180 different *ZEB2* mutations have been identified in MWS patients and no genotype-phenotype correlation was observed. The encoded protein ZEB2 (also called SIP1 or ZFHX1B) is a transcriptional factor, member of the two-handed zinc finger/homeodomain proteins. It comprises six functional domains, among which two zinc finger clusters localized at the N-terminus (N-ZF) and at the C-terminus (C-ZF) of the protein. Integrity of both ZF clusters is critical for recognition of its promoter gene targets. Here, we report the first missense mutations in zinc-finger domains of ZEB2. Patients presented with mild phenotype of MWS, with less marked facial features, moderate intellectual disability, and no or few congenital malformations. All mutations were located in the C-ZF (p.Tyr1055-Cys, p.Ser1071Pro and p.His1045Arg). Using *in vitro* experiments we showed that the underlying mechanism of these mutations is a loss of function and not a dominant negative effect. ZEB2 mutants failed to bind to its best-characterized target, the E-cadherin promoter, and impaired ZEB2 repressive function. Using zebrafish sip1b morphants, the most conserved ortholog of human ZEB2, we performed *in vivo* rescue tests with ZEB2 human mRNA carrying the mutations. Zebrafish rescue varied following the different *ZEB2* missense mutations and correlated with patients' phenotype. These data may improve our understanding of the role of ZEB2 zinc fingers, in the molecular pathology of MWS, and more broadly, in embryologic processes in which ZEB2 has a key role.

844F

Methylation status of a vault RNA, VTRNA2-1 (MIR886), is co-segregated with neural tube defects in Chinese Han families. *W. Li[1], C. Cai[2].* 1) Tianjin Medical University, Tianjin, China; 2) Tianjin Children's Hospital, Tianjin, China.

The etiology of neural tube defects (NTDs) is complex, although the connection between folic acid related genes and NTD is well documented, however, the epigenetic regulation of NTD remains enigmatic. We detected genomewide methylation status in 3 Chinese neural tube defects (NTDs) families. We found that the abnormal methylation of a vault RNA was co-segregated with NTDs status. DNA samples of 12 subjects in 3 families were collected. Each family had multiple NTD patients, including affected NTD sib pairs or parent-offspring pairs. Methylation status of 450,000 CpG islands was scanned in all 12 subjects, and Avg-beta values indicated the intensity of methylation signal of each loci (Illumina Genomestudio, version 2011). Methylation status of a vault RNA, VTRNA2-1, was co-segregated with NTDs. In Family A, 10 of 10 VTRNA2-1 CpG islands were highly methylated in affected siblings. In Family B, all 10 VTRNA2-1 CpG islands were hypermethylated in the affected father and proband ; the methylation level of the mother was very low. In Family C, all 10 VTRNA2-1 CpG islands were hypomethylated in both affected siblings; the father and the maternal uncle were hypermethylated on VTRNA2-1, and the mother had intermediated methylation. VTRNA2-1was transcribed by RNA polymerase III. In the general population, VTRNA2-1 is either hypo- or semimethylated in an allele-specific pattern. Previous studies suggested that VTRNA2-1might NFκB or regulates CXCL12 for cell migration. Both large-sample-size screening and functional analyses are needed to examine VTRNA2-1 methylation.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

845F

A High Frequency of Genomic Disorders in Patients with Congenital Kidney Malformations. *S. Sanna-Cherchi[1,2], K. Kiryluk[1], K.E. Burgess[1], M. Bodria[3], M.G. Sampson[4], D. Hadley[5], S.N. Nees[1], M. Verbitsky[1], V.J. Lozanovski[6], A. Materna-Kiryluk[7], B.J. Perry[1], N. Ristoska-Bojkovska[6], M. Zaniew[8], H. Flogelova[9], K. Drnasin[11], V. Goj[12], R.P. Lifton[13], L.N. Clark[14], M. Saraga[10,15], A.F. Dominiczak[16], L. Gesualdo[17], Z. Gucev[6], L. Allegri[18], A. Latos-Bielenska[7], D. Cusi[19], F. Scolari[20], V. Tasic[6], H. Hakonarson[4,5], G.M. Ghiggeri[3], A.G. Gharavi[1].* 1) Division of Nephrology, Columbia University, New York; 2) Department of Internal Medicine, St. Luke's-Roosevelt Hospital Center, New York; 3) Department of Nephrology and Laboratory on Pathophysiology of Uremia, G. Gaslini Institute, Genoa, Italy; 4) Department of Pediatrics, University of Pennsylvania, Philadelphia; 5) Department of Genetics, University of Pennsylvania, Philadelphia; 6) Division of Pediatric Nephrology, University Children's Hospital, Skopje, Macedonia; 7) Polish Registry of Congenital Malformations, Poznan, Poland; 8) Department of Pediatrics, Nephrology and Toxicology, District Children Hospital, Szczecin, Poland; 9) Faculty of Medicine, Palacky University, Olomouc, Czech Republic; 10) University Hospital Split, Department of Pediatrics, Split, Croatia; 11) Pediatric Outpatient Clinic, Solin, Croatia; 12) S.C. of Pediatrics, Fatebenefratelli Hospital, Milan, Italy; 13) Department of Genetics, Yale University, New Haven; 14) Taub Institute for Research on Alzheimer's Disease and the Aging Brain, New York; 15) University of Split, School of Medicine, Split, Croatia; 16) Institute of Cardiovascular and Medical Sciences, College of Medical Veterinary and Life Sciences, University of Glasgow, Glasgow, United Kingdom; 17) Department of Biomedical Sciences, University of Foggia, Italy; 18) Department of Clinical Medicine, Nephrology and Health Sciences, University of Parma, Italy; 19) Division of Nephrology, San Paolo Hospital and Department of Medicine, Surgery & Dentistry, University of Milan, Milan, Italy; 20) Chair of Nephrology, University of Brescia and Division of Nephrology, Montichiari Hospital, Italy.

Background: Congenital defects of the kidney and urinary tract are a major cause of pediatric kidney failure. The molecular diagnosis in the majority of these cases is unknown. Data from cytogenetic studies suggest that submicroscopic structural abnormalities may contribute to the disease pathogenesis. Methods: Genome-wide search for copy number variations (CNVs) was performed with high-density Illumina arrays. We examined the burden of large, rare CNVs in 192 patients with renal agenesis and hypodysplasia (RHD). We next searched for known and novel genomic disorders and replicated findings in 330 RHD cases from two independent cohorts (total of 522 RHD cases). Comparisons were made with 13,839 population controls genotyped with equal or higher resolution arrays. Results: CNV size was significantly larger in 192 RHD cases compared to 4,733 ethnicity-matched controls, based on every standard metric examined (e.g. average CNV size 366.1 kb in RHD vs. 197.1 kb in controls, P= 1.5 x 10-6). This excess of large CNVs was attributable to known and novel genomic disorders. Altogether, we detected CNVs diagnostic of 34 known genomic disorders, such as renal cyst and diabetes syndrome, in 55/522 RHD cases (10.5%); these disorders were present in 0.2% of 13,839 population controls (P=1.2 x 10-58). Another 32 RHD patients (6.1%) harbored large gene-disrupting CNVs that were absent or extremely rare in the 13,839 population controls (frequency <0.14%), identifying 38 potential novel or rare genomic disorders for this trait. The genomic imbalances were detected in RHD patients with and without extra-urinary tract defects. Conclusions: Up to 16.6% of patients with congenital kidney defects have a molecular diagnosis attributable to a genomic disorder. A search for genomic structural variants is indicated in this patient population to diagnose their specific genomic disorders, conduct adequate genetic counseling, and individualize medical care.

846F

Genome wide screening and proteome analysis of developing mice diastema region. *J.Y. Kim, W.J. Sohn, Y.M. Chae, Y.J. Jin, B.J. Han, Y.K. Lee, J.K. Jung, S.H. Kim, M.O. Kim, J.Y. Suh, S. Lee.* Kyungpook National University, Daegu, South Korea.

Different from humans, who have a continuous dentition of teeth, mice have only three molars and one incisor separated by a toothless region called the diastema in the hemi mandibular arch. Although tooth buds form in the embryonic diastema, they regress and do not develop into teeth. In this study, we evaluated the genes and proteins that modulate the diastema formation through comparative analysis with molar-forming tissue by genome wide screening method and liquid chromatography-tandem mass spectroscopy (LC-MS/MS) proteome analysis. From comparative and semi-quantitative proteome analysis and the microarray analysis, we identified 147 up- and 173 down-regulated signaling molecules in the diastema compared to the molar forming signaling molecules. Based on these evaluations, we selected and evaluated two candidate proteins, EMERIN and RAB7A, as diastema tissue specific markers. This study provides the first list of genes and proteins that were detected in the mouse embryonic diastema region, which will be useful to understand the mechanisms of tooth development.

847F

Foxl2, as Sox9, is involved in skeletal development through its role in cartilage maturation and bone mineralization. *L. Marcia[1], A. Puddu[1,2], M. Lovicu[1], R. Piras[1,3], D. Schlessinger[4], A. Forabosco[5], M. Uda[1], L. Crisponi[1], M. Marongiu[1].* 1) Istituto di Ricerca Genetica e Biomedica - CNR, Monserrato (CA), Italy; 2) Sardegna Ricerche, Pula (CA), Italy; 3) University of Cagliari, Cagliari, Italy; 4) Laboratory of Genetics, National Institute on Aging, Baltimore, Maryland, USA 21224; 5) Genomic Research Center, Cante di Montevecchio Association, Italy.

FOXL2 is a member of the family of forkhead domain transcription factors and if mutated, causes gonadal dysgenesis or premature ovarian failure (POF) in women, as well as eyelid/forehead dysmorphology in both sexes (blepharophimosis-ptosis-epicanthus inversus syndrome, BPES). Mice lacking *Foxl2* recapitulate relevant features of human BPES: males and females are smaller than wild type and show distinctive craniofacial morphology with upper eyelids absent. Mouse Foxl2 is required to prevent transdifferentiation of adult ovary to testis and its inducible deletion in adult ovarian follicles leads to immediate upregulation of testis-specific genes including the critical SRY target gene *Sox9*. *Foxl2* is thus responsible for female sex determination and maintenance, in contraposition with *Sox9* which leads to male determination and development. *SOX9* was first identified as a candidate gene for campomelic dysplasia (CD), a human syndrome affecting skeletal and testis development. In these patients most endochondral bones of the face fail to develop resulting in multiple defects such as micrognathia, cleft palate, and facial dysmorphia. Sox9 is also expressed in neural crest cells, and it is involved in skeletal, cartilage and bone formation. Haploinsufficiency of *Sox9* results in defective cartilage primordia and premature skeletal mineralization. Here we report the characterization of the skeletal phenotype of *Foxl2* knockout mice, hence demonstrating its role in skeletal development and strengthening its parallelism with *Sox9* function. *Foxl2* expression had been reported at 10.5 dpc in the 1st branchial arch and here we found its expression in cartilage cells involved both in endochondral and intramembranous ossification. We analyzed Foxl2 knock-out at several developmental stages: from 12.5 to 17.5 dpc, at P0, P7, 4 weeks and 6 months. By IHC we observed premature mineralization and cartilage maturation, and osteopenia. Confocal microscopy revealed a weak expression of Foxl2 in vertebrae cartilages and its co-localization with Sox9. Furthermore, down-regulation of Sox9 expression and upregulation of Col2a1 in the skull, both cartilage maturation markers, support the premature cartilage development. Overall, our results demonstrate that Foxl2, as Sox9, is thus directly involved in skeletal development and growth through its role in cartilage and bone maturation.

848F

Expression and functional analysis of MMACHC during mammalian development. *M. Moreno-Garcia, M. Pupavac, D.S Rosenblatt, L.A Jerome-Majewska.* Human Genetics, McGill University, Montreal, Canada.

Deficiencies of vitamin B_{12} or defects in its metabolism are known as inborn errors of metabolism. A series of inborn errors of vitamin B_{12} metabolism have been identified, designated *cblA-cblG*. These inborn errors can lead to abnormal development of the neural tube and other birth defects in humans, and result in elevation of either homocysteine or methylmalonic acid, or both, in blood and urine. The *cblC* disorder (OMIM 277400), in which there is both hyperhomocysteinemia and methylmalonic acidemia, is the most frequent inborn error of vitamin B_{12} metabolism and is due to mutation in MMACHC. Patients with this disease are unable to convert cobalamin into the two active forms, methylcobalamin and adenosylcobalamin, which are cofactors required by mammalian methionine synthase and methylmalonyl-CoA mutase, respectively. Clinical features of patients can include hematological, neurological, and ophthalmological findings, along with developmental delay. To elucidate the function of MMACHC and examine the expression pattern of MMACHC during mouse embryogenesis, we studied a gene-trap mouse line with a functional disruption of the mmachc gene and we determined sites of gene expression at 11.5 and 12.5 days post conception in WT mouse strain CD1 by in situ hybridization. Our hypothesis is that mice unable to produce functional MMACHC protein will have significant defects during organogenesis due to an inability to properly metabolize vitamin B_{12}. In situ hybridization analysis showed that Mmachc gene was expressed in head mesenchyme, dorsal root ganglia, heart, trachea, lung, esophagus, gut, mesonephric mesenchyme, and notochord during organogenesis in the mouse embryo. The gene-trap mouse line showed that heterozygous mutants were fertile and viable. However, homozygous embryos were only obtained at embryonic day 3.5. At the same time, we found by western blotting, that heterozygous mutants at E11.5 had decreased levels of MMACHC proteins and decrease in proliferation compared with the wild-type mice. Our results highlight that murine MMACHC is essential for early embryonic development.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

849F

Three novel human *VANGL1* gene mutations in 144 Slovakian, Romanian and German patients with neural tube defects. *O. Bartsch[1], I. Kirmes[1], A. Thiede[1], S. Lechno[1], H. Gocan[2], I.S. Florian[3], T. Haaf[4], U. Zechner[1], L. Sabova[5], F. Horn[6]*. 1) Institute of Human Genetics, University Medical Center of the Johannes Gutenberg University Mainz, Mainz, Germany; 2) Department of Pediatric Surgery, University of Cluj-Napoca, Cluj-Napoca, Romania; 3) Section of Neurosurgery, University of Cluj-Napoca, Cluj-Napoca, Romania; 4) Institute of Human Genetics, Julius Maximilians University, Würzburg, Germany; 5) 2nd Department of Pediatrics, Comenius University, Bratislava, Slovakia; 6) Department of Pediatric Surgery, Comenius University, Bratislava, Slovakia.

Human neural tube defects (NTDs) are among the most common and serious birth defects with an incidence averaging 1 per 1,000 pregnancies worldwide. NTDs are multifactorial in origin and result from defective neurulation. Neurulation is conventionally divided into a primary and a secondary phase. Primary neurulation creates the brain and nearly all of the human spinal cord, down to the upper sacral level. Defective primary neurulation typically leads to open forms of NTDs (including craniorachischisis, anencephaly and meningomyelocele). Secondary neurulation is defined as the formation of the caudal portion of the neural tube, and defective secondary neurulation results in closed forms of spina bifida, where the emerging spinal cord tissue fails to be separated from the other tissues of the tail bud (including “tethered” spinal cord). Numerous genetic and environmental factors are discussed to be relevant in the etiology of NTDs. In mice, mutants in more than 200 genes including the planar cell polarity (PCP) pathway are known to cause NTDs. It is thought that human NTDs arise in a similar fashion and heterozygous mutations in the human *VANGL1* (a gene in the PCP pathway) have been described in a small subset of patients with NTDs. We performed a mutation analysis of the human *VANGL1* gene in 144 unrelated individuals with NTDs from Slovakia, Romania and Germany and identified three heterozygous missense mutations: c.613G>A (p.Gly205Arg) with an open spina bifida (lumbosacral meningomyelocele), c.557G>A (p.Arg186His) with a closed spina bifida (tethered cord and spinal lipoma) and c.518G>A (p.Arg173His) with an unknown NTD. The c.613G>A mutation was also found in a healthy sibling. None of the mutations were described previously. Findings support that heterozygous *VANGL1* mutations represent hypomorphs or conditional mutants predisposing to NTDs and occur at a frequency of approximately 2% of open and closed spinal NTDs. The mutations (p.Arg173His, p.Arg186His, p.Gly205Arg) modified conserved regions of the VANGL1 protein and shared similarities with previously described mutants, providing further evidence for the presence of mutational hot spots in these patients. Bartsch et al., in press, Mol Syndromol.

850F

Identification of a new ENU-induced mouse model for neural tube defects. *S. Lachance[1,2], M.C. Guyot[1], M. Justice[3], Z. Kibar[1,4]*. 1) CHU Sainte-Justine Research Center, Montreal, Quebec, Canada; 2) Biomedical Science, University of Montreal, Montreal, Quebec, Canada; 3) Departement of Human and Molecular Genetics, Baylor College of Medicine, Houston, Texas, USA; 4) Departement of Obstetrics and Gynecology, University of Montreal.

Neural tube defects (NTDs) are among the most common congenital malformations in humans affecting 1–2 infants per 1000 births. NTDs are caused by a partial or complete failure of the neural tube to close during embryogenesis. The most common forms of NTDs in humans are anencephaly and spina bifida. Their etiology is a complex combination of environmental and genetic factors. In this study, we identified and characterized a novel mouse mutant, Skam26Jus that was generated by N-Ethyl-N-Nitrosuera (ENU) mutagenesis at the Mouse Mutagenesis and Phenotyping Center for Developmental Defects in Texas Medical Center in Houston. This mutant displayed a characteristic kinky or loop tail that is considered as the minimal sign if NTDs. Complementation of Skam26Jus with Looptail showed a genetic interaction between the gene mutated into Skam26Jus and the planar cell polarity gene Vangl2. This complementation led to spina bifida in 50% of embryos with a kinky or looptail appearance. Homozygosity mapping followed by targeted next generation sequencing of the candidate region will be used to identify the gene mutated in Skam26Jus. Our study describes Skam26Jus as a new mouse model for the study of human NTDs and provides an important tool for better understanding of their underlying molecular pathogenic mechanisms.

851F

Involvement of sperm microRNAs in unidentified male factor infertility. *M. G. Katz-Jaffe[1,2], B. R. McCallie[1], J. Stevens[2], W. B. Schoolcraft[2]*. 1) National Foundation for Fertility Research, Lone Tree, CO; 2) Colorado Center for Reproductive Medicine, Lone Tree, CO.

A complex population of RNA, including microRNA (miRNA), has been identified in the nucleus of mature mammalian spermatozoa which are potentially delivered to the oocyte upon fertilization. Preliminary evidence in the mouse has revealed a potential role for murine sperm miRNAs during early embryo cleavage divisions and in mediating epigenetic inheritance. However, the role of human sperm miRNAs during human embryo development remains controversial. Even if there is no functional role for human sperm miRNAs following fertilization, they may serve as a diagnostic tool for identifying male factor infertility. Current evaluation methods for male fertility use microscopic based assessments of sperm count, motility and morphology. Though informative, a normal sperm profile does not promise fertility. The aim of this study was to investigate human sperm miRNAs in relation to subsequent embryo development and potential unidentified male factor infertility. Real-time PCR was performed on sperm RNA using the TaqMan® Human MicroRNA Array (Life Technologies). Sperm samples (n=12) were collected during infertility treatment that included fertile donor oocytes to standardize for the female gamete. Profiling results indicated that 183 human miRNAs were expressed in these sperm samples, including miR-34c, which has been shown to be required for the first cleavage division in the early mouse embryo. Statistical analysis was performed comparing: A) sperm that resulted in good blastocyst development (>30%; n=6) and B) sperm that resulted in poor blastocyst development (<15%; n=6). Results revealed 18 differentially expressed sperm miRNAs. Ten showed increased expression and 8 displayed decreased expression in sperm samples that resulted in poor blastocyst development ($P<0.05$). Three differentially expressed sperm miRNAs belonged to the let-7 family (let-7b, let-7d and let-7e; $P<0.05$), a miRNA family previously identified to be associated with abnormal sperm function. Predicted target gene lists for the differentially expressed sperm miRNAs revealed principal genes involved in spermatogenesis and development. In summary, there is evidence that the expression level of miRNAs in mature human sperm could be reflecting the potential function and/or future competence of the gamete itself. Further studies to elucidate these differentially expressed sperm miRNAs and their target genes may assist in identifying unknown mechanisms associated with male factor infertility.

852F

Comparison of the osteogenic potential of adult stem cells from different sources. *R. Fanganiello[1], F.A.A. Ishiy[1], L.P. Capelo[1], M. Aguena[1], D.F. Bueno[1], B.V.P. Almada[1], M.T. Martins[2], M.R. Passos-Bueno[1]*. 1) Dept Genetics & Evolution, University of Sao Paulo, Sao Paulo, Sao Paulo, Brazil; 2) Department of Oral Pathology, School of Dentistry, University of São Paulo, São Paulo, Brazil.

Stem cells isolated from human exfoliated deciduous teeth (SHED) and from human adipose tissue (hASCs) are interesting subtypes of stem cells to be used in bone tissue engineering due to their osteoprogenitor properties and easy harvesting. In this study we aimed to compare the in vitro osteoblastogenic and the in vivo osteogenic potentials of SHED and hASCs. SHED and hASCs were each obtained from six healthy subjects. After 9 and 21 days of in vitro osteoblastogenic induction, SHED showed a 1.61-fold increase ($p<0.05$) in alkaline phosphatase activity and a 2.24-fold increase ($p<0.0001$) in matrix mineralization, respectively, when compared to hASCs. Amongst the six biomaterials screened, the best in vivo bone induction was observed with CellCeram in combination with SHED, with 8.4-fold increase in bone formation when compared to CellCeramTM free of cells. Next, we found that association of SHED to CellCeramTM induced 7.4 times ($p<0.0001$) more bone formation than hASCs under the same condition. Besides vascularization, dense and loose connective tissue were observed in all groups. Human lamin A/C immunostaining showed that SHED and hASCs induced recipient bone formation but did not differentiate directly into osteoblasts or osteocytes. These results allow us to conclude that SHED have an intrinsically higher in vitro osteoblastogenic potential and in vivo bone induction potential when compared to hASCs. This study also reinforces that the understanding and optimization of conditions to ideally match scaffolds with osteogenic stem cells will contribute to make stem cell-based strategies a better bid than the use of autologous bone substitute for rehabilitation of craniofacial bone defects.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

853F

The NINDS Repository's Publicly Accessible Collections of Highly Characterized Induced Pluripotent Stem Cells and Primary Fibroblasts. *C. Tarn[1], K. Fecenko-Tacka[2], M. O'Rourke[2], K. Panckeri[2], C. Rhoda[2], S. Heil[1], K. Hodges[3], M. Sutherland[4], R. Corriveau[4], M. Self[1].* 1) Coriell Institute for Medical Research, Camden, NJ; 2) Stem Cell Biobank, Coriell Institute for Medical Research, Camden, NJ; 3) Cell Culture Laboratory, Coriell Institute for Medical Research; 4) National Institute for Neurological Disorders and Stroke, NIH.

Human induced pluripotent stem cells (iPSCs), and fibroblasts from which iPSCs are derived, have become increasingly utilized resources for studying human disease. iPSCs may prove especially valuable in studying neurological disorders for which disease models are challenging, effective treatments are limited, and novel therapeutics are in great demand. The National Institute of Neurological Disorders and Stroke (NINDS) Repository at the Coriell Institute for Medical Research is a public resource established to provide a centralized and open collection of biological samples with associated phenotypic data to academic, non-profit, and industry investigators. Since 2011, more than 20 iPSC lines and 100 fibroblast lines from subjects harboring mutations associated with disorders such as Parkinson's disease, amyotrophic lateral sclerosis (ALS), Huntington's disease, and dementia, as well as from neurologically normal controls, have been made available via the NINDS Repository web-based catalog (ccr.coriell.org/NINDS). For a number of subjects, both the parental fibroblasts and corresponding iPSC lines are available. To ensure the quality of these valuable materials, all iPSCs and fibroblasts undergo stringent quality control testing including for viability, sterility (microbial and mycoplasma contaminations), and DNA fingerprint analyses for identity. In addition, a set of characterization criteria designed to assess and ensure pluripotency is completed for each iPSC line distributed by the NINDS Repository. These assays include embryoid body formation, cell surface antigen expression, and qRT-PCR analysis of pluripotency; G-banded karyotyping is also performed on each iPSC line to determine chromosomal integrity. The results are summarized within a document termed the Certificate of Analysis (CofA) which is displayed on the web catalog along with culturing protocols and other pertinent information. As knowledge of the applications and limitations of iPSC research emerges, it is important that relevant data and observations describing publicly available iPSCs and fibroblasts be documented and clearly presented to potential requestors. The NINDS Repository serves as an effective, centralized forum for displaying these critical data in addition to providing well characterized cell lines for both basic and applied research.

854F

Deciphering the *Hmx1* pathway in the mouse retina. *A. Boulling, D.F. Schorderet.* Institute for Research in Ophthalmology, Sion, Switzerland.

The *Hmx1* gene is responsible for the Schorderet-Munier-Franceschetti syndrome (MIM 612109), a new oculo-auricular defect identified in 2008, leading to eye developmental abnormalities and early retinal degeneration. Few information are known about the function of this transcription factor, excepted that it is expressed in eye and ear as in some other parts of the nervous system, and that it can recognize a CAAGTG binding site. Here, we identified *Hmx1* target genes in the retina, by using a comparative transcriptomic analysis of C57/Bl6, wild-type and Dmbo mice (carrying a *Hmx1* loss-of-function mutation), supplemented with the screening of the CAAGTG recognition motif (or CACTTG on the reverse strand) in the promoter of the differentially expressed genes. Comparative transcriptomic analysis was done at P15 and microarray data were analyzed with the DAVID algorithm and the Metacore software. Modulated genes were classified into relevant groups based on the adjusted P-Value. The Galaxy platform was used to generate an irrelevant random control group and to count CAAGTG motifs in the [-250,+200] region. Candidate target genes expression was checked by qPCR in mice retina at P15. Finally, the specific retinal cell type expression of each gene was correlated using the Gene expression profile database (http://www.fmi.ch/roska.data/index.php). Microarray results showed 146 modulated genes with $P<0.1$, and 30 modulated genes with $P<0.01$. First, the $P<0.1$ genes were classified into annotation clusters by means of the DAVID algorithm, underlining the enrichment of terms related to membrane and axonal proteins. Using the same genes, the Metacore software analysis revealed 3 enriched GeneGO Process Networks (Synaptogenesis, Visual perception, Synaptic contact). Then, the CAAGTG counting in the promoters of $P<0.01$ genes show 3 genes with 2 motifs (*Sgcg*, *Tshz2* and *Slc6a9*) that represents a 20-fold significant enrichment against a random control group, supporting the idea that these 3 genes are direct *Hmx1* targets. qPCR results confirmed microarray results and showed a dramatic Sgcg over-expression (about 1000-fold). Finally, we showed an overlap between *Hmx1* and candidate target genes expression in the amacrine cells, by inspection of the Gene expression profile database. In brief, *Sgcg*, *Tshz2* and *Slc6a9* represent *Hmx1* candidate targets in the mouse retina.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

855W

Evaluation of vitamin B12 effects on DNA damage induced by Paclitaxel. *O. Khabour[1], K. Alzoubi[2], M. Khader[2]*. 1) Medical Laboratory Sciences, Jordan Univercity of Science and Technology, Irbid, Jordan; 2) Clinical Pharmacy, Jordan Univercity of Science and Technology, Irbid, Jordan.

Paclitaxel is a cytotoxic drug used in the treatment of ovarian cancer, breast cancer and non-small cell lung cancer. Various studies have shown that paclitaxel could generate free radicals leading to irreversible cell injury. Recently, new studies have shown that vitamin B12 has antioxidative properties and can protect DNA from free radicals. In this study, we investigated the effect of paclitaxel on DNA by measuring sister chromatid exchanges (SCEs), chromosomal aberrations (CAs) and 8-hydroxy-2-deoxy guanosin (8-OHdG) in cultured human blood lymphocytes. In addition, we investigated the possible protective effect of vitamin B12 on chromosomal damage induced by paclitaxel. Our results showed that PAC significantly increased the frequencies of SCEs ($P < 0.001$) and CAs ($P < 0.001$) in human blood lymphocytes as compared to controls. These DNA damages were prevented by pretreatment of cells by vitamin B12. In addition, we show that paclitaxel induces increase in 8-OHdG, a marker of DNA damage, and this increase was prevented by vitamin B12. Therefore, vitamin B12 seems to protect against paclitaxel -induced DNA damage and genotoxicity.

856T

Evaluation of Vitamin B12 Effects on DNA Damage Induced by Pioglitazone. *K. Alzoubi[1], O. Khabour[2], N. Hussain[1], S. Al-azzam[1], N. Mhaidat[1]*. 1) Department of Clinical Pharmacy, Jordan University of Science and Technology, irbid, Jordan; 2) Department of Medical Laboratory Sciences, Faculty of Applied Medical sciences, Jordan University of Science and Technology, Irbid, Jordan.

Pioglitazone is a prototype of thiazolidinediones, used for the treatment of type 2 diabetes mellitus. Previous studies suggest that Pioglitazone might cause DNA damage by generation of oxidative species. In this study, we investigated the mutagenic effects of pioglitazone using sister chromatid exchanges (SCEs), and chromosomal aberrations (CAs) assays in cultured human lymphocytes. In addition, oxidative DNA damage was evaluated in cells culture by measuring 8-hydroxy-2'-deoxyguanosine (8-OH-dG) marker. We also investigated the possible protective effects of vitamin B12, which is associated with DNA repair, on DNA damage induced by pioglitazone. Treatment of the human lymphocytes with pioglitazone (100 uM) significantly increases the frequency of SCEs and CAs ($P < 0.01$). In addition, significant elevation in 8-OH-dG release from lymphocytes was observed after treatment with pioglitazone ($P < 0.01$). On the other hand, pretreatment of cultures with vitamin B12 (13.5 ug/ml) protected lymphocytes from the genotoxic effect of pioglitazone. Therefore, we conclude that pioglitazone is genotoxic, and it induces chromosomal and oxidative DNA damage in cultured lymphocytes and this toxicity is prevented by pretreatment with vitamin B12.

857W

Cytogenetic study of induced pluripotent stem cells (iPSCs) derived from senescent cells of elderly persons. *F. Pellestor[1], L. Lapasset[2], F. Becker[1], J. Puechberty[1], A. Schneider[1], S. Taviaux[1], P. Sarda[1], J.M. Lemaitre[2], G. Lefort[1]*. 1) Medical and Chromosomal Genetics, Hosp Arnaud de Villeneuve, CHRU Montpellier, Montpellier, France; 2) Laboratory of Plasticity of the Genome and Aging, Institute of Functional Genomics, Montpellier, France.

Reprogramming of adult somatic cells into induced pluripotent stem cells (iPSCs) provide a unique opportunity to generate patient-specific stem cells with potential application in regenerative therapies and without the ethical concerns of human embryonic stem cells (hESCs). However, important question persist about the safety of iPSCs, their chromosomal stability and their propensity to form tumor, in particular for iPSCs derived from senescent cells. In order to assess the chromosomal consequences of reprogramming process, we performed classical and molecular cytogenetic analysis on 4 iPSC lines generated from 74- and 96-years-old men's fibroblasts. Cytogenetic analysis was carried out using standard methods and R-banding. FISH assays with both centromeric and painting probes were used for chromosomal detection on interphase nuclei. A minimum of 20 metaphase spreads and 100 interphase nuclei per cell line were analyzed. The 4 cell lines exhibited normal karyotypes. In three iPSC lines, these investigations were completed by an analysis of the telomere sizes since senescence is characterized by telomere shortening. We measured in situ the length of telomere repeat domains using PNA pantelomeric probes and Metasystems image analyzer software Isis. The Isis software was used to quantify the fluorescence intensity of telomeres from at least 20 metaphases and nuclei from each cell line. In the 3 iPSC lines, we observed a significant increase in telomere length, when compared with their parental fibroblasts and control hESCs. These data were confirmed by telomere restriction fragments analysis (TRF). The perspective of using iPSCs in regenerative medicine is of particular interest in the context of age-associated disorders, but cytogenetic and genetic screening of iPSCs need to become standard practice before these cells might be used clinically.

858T

Secondary Aneuploidy in Samples of Balanced Chromosomal Rearrangements - No Increased Rate Detected. *O. Reish[1,2], R. Amiran[1], A. Kanesky[1], S. Gerafi[1], M. Mashevich[1]*. 1) Gen Inst, Assaf Harofeh Med Ctr, Zerifin, Israel; 2) Sackler School of Medicine, Tel Aviv University, Tel Aviv,.

Background: We previously have shown that the frequency of cells with aneuploidy for randomly chosen chromosomes was significantly higher in samples derived from trisomic and Turner's syndrome patients in comparison to samples derived from control subjects with normal karyotypes. **Aim:** To evaluate frequency of cells with aneuploidy among samples derived from healthy individuals who carry balanced constitutional rearrangements. **Methods:** 20 blood, 7 amniotic fluid and 3 chorionic villous samples carrying balanced translocations, were tested and compared with age and sex matched controls. FISH (fluorescent in situ hybridization) was applied at interphase, using randomly chosen chromosomes (#8, #15, #16), not related to the basic rearrangement. 200 cells per each chromosome in each sample were tested (600 cells per each sample). T test (two-tailed) at the 95% confidence level was applied. **Results:** Frequency of aneuploidy rate was not increased in test samples in comparison to control samples ($Pv = 0.95$) for all 3 evaluated chromosomes, with no significant difference between variably evaluated tissue types. Although monosomy rate was mildly increased in comparison to trisomy, it was not significant ($Pv>0.05$). **Conclusions:** Cells with balanced constitutional rearrangements are not associated with increased sporadic aneuploidy. This correlates with no increased risk for malignancies among carrier patients versus controls.

859W

Telomere “loss” in people with Down syndrome with and without mild cognitive impairment more than doubles after T-lymphocytes have been frozen in liquid nitrogen and short-term cultured with PHA. *E. Jenkins[1], L. Ye[1], S. Krinsky-McHale[1], W. Zigman[1], N. Schupf[1,2], W. Silverman[3]*. 1) Department of Human Genetics, New York State Institute for Basic Research Developmental Disabilities, Staten Island, NY; 2) Taub Institute for Alzheimer's Disease and Aging Research, Columbia University, New York, NY; 3) The Kennedy Krieger Institute and The Johns Hopkins University School of Medicine, Baltimore, MD.

We have previously shown that telomere PNA (peptide nucleic acid) probe (DAKO, Carpenteria, CA) signal number is significantly reduced in short-term whole blood cultures with phytohemagglutinin (PHA) from people with Down syndrome (DS) and dementia/mild cognitive impairment (MCI) in comparison to people with DS and no dementia/MCI (Jenkins et al., 2008, 2012). During the course of a study on telomere loss, we modified our method of culture to include freezing mononuclear whole blood cells, obtained via ficoll paque gradient centrifugation, in liquid nitrogen, until sex- and age-matched pairs of specimens could be obtained for parallel culture studies from individuals with DS with and without MCI. When we determined the amount of telomere [signal] loss in short-term T lymphocyte cultures after freezing, we observed a striking increase in the loss of telomeres in cultures from all people with DS with and without MCI, even though statistically significant differences persisted between samples from DS/MCI individuals and samples from DS individuals without MCI. The mean number of chromosome arms with no signal in two studies where the cells were not frozen was 12.0 (3.5) and 10.0 (3.7) for people with DS and MCI versus 5.9 (2.7) and 4.2 (2.8) for people with DS without MCI, respectively, while for two other studies that involved freezing the samples, the corresponding results were 26.4 (9.3) and 20.1 (6.2) for people with DS and MCI, and 18.8 (5.50) and 9.3 (4.3) for people with DS and no MCI, respectively. We hypothesize that the etiology is related to something in the freezing process. We are anxious to investigate this further to isolate the cause since there are so many critical processes that involve the freezing/thawing cycle that may be affected including embryo implantation. Supported in part by New York State Office for People with Developmental Disabilities Institute for Basic Research (IBR), Alzheimer's Association grants: IIRG-07-60558, IIRG-96-077; NIH grants: PO1-HD35897; RO1-HD37425, RO1-AG014673, and P30-HD024061.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

860T

Study of aging consequences in human mesenchymal stem cells during tandem passages. *H. Pour-Jafari[1,2], B. Pourjafari[2], F. Talebzadeh[2], A.R. Zamani[3], M. Hosseinipanah[4], H. Fazli[2]*. 1) Molec Med & Gen Dept, Sch Med, Hamadan Un, Hamadan, Iran; 2) Research Center for Mol Med., Hamadan University of Medical Sciences, Hamadan, Iran; 3) Immunology Dept., School of Medicine, Hamadan University of Medical Sciences, Hamadan, Iran; 4) Anatomy Dept., School of Medicine, Hamadan University of Medical Sciences, Hamadan, Iran.

Stem Cells by their unique characteristics are determined. One of which is high proliferation ability and the other is differentiation to various tissues. Every day around the world including Iran the results of new works released that show the growing momentum of works on the topic used in stem cell therapy. Working with stem cells from different origins is in progress, including the cells originated from fetal and adult origin (such as bone marrow, dental root, cord and etc.). Umbilical cord as a rich source of stem cells originated from human, without using invasive methods, is available in a variety of research using stem cells and then used for cell differentiation and subsequent therapy. On the other hand the fact that each of the mentioned cell after a few days growth and proliferation, will increase and will fill bottom of the flasks in tandem passages. As results of successive cultures do develop stem cells are aging and this caused damage to genetic material that can be evaluated in primary (passages numbers 1–5), middle (passages numbers 8–12) and late passages (passages numbers 15–18) by the Single Cell Gel Electrophoresis assay (comet assay). Also their karyotypes were examined; Solid technique and G-Banding staining for chromosome analysis in the stem cells were employed. Several microscopic slides from each passage were prepared. Then 20 cells were selected from those slides related to each passage randomly. In all selected cells, damages were examined and their degrees of damages were scored based on the standard patterns, from 0 to 4. The averages were compared in different groups. Average scores related to three studied different cell groups, primary, middle and late passages, were 0.4, 2.8 and 3.6 respectively. Prepared karyotypes from cells belonging to passages 1–5 were normal but in aged passages they were abnormal, numerically and also structurally. Our results showed that the aging phenomenon and its consequences such as DNA damages should be considered in works with tandem passages of stem cells. We concluded that the stem cells taken from human umbilical cord blood storage and passage in our center conditions just in the early passages (below passage five) of intact DNA is entitled. Therefore, for research work such cells must be used. We recommend because of the expensive materials needed for passage and maintenance of these cells, passages more than five should be used only for educational purposes in the laboratory.

861W

Telomere structure and function in Cornelia de Lange syndrome preserved despite evidence for premature aging. *A.D. Kline[1], H. Riethman[2], K. Ravichandran[2], M. Kaur[3], A. Kimball[1], I.D. Krantz[3]*. 1) Harvey Inst Human Gen, Greater Baltimore Med Ctr, Baltimore, MD; 2) Wistar Inst, Philadelphia, PA; 3) Div Hum Genetics and Molec Biology, The Children's Hosp of Philadelphia, Philadelphia, PA.

Premature aging has been associated with abnormalities in genome stability, telomere length, integrity and stability, DNA repair, and oxidative stress. Evaluations of multiple individuals with Cornelia de Lange syndrome (CdLS) through an aging clinic, and at regional and national meetings, have shown evidence for premature aging, including facial features older than expected for age with wrinkles and sagging of skin, development of prematurely grey hair, Barrett's esophagus and cholecystitis occurring at a younger age than expected, prostatic enlargement by the early 40's, early development of osteoporosis, decline in adaptive skills and development of impaired corneal reflex with some evidence for neuropathy and autonomic dysfunction. In order to pursue the etiology of the premature aging seen in CdLS, we have first undertaken evaluation of telomeres in individuals of different ages. Cohesin's role in maintenance of telomeres has been unknown, although cohesion at the telomeres is known to play a crucial role in chromosome structure and genomic stability. Telomere characteristics from 11 lymphoblast and fibroblast cell lines from individuals with CdLS with mutations in one of the genes associated with cohesin were compared to 6 age-matched control cell lines without CdLS. The telomere repeat lengths were found to be similar to those of controls, with those from blood slightly longer than from skin in both study and control cells. A telomere stability assay, in which the lengths of individual telomeres are compared using a PCR-based assay showed that there is no difference in the relative fraction of short telomeres in CdLS compared to the controls. Thus, there is no evidence that there are changes in the telomeres in CdLS with aging, including DNA breakage or damage in the telomeres, or shortening of the telomeres. DNA repair in fibroblast cells as compared to controls will be the next step in the evaluation of premature aging.

862T

Identifying Genes Associated with Hearing Loss through Chromosomal Translocations. *B.B. Currall[1,3], N.G. Robertson[1], A.M. Lindgren[1], M.E. Talkowski[2,3], C.C. Morton[1,3]*. 1) Brigham and Women's Hospital, Boston, MA; 2) Massachusetts General Hospital, Boston, MA; 3) Harvard Medical School, Boston, MA.

Chromosomal translocations, an abnormal exchange of genetic material between nonhomologous chromosomes, are estimated to be as frequent as 1 in 500 newborns and may result in abnormal phenotypes as a consequence of disruptions or dysregulations of genes. By examining the connection between chromosomal translocations and their associated abnormal phenotypes, the Developmental Genome Anatomy Project (DGAP, www.DGAP.harvard.edu) aims to identify genes that are necessary for normal human development. Through DGAP, we have identified several new genes that are putatively important for development of normal hearing. DGAP176 is a male with bilateral congenital sensorineural hearing loss, cognitive impairment, lax joints, coordination difficulties, plagiocephaly and several major craniofacial anomalies. Karyotyping reveals an inversion in the X chromosome inherited from the mother, 46,Y,inv(X)(q13q24)mat, which results in a loss of genetic material including one allele of the POU3F4, a gene known to be associated with hearing loss (DFNX2). DGAP056 is a male with profound bilateral congenital sensorineural hearing loss, mitral valve prolapse, hypospadias, early onset prostate cancer (age 38), and several minor craniofacial anomalies. Karyotyping reveals a translocation between chromosomes 2 and 13, 46,XY,t(2;13)(p24.1;q22.3), which disrupts an unannotated gene, C2ORF43. Analysis by qPCR shows a reduction in the amount of C2ORF43 transcripts. Both in situ hybridization and immunohistochemisty in mice show staining in the developing and adult inner ear structures. We have recently developed a knockout mouse model to explore further the correlation between C2ORF43 and hearing loss as well as other pathological findings. DGAP120 is a male with moderate mid-frequency hearing loss, mild learning disabilities, intermittent extropia, plagiocephaly, and minor craniofacial anomalies. Karyotyping reveals a translocation between chromosomes 6 and 11, 46,XY,t(6;11)(q23~q25;q21) or (q25;q25). Next generation sequencing of the breakpoints demonstrates that the unannotated gene C6ORF103 is disrupted. Further analysis is underway to determine the association between C6ORF103 and hearing loss. DGAP has provided tools and methods to identify and characterize genes involved in hearing loss. Such studies allow us to gain not only insight into the basic biology of the auditory process, but also to investigate new therapies for those who suffer from hearing loss.

863W

An unbalanced whole-arm translocation der(1;13)(q10;q10) in Burkitt lymphoma. *Y. Seok, S.J. Park, S.J. Kim, E.Y. Lee, J.R. Choi.* Department of Laboratory Medicine, Yonsei University College of Medicine, Seoul, Korea.

Structural rearrangements involving the long arm of chromosome 1 are common secondary aberrations in Burkitt lymphoma. Among unbalanced translocations of 1q, whole-arm translocation involving breakage and reunion of nonhomologous chromosomes at their centromeres is relatively rare. We describe here a novel unbalanced whole-arm translocation der(1;13)(q10;q10) in a case of Burkitt lymphoma. A 15-year-old male sought care for epigastric pain. Computed tomography images showed hepatosplenomegaly with innumerable tiny nodules and mesenteric lymphadenopathy. Core biopsy of the nodules confirmed Burkitt lymphoma. Bone marrow aspiration and biopsy analysis from both sides revealed normocellular marrow with significant infiltration of atypical lymphocytes, the majority of which were small to medium sized, with small amount of vacuolated basophilic cytoplasm and distinct nucleoli. Immunophenotyping of the atypical lymphocytes by flow cytometry identified B-cell phenotype with expression of CD10, CD19, CD20, CD79a, and lambda-restricted surface immunoglobulin. In addition to the primary aberration t(8;14)(q24;q32), der(1;13)(q10;q10) was detected by cytogenetic study of his bone marrow sample. By acquisition of the long arm of chromosome 1, the der(1;13)(q10;q10) resulted in trisomy 1q. The patient just recently completed induction chemotherapy with cyclophosphamide, vincristine, and prednisolone (COP). Most 1q abnormalities consist of partial duplications of 1q and have usually been reported to be associated with a poor clinical outcome in Burkitt lymphoma. However, a whole-arm translocations of chromosome 1q is unusual and so far poorly understood finding. We believe that this case represents an acquired cytogenetic abnormality which has not yet been described in Burkitt lymphoma. Keywords: Whole-arm translocation, 1q, Burkitt lymphoma.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

864T

Two Sibs with Trisomy 2q 34-qter. *Y. Wang[1], X. He[2], Y. Zhou[3], C. Wang[2], X. Yang[2], L. Zhu[2], Y. Yang[2], H. Wu[2], X. Liu[2], D. Shrestha[1], C. Lee[1], E. Carter[1], J. Martinez[1]*. 1) Pathology, University of South Alabama, Mobile, AL; 2) Cent. of Clinical Genetics, Bayi Children's Hospital, Beijing, China 100700,; 3) Biology, Baylor University, Waco, TX 76706.

The phenotype of the partial duplication at the long arm of chromosome 2, known as partial trisomy 2q, has been well described in the literature. The condition for most cases comprises of facial anomalies, visceral abnormalities and mental retardation. The phenotypic variability depends upon the breakpoints, exact portion of the duplication region involved as well as loss of another chromosome segment. We report on a brother and sister with facial anomalies and mental retardation. The clinical findings in both sibs consisted of prominent forehead, hypertelorism, low-set ears, broad nasal bridge, long philtrum, and the thin upper vermillion lip. The sister also had nystagmus, bilateral single palmar creases, and short of both 5th fingers with clinodactyly. The brother had only short of both 5th fingers. Developmental milestones of both were delayed.They also had disproportionate delay of speech. Their IQ examination showed moderate mental handicap. A brain computed tomography (CT) showed normal anatomy and they have no other visceral abnormalities. The sib's maternal uncle had similar facial anomalies and mental retardation. The patients had a karyotype with duplicated material of unknown origin in the short arm of chromosome 3. A detailed analysis of the constitutional chromosomal changes in two sibs and their parents was conducted by as microarray comparative genomic hybridization (array CGH), multiplex ligation-dependent probe amplification (MLPA) and linkage analysis using 16 short tandem repeat (STR) markers on chromosome 2. The addition of the duplicated material was found originating from the long arm of chromosome 2 (q34-qter) and size was about 31.523Mb. The parental chromosome studies showed that the mother was the carrier of a balanced translocation chromosome of 2 (q34-qter) into terminal 3p. MLPA of both sibs presented deletion of 3p CHL1 gene. The phenotype of the terminal deletion of chromosome 3p including CHL1 gene has been characterized as causing growth retardation, developmental delay, mental retardation, dysmorphisms, microcephaly and ptosis. It has been suggested that variable phenotype of patients with partial trisomy 2q results from the phenotypic expression of any accompanying chromosomal loss. Therefore, in order to better understand the phenotype-genotype relationship, it is important to employ high resolution genetic techniques, such as array CGH, to better characterize chromosomal duplications.

865W

Subtelomere sequences on human acrocentric chromosome short arms. *S. Baaj, A. Clarke, P. Thakkar, E. Chan, M. Khan, S. Chawla, J. Leya, J. Doering.* Dept. Biology, Loyola University Chicago, Chicago, IL. USA.

The 10–15% of the human genome containing heterochromatic regions has yet to be characterized. We are determining the structure of the short arm of human chromosome 21 (HC21p) as model for the organization of these regions. While the subtelomeres of most chromosomes have been extensively studied, these regions remain poorly characterized on the acrocentric p arms. Thoraval et al. (1996) identified a 580bp sequence on acrocentric p arms, as well as HC3 and HC4. Using telomere-anchored PCR, with a degenerate telomere primer and a primer specific to the 580bp sequence, they obtained a product from total genomic DNA, thus claiming that the 580 bp sequence was subtelomeric. We used the same assay with input DNAs from human chromosome-containing hybrid cell lines, and found no PCR product from distal HC21p or from any other acrocentric chromosome. PCR products were obtained from intact HC21 and total human genomic DNA. Thus the sequence organization Thoraval et al. studied is not subtelomeric on acrocentric chromosomes, but seems to occur at a number of other loci throughout the genome. Searches of several human sequence databases confirmed that the 580 bp and degenerate telomere sequences are spread throughout the genome. Regions flanked by inverted degenerate telomere sequences are also found at a number of genomic loci that are not part of subtelomeres. Using Southern blots we showed that the 580 bp sequence is part of a 6kb tandemly repetitive element that is primarily located on the acrocentric short arms. We cloned and sequenced a copy of this element from an HC21-specific phage lambda library. Two tandemly repetitive sequences exist within the 6kb repeat: an AGGA repeat and a 147bp repeat, also located at the proximal end of the rRNA gene cluster. Additionally, there are two copies of a 500 bp direct repeat in the 6 kb element that share 80% sequence identity. PCR primers for different regions in the 6kb sequence were designed to test its presence and degree of sequence conservation on HC21 and the other acrocentric chromosomes. Several clones of the PCR products were sequenced. Copies of the 6kb repeat are highly conserved (98%) on all acrocentric chromosomes and bear 96% sequence identity to repeats found in BACs from HC22, HC4 and proximal and distal HC21p. Analysis of these BACs suggests that the 6kb tandem repeats may be subtelomeric while single or partial 6kb units are present throughout the genome.

866T

Multiple chromosomal losses and gain including the deletion of 12p13 region defined by cytogenetic microarray in a case of simultaneous T-cell prolymphocytic leukemia and renal cell carcinoma. *J. Kim[1], J. Moon[1], M. Nam[1], J. Kim[2], S. Yoon[1]*. 1) Laboratory Medicine, Koure Univ. Ansan Hosp., Ansan, Gyeonggi, South Korea; 2) Laboratory Medicine, Koure Univ. Anam Hosp., Seoul, South Korea.

T-cell prolymphocytic leukemia (T-PLL) is a rare form of mature T-cell lymphoproliferative disease which occurs mainly in the elderly. T-PLL demonstrated highly recurrent submicroscopic gains and losses previously established, in regions including the *ATM* (11q22–23 region, frequency ~95%), *ETV6* (12p13, >10%), and *CDKN1B* genes (12p13, ~48%). The *CDKN1B* gene, in particular, is a tumor suppressor gene that encodes the cyclin-dependent kinase inhibitory protein $p27^{KIP1}$, which is associated with tumor size and poor prognosis in patients with renal cell carcinoma. We report a case of simultaneous development of T-cell prolymphocytic leukemia and renal cell carcinoma with multiple chromosomal losses and gain including the deletion of 12p13 region defined by cytogenetic microarray. A 55-year-old female presented with absolute lymphocytosis (66%, 15,150/μL) at a routine blood exam. On physical exam, she had neither palpable lymphadenopathy nor skin lesions. Blood biochemical analysis and serologic studies were nonspecific. A computed tomography scan showed no hepatosplenomegaly, but revealed a left kidney mass later diagnosed to be renal cell carcinoma (RCC). Examination of a bone marrow biopsy demonstrated diffuse infiltration of abnormal lymphoid cells. Flow cytometry analysis indicated that the lymphoid cells showed aberrant T cell immunophenotype with positivity for CD3, CD5, CD7, and coexpression of CD4/8. The conventional cytogenetic analysis was normal, but cytogenetic microarray showed complex chromosomal aberrancies including deletion of the 8p11.23-p23.3, 11q14.1-q25, 12p13.2, 22q11.1-q12.2, and Xp21.3-q28, and gain of the 18q21.33–22.2 and 22q12.2-q13.313. Further research needs to be done whether the chromosomal loss and gain has also contributed to malignant transformation of RCC.

867W

A Case of Therapy-related acute myeloid leukemia with t(19;21)(q13.4;q22). *M. Nam[1], J. Moon[1], J. Kim[2], J. Kim[1], S. Yoon[1]*. 1) Laboratory Medicine, Korea University Ansan Hosp., Ansan, Gyeonggi, South Korea; 2) Laboratory Medicine, Korea University Anam Hosp., Seoul, South Korea.

AML1 (= RUNX1=CBF2) is a gene essential for normal hematopoietic development. It is the most frequent target for genetic alternations in acute myeloid leukemia (AML), and *AML1* gene rearrangement is common in therapy-related myeloid neoplasm associated with topoisomerase II inhibitors. T(19;21)(q13.4;q22), a very rare cytogenetic abnormality that has only once been reported in a case of radiation-associated acute leukemia, has been previously demonstrated to act in leukemogenesis by blocking myeloid differentiation. We now present a patient in whom AML with t(19;21) developed after six years of high-level chemotherapy (R-CHOP regimen, including both alkylating agents and topoisomerase II inhibitors) for malignant lymphoma. A 64-year-old male was diagnosed with non-Hodgkin Lymphoma (diffuse large B-cell lymphoma, stage III) on 2004, when the bone marrow biopsy showed normal marrow with no involvement of lymphoma. He was treated with the R-CHOP regimen chemotherapy and autologous PBSCT at primary remission. The patient responded well to the induction chemotherapy and was asymptomatic until 4 years later, he developed pancytopenia, and the bone marrow biopsy findings done on 2009 showed normal marrow with no involvement of lymphoma (Karyotype: 46,XY). At about this time, the patient was found to have dilated cardiomyopathy. On May, 2010, the patient showed infection signs and was admitted for the treatment of pneumonia. An initial complete blood count revealed aggravation of the pancytopenia, but blasts were not seen on the PB smear. A bone marrow biopsy revealed AML, FAB type M5b. The Hemavision multiplex RT-PCR showed no fusion transcript, and no *MLL* gene rearrangement was noted on fluorescent in situ hybridization (FISH) analysis. Subsequently, conventional cytogenetic analysis showed a t(19;21)(q13.3;q22), and *AML1* gene rearrangement was confirmed by FISH analysis. Twenty-three days after the AML induction chemotherapy, the patient died from uncontrolled pneumonia and sepsis.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

868T

Detection of Donor Cells with a Clonal Abnormality 20 years after Transplantation: Is this Evidence of Donor Cell Leukemia (DCL)? *G. Velagaleti, V. Ortega, D. Means.* Dept Pathology, Univ Texas Hlth Sci Ctr, San Antonio, TX.

A vast majority of patients achieve engraftment following allogenic bone marrow transplantation (BMT) within 2 weeks. DCL is a rare complication following BMT with unknown etiology. Here, we report on a patient who showed no cytogenetic evidence of engraftment for well over 20 years following which donor cells were detected with a clonal abnormality. A 59 year old female was diagnosed with CML in 1986 and underwent an allogenic BMT from her brother in 1987. Subsequently she developed graft-versus-host disease of the eyes, mouth and skin in 1990 suggesting possible rejection of donor cells followed by relapse of CML and progression to accelerated phase in 2000. Chromosome analysis demonstrated two Ph+ chromosomes, 46,X,-X, t(1;7)(p10;p10),add(4)(p16),t(9;22)(q34;q11.2),-15,+3mar[16]/46,XX,der(9) add(9)(p24)t(9;22)(q34;q11),add(18)(p11.3),-19,der(22)t(9;22)[4]. Treatment with Gleevec was unsuccessful due to intolerance. Follow-up studies in 2006 revealed persistence of known Ph+ clones and an additional Ph+ clone. Upon treatment with dasatinib she achieved partial remission and chromosome studies in 2007 showed, for the first time, donor cells with gain of chromosome 8 (T8) along with recipient cells with the Ph+ clone. The karyotype was interpreted as, 46,XX,t(9;19)(p24;p13.1),t(9;22) (q34;q11.2),t(14;18)(q24;11.2)[2]//47,XY,+8[3]/46,XY[20]. Subsequent karyotypes demonstrated only donor cells with T8 clone in 15–35% of cells. Constitutional T8 was ruled out as the donor's blood karyotype was normal. T8 is a unique additional anomaly in 10% of CML cases frequently emerging at the time of acute transformation and may arise after imatinib treatment. Although engraftment is an immediate response, the cytogenetic evidence of donor cells did not become apparent in our case due to the proliferative dominance of Ph+ clones. Dasatinib treatment in our patient effectively removed the dominant Ph+ clones facilitating the detection of donor cells with T8. Since the etiology of DCL is unclear, we are not certain as to the significance of the T8 in donor cells in our patient and we continue to monitor the patient for high-risk secondary MDS and AML.

869W

Xp22.31 duplications: Indistinctly benign or inconspicuously pathogenic? *E.D. Esplin[1], B. Li[2], V. Cox[2], R. Clark[3], C. Curry[4], A. Slavotinek[2], L. Hudgins[1].* 1) Division of Medical Genetics, Department of Pediatrics, Stanford Univ Medical Center, Stanford, CA; 2) Division of Medical Genetics, Department of Pediatrics, University of California San Francisco, San Francisco, CA; 3) Division of Medical Genetics, Department of Pediatrics, Loma Linda University, Loma Linda, CA; 4) Division of Medical Genetics, Department of Pediatrics, UCSF Fresno, Fresno, CA.

Submicroscopic duplication of Xp22.31 has been reported as either a possible cause of development delay or a benign variant. Though multiple cases have been reported, characteristic features attributable to this duplication have yet to emerge. The pathogenicity of chromosome Xp22.31 duplications remains unclear. We report a series of 6 new cases of Xp22.31 duplication, the largest case series thus far, and offer insights as to whether this duplication is associated with a distinct phenotype. Patients were referred for genetics evaluation for developmental delay, autism, dysmorphic facies and talipes equinovarus. They include a first born female, two males with healthy siblings and a family with three siblings. The Xp22.31 duplications in these patients were defined by array CGH analysis with breakpoints at 6,461,814-8,086,161 and 6,477,036-8,057,482 approximating 1.6 Mb. One patient has two duplications (6,461,866-7,704,162 and 7,920,089-8,124,774) totaling approximately 1.2 Mb and spanning about 1.6 Mb. In all of our cases, the Xp22.31 duplication was found to be of maternal inheritance. The most consistent clinical features in our cases include: developmental delay or learning difficulties of varying severity in all 6 cases, seizures in 3, talipes equinovarus in 2 and feeding difficulties in 2. Individual patients also demonstrated autism spectrum disorder, bilateral elbow contractures, hearing loss and dysmorphic facies. One recent case report, with breakpoints similar to our series, describes a patient with an approximately 1.5 Mb duplication of Xp22.31 at 6,526,735-8,101,017, with bilateral talipes valgus, hypotonia, and developmental delay. One survey of microarray results found 23 patients with Xp22.31 duplications similar to ours. Limited clinical data were available: intellectual disability (13/23), GE reflux or feeding difficulties (6/23) and seizures (3/23). A genomic instability study reported 14 patients with Xp22.31 duplications of about 1.6 Mb, including breakpoints 6,455,604-8,109,387. Limited clinical descriptions reported developmental delay (8/14), autism spectrum (7/14), hypotonia (4/14), feeding difficulty (3/14), and seizures (2/14). The findings in our case series provide further evidence suggesting this duplication may act as a risk factor for neurocognitive disorders, seizures, equinovarus and feeding difficulties, the severity of which may be mitigated by modifiers in the genomic background.

870T

Refining the critical region associated with Jacobsen Syndrome through the identification of a small interstitial deletion at 11q23.4-q24.1 by CGH array. *S. RASKIN[1, 2], H. SALOMAO[1], F.R. FAUCZ[1], A. BONALUMI[1], J. SOUZA[1], P. GROSSFELD[3], J. ROSENFELD[4], V. SOTO-MAIOR[1].* 1) Group for Advanced Molecular Investigation (NIMA), Graduate Program in Health Sciences (PPGCS), Center for Biological and Health Sciences (CCBS), Pontifícia Universidade Católica do Paraná (PUCPR) Curitiba, PR, Brazil; 2) Genetika - Centro de Aconselhamento e Laboratório de Genética, Curitiba, PR, Brazil; 3) Division of Pediatric Cardiology, Department of Pediatrics, Rady Children's Hospital of San Diego, California, USA; 4) Signature Genomic Laboratories, PerkinElmer, Inc., Spokane, Washington, USA.

Jacobsen Syndrome (JBS) is characterized by pre- and postnatal physical growth retardation, psychomotor retardation, skull deformities, thrombocytopenia and characteristic facial dysmorphism. A subset of patients have heart, central nervous system and other malformations. It is usually caused by 11q terminal deletions, with sizes varying from 5 to 16 Mb, but more recently, interstitial deletions have been identified. We report a four-year-old Caucasian female born from healthy non-consanguineous parents, with a clinical diagnosis of JBS, based on postnatal physical growth retardation, hypotonia, psychomotor retardation, speech delay, facial asymmetry, ogival palate, thrombocytopenia and microcephaly. Oligonucleotide-based aCGH was performed using a 135K-feature whole-genome microarray (SignatureChip Oligo Solution version 2.0, Signature Genomic Laboratories, Spokane, WA, U.S.A.), and identified a “de novo” 1.4 Mb deletion at 11q23.4-q24.1 including the GRIK4, TBCEL, TECTA, SC5DL, SORL1, MIR125B1, MIR100HG and BLID genes. A second “de novo” deletion of approximately 9.3 Mb at 2q14.1q14.3 that includes the genes DDX18, CCDC93, INSIG2, EN1, MARCO, C1QL2, STEAP3, DBI, TMEM37, SCTR, TMEM177, PTPN4, EPB41L5, RALB, INHBB, GLI2, TFCP2, CLASP1, MKI67IP, TSN and CNTNAP5 was also detected, and although mutations of GLI2 have been associated with holoprosencephaly and deletions of CNTNAP5 have been implicated as a risk factor for autism and dyslexia, normal individuals with overlapping deletions have also been described in the published literature. This patient has, therefore, the smallest 11q interstitial deletion reported to date, but many features of JBS. She has no congenital heart malformation, consistent with the previously published papers implicating ETS-1 as the cardiac gene. Against previous knowledge, direct deletion of the platelet function critical genes ETS-1, FLI-1, NFRKB and JAM3 may not be crucial for thrombocytopenia, but deletion of GRIK4, TBCEL, TECTA, SC5DL, SORL1 and/or BLID and/or a positional effect may be. We are aware that careful in the interpretation of this results is needed, as we report a single case and we do not know the exact impact of the second deletion detected, so reports of further cases with similar 11q deletion size, position and gene content, but no other chromosomal abnormality are needed to confirm our findings. Our results reiterate the benefits of array-CGH to clarify genotype/phenotype associations.

871W

Triploid-diploid mixoploidy presenting with gastrointestinal complaints and growth failure. *F.I. Bhamani[1], G.L. ELsner[2], M.B. Beg[3], R.R. Lebel[4].* 1) Ziauddin University Hospital, Karachi, Sindh, Pakistan; 2) Section of Medical Genetics, SUNY Upstate Medical University, Syracuse, NY, MS2; 3) Division of Pediatric Gastroenterology, SUNY Upstate Medical University, Syracuse, NY, MD; 4) Section of Medical Genetics, SUNY Upstate Medical University, Syracuse, NY, MD.

A female infant was delivered by cesarean section at 32 weeks of gestation because of growth arrest and poor movement patterns. Her mother was a 26 year old G2P2 woman of English and Welsh ancestry who had a Wilms tumor when she was 2 years of age. The father was 41 years old with type II diabetes, hypercholesterolemia, and hypertension; his family history was notable for multiple close relatives with cancers of the breast, colon, and pancreas. The infant had feeding problems which were based on gastroesophageal reflux, laryngomalacia, and decreased gut motility. Hypotonia was notable from the outset, and she eventually displayed significant delays in both motor and cognitive milestones. Meanwhile, lymphocytes had yielded a normal karyotype (46,XX) but at two years of age she underwent a skin biopsy and mosaicism for a 68,XX cell line was discovered in fibroblasts. At five years of age, the patient is very short of stature but has a weight at the 25th percentile and head circumference at the 90th. Other phenotypic features include: low set ears, piebald irides and scalp hair, eyelid ptosis, strabismus, broad nasal bridge, anteverted nares, upswept eyebrows, hypoplastic teeth, pectus excavatum, hypoplastic labia, scoliosis, 3–4 finger syndactyly, 2–3 toe syndactyly. We present this case with a review of the literature for mixoploidy (the rare event of mosaicism for diploid and triploid cell lines). We also compare the clinical phenotype with that typical of children affected by the Russell-Silver syndrome. The complexity of the gastrointestinal problems make this case unusual.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

872T

Trisomy 14 mosaicism in a girl: Clinical cytogenetic and molecular analysis. *R. Cruz-Alcivar[1], E. Lieberman[1], C. Salas[1], P. Navarrete-Meneses[1], C. Cantú[3], K. Buiting[4], C. Durán-McKinster[2], V. Del Castillo-Ruiz[1], P. Pérez-Vera[1], Alcivar.* 1) Laboratorio de Cultivo de Tejidos, Departamento de Investigación en Genética Humana, Instituto Nacional de Pediatria, Mexico DF, México; 2) Departamento de Dermatología, Instituto Nacional de Pediatria, Mexico DF, México; 3) Genomi-k Monterrey NL, Mexico; 4) Institut für Humangenetik Universitätsklinikum Essen, Essen, Germany.

Trisomy 14 mosaicism is a rare cytogenetic abnormality, at present approximately 26 patients have been reported. These patients exhibit growth and developmental delay with variable severity, congenital anomalies, pigmentary skin lesions, and dysmorphic features. We report a 13 year old female from non-consanguineous healthy parents who had one previous miscarriage. She has one sister with hip dysplasia. At six months old she had short stature, cleft palate, face asymmetry, hyperpigmented linear streaks in arms and legs. Developmental delay, oesophagical reflux and hip dysplasia were documented. At present, she has mild facial dysmorphism, moderate mental retardation, bilateral conductive deafness and nocturnal enuresis. She has undergone palate closure, Nissen funduplication, hip myotomy, brain MRI shows arachnoid cysts. At 1 year and 2 months of age a GTG-banding chromosome analysis of peripheral blood lymphocytes was performed, revealing mos 47,XX,+mar[20]/47,XX,+14[1]/46,XX[4]. Four months later, karyotype in fibroblasts from hyper and hypogimented skin were obtained showing two cell lines, 47,XX,+mar and 46,XX. Parental karyotypes were normal. At 12 years old the patient was analyzed with a DNA array-Oligo 180k (Agilent), revealing a trisomy 14 mosaicism (arr 14q11.1q33.32(18,127,052–106,358,550)x2~3); the marker chromosome is originated from chromosome 14 and probably duplicates the proximal region between 19.5 Mb and 21.3 Mb. Metaphase FISH analysis performed with DNA BACs probes for 14q11.2 and 14q32.33 revealed the presence of three cell lines one normal (55.8%), one with +14 (9.6%) and another with +del(14q)(34.6%). Original peripheral blood slides were re-evaluated increasing the analyzed metaphases. The result obtained combining the performed analyses was: mos 47,XX,+del(14)(q11.2)[45]/47,XX,+14[10]/46,XX[45]. Patient was also analyzed for maternal uniparental disomy 14, but it was not showed evidence for abnormal methylation. In our knowledge, this is the first case of complete trisomy 14 coexisting with a cell line with partial proximal trisomy. The complete delineation of this proximal region might be useful for finding important genes associated with the phenotype of the proband.

873W

Low Klinefelter mosaicism in a male with azoospermia detected by FISH. *P. A. Delgado[1], B. Crandall[2], N. Rao[3], P. Bui[3], C. A. Tirado[3]*. 1) Department of Ecology and Evolutionary Biology, UCLA; 2) Department of Medical Genetic and Prenatal Diagnosis, UCLA; 3) Department of Pathology and Laboratory Medicine, UCLA.

Gonosomal aneuploidies such as Klinefelter syndrome (47,XXY) are the most frequent chromosomal aberrations in infertile men. Herein we present a 39 year old male with history of infertility with mild left sided gynecomastia, but without facial dysmorphic features noted. A physical exam revealed abnormally small testes, and a bilateral testicular biopsy showed absence of spermatogonia. Follicle-stimulating hormone (FSH) was 56.1 mIU/ml (very high), and his testosterone was minimally low at 247.7 (normal 249 to 836), for which he received testosterone therapy. Azoospermia, gynecomastia, very small testes and high FSH levels, low testosterone levels are all associated with Klinefelter syndrome. Cytogenetic analysis of the peripheral blood showed an apparently normal male chromosome complement in all cells examined. However, FISH studies using the XRY and the CEP X revealed a small cell population (3.6%) with two signals for the X chromosome and one signal for the SRY gene, consistent with low level mosaicism for the XXY cells. Klinefelter-like patients with a normal karyotype are usually mosaic. Our patient showed a very low level of mosaicism which may be attributed to the mosaicism in his germinal cells. The patient Klinefelter patients with low level mosaicism have been reported and may be likely to have such mosaicism in their germinal cells.

874T

A balanced translocation t(2;16)(q35;q24.1) associated with limb defects and craniofacial dysmorphism through presumptive misregulation of the *IHH* (Indian hedgehog) gene. *A.C.S. Fonseca, S. Antonini, A. Bonaldi, P.A. Otto, A.M. Vianna-Morgante.* Department of Genetics and Evolutionary Biology, University of São Paulo, São Paulo, São Paulo, Brazil.

We report an apparently balanced de novo translocation t(2;16)(q35;q24.1) in an 8-year-old girl with abnormal number and insertion of digits affecting the four extremities, club feet, and craniofacial dysmorphism. No microimbalances were detected by array CGH (180K whole-genome platform; Agilent) that could explain the phenotype. The chromosome 16 breakpoint was mapped by FISH to a 134 kb gene negative region (chr16: 86,058,168–86,192,535; hg19). The chromosome 2 breakpoint was mapped to a 40 kb "gene desert" (chr2: 220,942,287–220,983,218), ~1Mb upstream from the *IHH* (Indian hedgehog) gene (chr2:219,919,142–219,925,238), which encodes a hedgehog family signaling molecule involved in chondrocyte differentiation and cortical bone formation. Considering the phenotype of the patient and absence of direct involvement of any structural genes, we speculate that *IHH* might be involved through disruption of its regulation by the translocation. Limb and craniofacial defects have been associated with copy number variations upstream from the gene, in mouse (doublefoot mouse mutant; Babbs et al. Mech Dev 2008; 125:517) and in humans, and a putative regulator was identified ~50 kb 5' of *IHH* (Klopocki et al. Am J Hum Genet 2011; 88:70). A more distal *IHH* regulatory region might have been disturbed by the t(2;16) here described. Two enhancers, hs1507 (chr2:220,713,868-220,717,300) and hs1635 (chr2:220,771,085–220,773,279), which drive gene reporter expression in limb and/or facial mesenchyme in mouse embryos (VISTA enhancer browser), appear as potential *IHH* regulators; mapped ~200 kb proximal to the der(2) breakpoint, one or both might have become functionally altered. Alternatively, a more distally located regulatory element might have been disrupted by the breakpoint or become isolated from *IHH* through translocation to chromosome 16. Although no genes on chromosome 16 seem to have contributed directly to the phenotype, an enhancer mapping ~300 kb distal to the der(16) breakpoint (hs1;chr16:86,430,087-86,430,726), which drive gene reporter expression in the limb, cranial nerve, hindbrain and neural tube in mice embryos, was translocated to the der(2), and this might have impacted *IHH* expression. Financial support: FAPESP (CEPID 98/14254-2; 2011/ 14293-4).

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

875W

Detection of parental origin and meiotic error type in abortions and born alive with aneuploidies. *S. Frias[1,4], S.P. Mendoza-Constantino[1], S. Sánchez[1], L. Torres[1], B. Molina[1], P. Grether[2], D.G. Mayén[3], R. Meléndez[3], E. Yokoyama[1], C. Villarroel[1], E. Lieberman[1], V. Del Castillo[1]*. 1) Human Genetics, Instituto Nacional de Pediatría, D.F., Mexico; 2) Instituto Nacional de Perinatología, México D.F. México; 3) Hospital Angeles Lomas, Huixquilucan, Edo. Mexico, México; 4) Instituto de Investigaciones Biomédicas UNAM, D.F. México.

Aneuploidies have been identified in at least 5 percent of clinically recognized pregnancies and are the leading genetic cause of spontaneous abortions and birth defects. To date there is no a known reason why an aneuploid fetus continues to develop or is aborted. The meiotic error type and parental origin of aneuploid chromosome could be factors of lethality or fetal survival. We detected the parental origin of aneuploidy and meiotic error in 14 abortions and 31 live born with trisomy of chromosomes 13,18,21 and monosomy X. A segregation analysis using at least five short tandem repeat (STR) markers per chromosome was made with Aneufast Kit for QF-PCR in an ABI PRISM 310 Genetic Analyzer. A comprehensive analysis of all STRs on each chromosome involved in aneuploidy was made. The allele more proximal to the centromere was considered as the main marker to determine the type of meiotic error. For monosomy X, we determined the origin of the chromosome X and therefore nondisjunction occurred in the parent that alleles absent belong to. To determine the meiotic error type, we compared all the alleles in three chromosomes involved in trisomy. In the case of monosomy X, it was not determined. (Table 1). The maternal origin and MI nondisjunction, were the most frequent in samples of live born as reported in the literature. The error in MII was more frequent in abortions than in live born, although this should be confirmed by increasing the number of cases studied. If so, one might assume that the presence of two identical chromosomes, generates more areas of homozygosis than in type MI errors and this could have deleterious consequences and lead to spontaneous abortion.Project funded by FONCICYT 95419.

TABLE 1 T13 live-born	T18 live-born	T21 live-born	45,X live-born	T13 miscarriage	T18 miscarriage	T21 miscarriage
Parental Origin 6mat/ 0pat	3mat/ 0pat	10mat/ 1pat	2mat/ 9pat	1mat/ 2non deter	5mat/ 3non deter	3 non determine
Meiosis error 4 MI/ 2MII	2 MI/ 1MII	8 MI/ 2MII	—	2MI/ 1MII	5MI/ 3MII	2MI71MII

876T

Non-mosaic trisomy 9: A new case report. *T. Greer[1], B. Powell[1], N. Hauser[1], J. Shen[1], L. Mahon[2], R. Owen[2], T. Sahoo[2], A. Anguiano[2], F. Boyar[2]*. 1) Med Gen/Metabolism, Children's Hosp Central CA, Madera, CA; 2) Quest Diagnostics, San Juan Capistrano, CA.

Background: Trisomy 9 is a chromosomal abnormality rarely seen in live-born infants. It is believed non-Mosaic Trisomy 9 is ultimately lethal, and surviving infants with Trisomy 9 represent mosaicism. Previously, conventional karyotype was used to evaluate patients to determine if genetic mosaicism was present. Microarray technology is now able to detect low levels of mosaicism, providing a better method to truly determine if mosaicism is present. Case Report: The infant was born to a 23yo G3P1 female. Pregnancy was complicated by abnormal ultrasound findings showing bilateral ventriculomegaly and premature closure of the ductus. The mother declined aneuploidy screening. Born at 39 weeks, the infant required intubation for stridor and poor respiratory effort. Physical features included: scaphocephaly, frontal atrophy with deep forehead creases, microphthalmia with small palpebral fissures, broad nasal bridge, short philtrum, small mouth, retrognathia, low-set crumpled ears, loose nuchal skin, poorly developed phallus, and undescended testicles. Lower extremities were breech with retroflexion of the knees, feet bilaterally were dorsiflexed, with excessive plantar creases, and prominent calcanei. Upper extremities revealed broad thumbs with short fourth metacarpals. Abdominal ultrasound revealed a left bladder diverticulum. Brain ultrasound revealed partial agenesis of the corpus callosum, no ventriculomegaly. Chest radiographs revealed congenital scoliosis with posterior rib fusions at T5-T6 and vertebral fusions at T11-T12, with a right sided pneumothorax. An ophthalmologic examination was performed showing right Peter's anomaly with left clinical anophthalmia/synophthalmia. Genetic Testing: Quest Diagnostics performed an oligo-SNP assay using a microarray containing over 2.67 million probes, (1.9 million copy number probes, 750 thousand SNP probes). Analysis of the child's blood revealed arr 9p24.3q31.3(353,113-113,023,818)x3, suggestive of non-Mosaic Trisomy 9. To verify, 100 interphase and 50 metaphase cells were studied by FISH using chromosome 9 centromere and 9p21 region probes (Abbott Molecular/ Vysis; CDKN2A and D9Z1). All cells showed Trisomy 9. Conclusion: In this report we describe an infant found to have non-Mosaic Trisomy 9 by microarray analysis on leukocytes with no evidence of mosaicism. This report adds to the previous reports describing characteristics of infants born with this extremely rare condition.

877W

A Unique Phenotype of 2q24.3–2q32.1 Duplication: Infantile-Onset Epilepsy without Mesomelic Dysplasia. *B. Lim[1], J. Ko[1], J. Choi[2], M. Woo[2], J. Chae[1]*. 1) Department of pediatrics, Seoul Natinoal University Hospital, Seoul, South Korea; 2) 3Institute of Reproductive Medicine and Population, Medical Research Center, Seoul, South Korea.

The voltage-gated sodium channel genes and HOXD genes are clustered on chromosome 2q, and duplication of this region is associated with 2 clinical phenotypes: early-onset epilepsy and mesomelic dysplasia Kantaputra type, respectively. We report a case involving 2q24.3–2q32.1 duplication encompassing both the voltage-gated sodium channel and HOXD gene clusters, which were detected by a comparative genomic hybridization array. The associated clinical features were early infantile-onset epilepsy, hypoplastic left heart syndrome and global developmental delay. However, no features of mesomelic dysplasia were found. A fluorescent in situ hybridization study showed that the noncontiguous insertion of the duplicated chromosome 2q segment into chromosome 6q was inherited from the father, who has a balanced insertional translocation. The unique genotype-phenotype correlation in the present case suggests that dosage-sensitive effects might apply only to the voltage-gated sodium channel genes.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

878T

Importance of chromosome breakage testing in Fanconi's anemia - involvement of specific chromosomes. *V. Mohan[1, 2], G. Archana[2], T. Preethi[2]*. 1) Genetics, Vasavi Medical Research Centre, Hyderabad, India; 2) Cytogenetic and molecular laboratory, Tapadia Diagnostic Centre, Hyderabad, India.

Chromosomal breakage syndromes are a group of genetic disorders that include Ataxia telangiectasia, Bloom's syndrome, Fanconi's anemia. Fanconi's anemia (FA) is a paradigm of aplastic anemia and has the similar features of pancytopenia but is characterized by spontaneous or induced chromosomal instability and a variety of congenital anomalies. A cytogenetic breakage study is essential for the diagnostic differentiation between FA and the idiopathic aplastic anemia. Many cases of FA are not diagnosed in a timely manner. FA should be tested in any infant born with the thumb and arm abnormalities, with signs of progressive bone marrow failure, in cases of adult-onset aplastic anemia and elevated hemoglobin F levels, solid tumors presented at an early age in the absence of other risk factors or even with unusually toxic reactions to chemotherapy or radiation. It is absolutely essential to rule out FA before contemplating bone marrow transplantation for aplastic anemia. The regimen used to prepare patients for transplant is very different for FA patients as FA patients tolerate radiation and chemotherapy very poorly. In culture, exposure to claustogenic agents such as mitomycin C (MMC) induces a higher rate of chromosomal abnormalities in patients with FA than in normal individuals. Spontaneous chromosomal breaks in these patients are also higher than normal. The adequate dose of MMC used in our tests for diagnosis of FA was 25 ng/ml. It is important to establish the adequate dosage and time of exposure to the claustogenic agent using known positive cases and normal controls. An sex/age matched normal control as well as a culture without MMC from the patient is used to determine the baseline breakage rate. Cytogenetic analysis was performed in samples from 12 FA patients in our laboratory who were tested under MMC stress. We present the results of G -banded chromosomes from these samples indicating specific chromosomes involved in damage.

879W

Collaborative Multicenter Cytogenetic Study of Microdeletion Syndromes - Brazilian Network. *M. Riegel[1,2], N. Barcellos[3], R. Mergener[1], R. Gus[1], R. Giugliani[1,2,3], Brazilian Network members.* 1) Medical Genetics Service, Hospital de Clínicas de Porto Alegre-RS, Brazil; 2) Post-Graduation Program in Genetics and Molecular Biology, Universidade Federal do Rio Grande do Sul, Brazil; 3) Post-Graduation Program in Medical Sciences, Universidade Federal do Rio Grande do Sul, Brazil.

It has been estimated that the frequency of microdeletion syndromes (MS) is over 1:1000 liveborns. Although the clinical picture of many MS is well defined, the diagnostic facilities are available only in few centers of the public health system in Brazil; systematic studies are limited to some specific common microdeletions, and the actual overall frequency of these syndromes in Brazil has not been reported. With the main goal of joining efforts for the diagnosis and research in MS we proposed a collaborative multicenter study with the following goals: 1) To build a database for the diagnostic studies and information (abstracts, publications) performed by Brazilian research groups, 2) To support the cytogenetic diagnosis of suspected cases and; 3) to support the development of public cytogenetic laboratories. The RedeBRIM (Brazilian Network of Reference and Information in Microdeletion Syndromes) is formed by a coordinating center (SGM/HCPA) and 25 participating centers which operate in the Brazilian Public Health System (SUS). Communication with the centers is coordinated by the administrative basis of the project (Molecular Cytogenetics Lab/HCPA). We hope that this initiative will allow the development of a database for reference and information on MS; to contribute to the knowledge, assistance and research in the area of human cytogenetics and to lead to the stablishment of an organized network for research and diagnostics of microdeletion syndromes involving public health services in Brazil. Financial support : CNPq/National Council for Scientific and Technological Development - Brazil.

880T

identification of chromosome aberration in fragile x syndrome patients in tamil nadu population south India using cytogenetic techniques. *A. Sundramoorthy.* Bharathiar University, Unit of Human Genetics, Department of Zoology, Coimbatore, Tamil Nadu, India.

identification of chromosome aberration in fragile x syndrome patients in tamil nadu population south India using cytogenetic techniques arun s1*, balachanadar v1, sasikala k1, manikantan p1, arun m1, karthick kumar a1, mustaq ahamed s1 1human genetics laboratory, department of zoology, bharathiar university, coimbatore - 641 046, tamil nadu, india objective: Fragile X syndrome is the most common inherited form of X-linked mental retardation (XLMR). The aim of this study was to screen for and identify fragile X syndrome-affected individuals using Cytogenetic approaches at a special education school for the mentally retarded in Tamil Nadu, South India. materials and methods: 3 micro litres of blood sample was collected from 45 individuals with mental retardation, including 20 males and 25 females, was analyzed using Cytogenetic techniques and Microscope coordinates of all digitized G-banded preparations were recorded so that the metaphase cells analyzed by G-banding technique. results: A total of 75 individuals with mental retardation were in the category of mild to moderate the syndrome. Sixteen males (6.25%) and two female (0.75%) were found to express fragile X site at q27.3. The frequency of fragile X-positive cells in males ranged between 7% and 57%, while in the females it was between 14% and 20%, respectively. conclusion: We established a cytogenetic report the incidence of fragile X syndrome in a mentally retarded population in Tamil Nadu south India. The frequency of fragile X positive cases found in this study is similar to other reports of fragile X syndrome in preselected patients. The Cytogenetic, molecular study and genetic counselling for other family members will continue in the future research. Keywords: Chromosome aberration test, DNA damage, Lymphocytes; Fragile x syndrome *Corresponding author: Arun S, E-mail: asmarun@gmail.com.

881W

A Defect in the TUSC3 Gene is Associated with Syndromic Intellectual Disability. *A. Battaglia[1], L. Bernardini[2], S. Loddo[2], V. Doccini[1], V. Parisi[2], A. Novelli[2], T. Filippi[1]*. 1) Dev Neurosciences, Stella Maris nst/Univ Pisa, Calambrone, Pisa, Italy; 2) Mendel Lab. IRCCS Casa Sollievo della Sofferenza Hospital San Giovanni Rotondo (FG), Italy.

Defects in the TUSC3 gene have been identified in nonsyndromic autosomal recessive mental retardation (ARMR) individuals, due to either point mutations or intragenic deletions. We report on a boy with a homozygous microdeletion 8p22, sizing 203 Kb, encompassing the first exon of the TUSC3 gene, detected by SNP-array analysis (Human Gene Chip 6.0; Affymetrix). Both non-consanguineous parents, coming from a small Sicilian village, were heterozygous carriers of the microdeletion. The propositus had a few dysmorphic features (mid-parietal whorl; mildly prominent forehead; hypoplastic philtrum; mildly anteverted ears; hirsutism on limbs and dorsum; small hands; superficial venous reticulum; left cryptorchidism), and a moderate cognitive impairment. Verbal communication was impaired, with an inappropriate phonetic inventory, important phono-articolatory distortions, and bucco-phonatory dyspraxia. Comprehension was possible for simple sentences. Behavior was characterized by motor instability, high tendency to irritability and distraibility, anxiety traits, and an oppositional-defiant disorder. His parents were of normal intelligence. TUSC3 is thought to encode a subunit of the endoplasmic reticulum-bound oligosaccharyltranferase complex that catalyzes a pivotal step in the protein N-glycosylation process. This is the fourth case with alteration of TUSC3 causing ARMR and the third family in which a deletion has been described. Although the pathogenic mechanism has not been clarified yet, our report argues for a more prominent role of TUSC3 in the etiology of intellectual disability and for the possibility that deletions encompassing this gene could be more common than expected.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

882T

An apparently balanced translocation t(5;6)(q35.2;p22.3) disrupting the *CDKAL1* gene in a girl with clinical features of Silver-Russell syndrome. *A. Bonaldi, S.A.S. Fonseca, A.C.S. Fonseca, A.M. Vianna-Morgante.* Department of Genetics and Evolutionary Biology, University of São Paulo, São Paulo, São Paulo, Brazil.

Silver-Russell syndrome (SRS) is characterized by severe intrauterine and postnatal growth retardation associated with a typical small triangular face and other variable features. Most frequently, SRS is caused by altered gene expression on chromosome 11p15 due to hypomethylation of the telomeric imprinting center that is present in about 40% of the patients. Maternal uniparental disomy of chromosome 7 (matUPD7) accounts for 5–10% of SRS cases. Thus, about half of the SRS patients remain without a causative diagnosis. We report on a girl with the clinical diagnosis of SRS carrying a de novo apparently balanced translocation t(5;6)(q35.2;p22.3). She was born at term, with 2,300 g and 42 cm (<3rd centile), her stature and weight being below the 3rd centile until her last clinical evaluation at 3 8/12 years of age. No (epi)genetic defect on 11p15 or matUPD7 were found. Array-CGH analysis (180K whole-genome platform; Agilent) did not detect any micro-imbalances that could explain her phenotype. Chromosome 5 breakpoint was mapped by FISH to a ~93 kb gene negative segment (chr5:173,684,023–173,777,440; hg19). On chromosome 6, the breakpoint was mapped to a ~152 kb segment (chr6:20,820,370-20,972,522) within the *CDKAL1* (cyclin-dependent kinase 5 regulatory subunit associated protein 1-like 1) gene (chr6:20,534,688-21,232,634). *CDKAL1* encodes a highly conserved 579-residue protein that contains the conserved domains of the MTTase family. Variants within *CDKAL1* intron 5 have been reproducibly associated with elevated risk for developing type 2 diabetes, and with decreased insulin secretion and reduced birth weight. Our finding that one of the translocation breakpoints disrupted *CDKAL1* prompted our screening for *CDKAL1* point mutations in SRS patients negative for (epi)mutations at 11p15 and matUPD7. In a small group of 30 patients no potentially pathogenic variants were detected. In addition to the possible phenotypic effect of *CDKAL1* haploinsufficiency, the translocation might have impacted the regulation of developmental genes in the vicinity of breakpoints (e.g. *SOX4* on chromosome 6; *MSX2* on chromosome 5). Interestingly, the enhancer hs1340 (VISTA Enhancer Browser) is localized in *CDKAL1* intron 9, in the breakpoint region, and might have been disrupted or isolated from its target. Financial support: FAPESP (CEPID 98/14254-2; Student fellowship 2011/12486-0).

883W

Report of the first case of robertsonian translocation in Down-Turner mosaicism (mos 45, X / 46,XX, + 21, rob (21;21)(q10;q10)) with clinical evolution. *M.F. Carvalho[1,2,3], E.F. Carvalho[4], K.M Carvalho[1].* 1) Unidade de Genética Médica, Universidade Estadual do Ceará-UECE/APAE-CE, Fortaleza, Ceará, Brazil, MD, PhD; 2) Unidade de Genética Médica, Christus, Fortaleza, Ceará, Brazil, MD, PhD; 3) Genpharma, Fortaleza, Ceará, Brazil,MD; 4) Unidade de Genética- Instituto da Criança do Hospital das Clínicas, Universidade de São Paulo,MD.

Double aneuploidy involving both autosomal and sex chromosomes is very rare. Down's/Turner's mosaic, occurs in about 1 in 2 000 000. We report the first case of Down's/Turner's mosaic with robertsonian translocation. The patient was the first child of non consanguineous parents. It was a female whom born a term, by uneventful cesarean section, weighting 2,220g and with length of 46 cm, without neonatal complications. At birth, the pediatrician made the diagnosis of Down syndrome. At 3 months, in consultation with routine pediatric a heart murmur was heard, so the patient was referred to a cardiologist. Echocardiography was performed and diagnosed congenital heart disease (patent ductus arteriosus). There were performed two surgeries to repair the heart defect, one with 6 months of life and another with 2 years. She also had mild developmental delay. At 3 years old she was examined at our outpatient Genetic unit. The patient showed more clinical findings of Down syndrome than Turner syndrome: low weight and height for age, microcephaly, flat facial profile, upslanting palpebral fissures, epicanthal folds, short nose with depressed nasal bridge, hypotonia with tendency to keep month open and protrude the tongue, short neck, single palmar creases, and prominent ears. Abdominal ultrasound and thyroid function were normal. Cytogenetic analysis of peripheral blood preparations by using G-banding revealed mosaicism with 2 cell lines (mos 45, X [21] / 46,XX, + 21, rob (21;21)(q10;q10)[9]). Additional genetic studies (karyotypes) are required to define the cause which probably originated this double aneuploidy with this translocation. So we present the first case related of Down-Turner mosaicism with robertsonian translocation and we review the previous reports of Down-Turner syndrome.

884T

Copy number variation in imprinted regions: Narrowing critical imprinted intervals. *B.C. Ballif[1], N.J. Neill[1], J.E. Fox[2], M. Descartes[3], F. Brewer[3], T. Stroud[4], J.B. Ravnan[1], L.G. Shaffer[1], J.A. Rosenfeld[1].* 1) Signature Genomic Laboratories, PerkinElmer Inc, Spokane, WA; 2) Division of Medical Genetics, Steven and Alexandra Cohen Children's Medical Center of New York, Manhasset, NY; 3) Department of Genetics, University of Alabama at Birmingham, Birmingham, AL; 4) Division of Developmental Pediatrics, University of Missouri School of Medicine, Columbia, MO.

While it is known that uniparental disomy (UPD) of certain chromosomes with imprinted genes is a mechanism of disease, these imprinted genes can be disrupted in other ways, including variation in their copy number, which can also cause similar phenotypes. With the increasing use of single nucleotide polymorphism (SNP)-based microarrays, which have the ability to detect segments of uniparental isodisomy, a more precise understanding of the critical imprinted regions and genes is necessary for accurate diagnostic interpretations and phenotype predictions. We surveyed our database of over 53,000 probands studied by microarray-based comparative genomic hybridization (aCGH) and identified over 50 individuals with small (<1.5 Mb) copy number variations in or near imprinted genes in 10 distinct chromosomal regions. We gathered additional clinical information on three individuals with small deletions involving the imprinted region on 14q32, two of whom have the previously described, recurrent 1.1-Mb deletions of 14q32.2q32.31. Using SNP arrays, we determined the parent of origin for cases in which parental specimens were available. The parental origins correlated with the individuals' phenotypes, with a 1.1 Mb deletion on the paternal chromosome in one case being associated with a UPD(14)mat-like phenotype including precocious puberty, feeding problems, and growth and developmental delay. In a second patient, the deletion was 120 kb in size and included a portion of RTL1/RTL1as, maternally expressed MEG8, and one of the maternally expressed snoRNA clusters in the region; paternally expressed DLK1 and maternally expressed MEG3 were not deleted. The clinical features of this patient with the smallest deletion included a bell-shaped thorax, polyhydramnios, and rib abnormalities, which are consistent with a UPD(14)pat-like phenotype and over expression of RTL1. The potential influence of reduced MEG8 and snoRNA expression on this patient's phenotype is uncertain. To our knowledge, this represents the smallest maternal deletion of this region and further refines the clinical features associated with deletion of these imprinted genes.

885W

Detection of Partial Tetrasomy 12pter-12p11.22 Region by Array CGH and FISH in a Newborn with Multiple Congenital Abnormality. *S. Ebrahim[1,2], D. Stockton[3], AM. Jay[3], M. Hankerd[2], D. Schloff[2], M. Kristofice[2], AN. Mohamed[1,2].* 1) Department of Pathology, Wayne State University School of Medicine, Detroit, MI; 2) Detroit Medical Center University Laboratories, Cytogenetics laboratory; 3) Division of Genetic & Metabolic Disorders, Children's Hospital of Michigan. USA.

We report on the molecular detection and characterization of partial tetrasomy of chromosome 12p13.33–12p11.22 region in a 4-day-old girl without an obvious clinical diagnosis of Pallister-Killian syndrome (PKS). Her prenatal course involved an abnormal ultrasound showing short femurs and flat facial features. An amniocentesis was suggested but was not done. The infant was born at 39 and 3/7 weeks of gestation to a 19-year-old gravida 2, para 2 mother via cesarean section. At birth dysmorphic features, an imperforate anus, rectovaginal fistula and tethered cord were noted. Blood chromosome analysis and oligonucleotide-based array comparative genomic hybridization (array-CGH) was performed. High resolution chromosome analysis of PHA stimulated peripheral blood culture revealed a normal 46,XX female karyotype in all cells examined. Array-CGH detected a 29.23 Mb copy number gain involving chromosome 12p11.22–12p13.33 short arm region. FISH using DNA probes specific for chromosome 12p13 region and CEP-12 alpha-satellite probe localized to chromosome 12p11.1~q11 centromere region was performed on three sample types. The hybridization demonstrated four copies for 12p13 in 6.5%, 36%, and 86% on cultured and uncultured blood samples and buccal mucosa interphase cells respectively while only two copies of CEP12 were present in 100% of cells. This demonstrates the patient to be mosaic tetrasomic for part of chromosome 12 short arm and disomic for CEP 12 chromosome/no extra CEP-12/alpha-satellite DNA, likely due to neocentromere formation (analphoid marker chromosome formation) or due to duplication within the same chromosome. A skin biopsy was requested for chromosome analysis to clarify the abnormalities. PKS is a clinically characteristic rare genetic disorder caused by tissue-limited mosaicism for tetrasomy 12p, normally as a result of the presence of a 12p isochromosome. Few patients with mosaic partial tetrasomy 12p and variable PKS phenotypes have been identified. This case demonstrates the effective use of array CGH in conjunction with FISH for the diagnosis of PKS in peripheral blood samples in a newborn without an obvious clinical diagnosis of PKS.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

886T

Duplication Xq28 and Deletion of SHOX in a family with short stature, intellectual and developmental disability. *E.F. Andersen[1,2], E.E. Baldwin[3], S. Ellingwood[4], R. Smith[4], A.N. Lamb[1,2]*. 1) Cytogenetics Laboratory, ARUP Laboratories, Salt Lake City, UT; 2) Department of Pathology, University of Utah, Salt Lake City, UT; 3) Technical Operations, ARUP Laboratories, Salt Lake City, UT; 4) Department of Pediatrics, Division of Genetics, Maine Medical Center, Portland, ME.

Copy number changes on the X chromosome may be associated with developmental delays and intellectual disability (DD/ID) in a manner specific to both sex chromosome complement and gene content. We report a familial case of a 774 kb Xq28 duplication (154,459,273–155,233,730) and a 570 kb Xp22.33 deletion (168,546–738,792) detected by CMA. The proband, a 20-month-old female, presented with failure to thrive, short stature, gross motor and speech delays, and mild dysmorphic facial features. Two brothers, ages 12 and 16 years, also with short stature, dysmorphic features and a history of developmental delays, were also noted to have intellectual disability and ADHD. The 12-year-old also exhibited aggressive behaviors and was diagnosed with a mood disorder. In addition, the 16-year-old had a 722 kb gain of 2q33.3, encompassing 9 genes, of unknown clinical significance. These findings, concomitant deletion and duplication of the terminal ends of a chromosome, are suggestive of a recombination event in a balanced pericentric inversion or inheritance of a recombinant X chromosome. Parents were not tested; however the pattern of inheritance is suggestive of maternal transmission. The Xp22.33 deletion encompasses 5 genes within the PAR1 region including *SHOX,* a gene known to cause short stature when haploinsufficient. The duplicated region includes the PAR2 region, which is not known to be dosage sensitive, and 14 proximal genes, which individually have not been associated with functional disomy. El-Hattab *et al*. (*J Med Genet*, 2011) report a recurrent interstitial Xq28 duplication that partially overlaps with the duplication detected here, associated with intellectual disability, behavioral abnormalities, and characteristic facial features in males and in females with unfavorably skewed X chromosome inactivation; many of these features were also noted in the present family. Minimally, the region of overlap of these duplications is 101 kb and includes *RAB39B*, the 3' end of *VBP1* and part of *CLIC2*. Hemizygous mutation of *RAB39B*, a neuronally enriched small GTPase involved in vesicular trafficking, is associated with ID, autism, epilepsy, and macrocephaly (Giannandrea *et al*., 2010); however, functional disomy of *RAB39B* has not been established. We suggest that duplication of *RAB39B* may be responsible for the dysmorphic features, DD/ID, and other features observed in patients with duplications of the distal region of Xq28.

887W

Deafness and Infertility Syndrome (DIS): an under-diagnosed cause of hearing loss? *U. Aypar[1], N. Hoppman-Chaney[1], D. Babovic-Vuksanovic[2]*. 1) Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN; 2) Department of Medical Genetics, Mayo Clinic, Rochester, MN.

Hearing loss is the most common birth defect and the most prevalent sensorineural disorder in developed countries. More than 50% of prelingual deafness is genetic, most often autosomal recessive and nonsyndromic, of which approximately 50% is caused by mutations in GJB2 and GJB6. Sensorineural hearing loss and male infertility (Deafness-Infertility Syndrome; DIS) is a contiguous gene deletion syndrome resulting from homozygous deletion of the CATSPER2 and STRC genes on chromosome 15q15.3. Females with DIS have only hearing loss and are fertile. DIS has been recently described in three consanguineous families and only one patient from a nonconsanguineous family, implying a low carrier frequency. Recently, a patient with hearing loss presented to the medical genetics department at Mayo Clinic. Due to the presence of macrocephaly in addition to hearing loss, she was tested by array comparative genomic hybridization (array CGH) and was found to be homozygous for deletion of both CATSPER2 and STRC, consistent with a diagnosis of DIS. Her parents are not consanguineous, and FISH testing demonstrated that both were heterozygous carriers, ruling out uniparental disomy. Upon receiving this result, we retrospectively examined our internal database of patients tested by array CGH in order to determine the carrier frequency of this deletion in our patient population. We discovered that 56 of our 5151 patients (1.09%) are heterozygous for this deletion, excluding one patient who had a sibling with the same deletion. If this number is representative of the general population, this implies a 1% carrier frequency in the general population and a prevalence of DIS of 1 in 40,000 individuals. These results are consistent with the studies performed by Knijnenburg et al. in which they determined the carrier frequency of this deletion in a much smaller sampling of normal individuals to be approximately 1.6%. Therefore, we propose that DIS is a greatly underdiagnosed cause of deafness and should be considered as a differential diagnosis in all children with hearing loss. Likewise, current molecular genetic testing panels for hearing loss should be expanded to include deletion/duplication analysis for this region since chromosomal microarray is the only platform currently capable of detecting deletions that result in DIS in clinical laboratories.

888T

Identification of 90 first-reported abnormal chromosome karyotypes in population of Eastern China with G-band and/or C-band banding techniques. *L. Li[1], S. Zheng[1], J. Zhang[1], W. Fan[2]*. 1) 1. Genetics Counselling Department, Genetics Institure of LinYi Peoples' hospital, 27# Eastern Liberation Road, LinYi, Shandong, 276003, China; 2) 2. School of Life Science, Hebei University, 180# Eastern Wusi Road, Baoding,Hebei Province 071002, China.

[Abstract] Purpose: To characterize the chromosomal structure abnormality in genetic counseling population in Shandong Province of Eastern China, and to understand their association features with the corresponding diseases. Methods: Peripheral blood lymphocytes were cultured and analyzed by the G-band and/or C-band techniques. Abnormal karyotypes were recorded according to International System for Human Cytogenetic Nomenclature. Results: A total of 1046 abnormal karyotypes were identified in the past 10 years from 7566 patients who suffered from abnormal pregnancy, infertility, primary/secondary amenorrhea, mental retardation/congenital malformation and others, and the average positive rate is 13.82% (1046/7566). Among them, 282 abnormal karyotypes were from 5178 patients with abnormal pregnancies (5.45%); 74 from 223 patients with male infertility (33.18%); 121 from 209 patients with primary/secondary amenorrhea (57.89%); 524 from 811 patients with putative mental retardation and congenital malformation (64.61%); 45 from 1145 patients with variety of other disorders (3.93%). Within the 1046 abnormal karyotypes described above, 90 are the first reported cases in the world (8.6%; 90/1046) that has been confirmed through consulting with the International Human Chromosomal Abnormalities Directory of Registration. These 90 first-reported karyotypes came from 41 males and 49 females, which include 79 balanced translocations, 5 cases of pericentric inversion, 1 case of paracentric inversion/insertion translocation, 1 case of partial trisomy of 6q, 1 case of partial trisomy of 7p, 1 case of partial haploid of 15/16, and 1 case of false dicentric chromosome of 7 respectively. Very interestingly, we have identified a familial 46, XX testicular disorder of sex development (DSD) containing 3 DSD patients in two generations from a family of three generations. Conclusions: Chromosome abnormalities show a higher incidence in Shandong province of Eastern China due to isolated living status of the residents. It is particularly indispensable to perform chromosomal karyotype analysis for any patients with putative developmental disorders in this area, and it also offers a good resource for human genetics study, especially the cytogenetic analysis and disease-gene identification.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

889W

Genotype-phenotype correlation in individuals with MEF2C mutations and/or 5q14.3-q15 deletion syndrome. *E. Pallesi-Pocachard[1], J. Andrieux[2], A-M. Bisgaard Pedersen[3], A. Carabalona[1], R. Steensbjerre Moller[4], P-H. Kaad[5], E. Parrini[6], D. Heron[7], B. Keren[7], B. Benyahia[7], M. Walsh[8], N. Sobreira[8], D. Batista[9], V. Malan[10], A. Represa[1], T. Wang[8], R. Guerrini[6], M. Kirchhoff[3], C. Cardoso[1], M. Holder-Espinasse[11]*. 1) INMED, INSERM U901, Université de la Méditerranée, Campus de Luminy, Marseille, France; 2) Plateforme de génomique, Centre de Biologie Pathologie, CHRU, Lille, France; 3) Department of Clinical Genetics, Rigshospitalet, Denmark; 4) The Wilhelm Johannsen Centre for Functional Genome Research, Department of Molecular and Cellular Medicine, The Panum Institute, University of Copenhagen, Copenhagen, Denmark; 5) Department of Paediatrics, Hjoerring Hospital, Hjoerring, Denmark; 6) Pediatric Neurology and Neurogenetics Unit and Laboratories, Children's Hospital A. Meyer-University of Florence, Florence, Italy; 7) Génétique Moléculaire et Chromosomique, Groupe Hospitalier Pitié-Salpêtrière, Paris, France; 8) Institute of Genetic Medicine and Department of Pediatrics, Johns Hopkins University School of Medicine, Baltimore, USA; 9) Kennedy Krieger Institute and Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, USA; 10) Département de Cytogénétique et de Génétique, Hôpital Necker-Enfants Malades, Paris, France; 11) Service de Génétique Clinique, Hôpital Jeanne de Flandre, CHRU de Lille, France.

Intellectual Disability (ID) is the most common developmental disorder, affecting cognitive and adaptive functions with a frequency of approximately 2–3% in the general population. Microarray genomic hybridization has shown that submicroscopic genomic imbalances cause ID in at least 10% of idiopathic cases. Recently, microdeletions of the 5q14.3 region including the MEF2C gene have been described in patients featuring severe ID, stereotypic movements, epilepsy and cerebral malformations. Point mutations in MEF2C have been observed in a small number of patients with ID. However, the phenotypes caused by MEF2C point mutations are not different from those caused by 5q14.3 deletions. The protein encoded by the MEF2C gene has a crucial function in the homeostatic control of activity-dependent synaptogenesis that plays a role in the establishment of functional neuronal circuits during development and memory storage. In brain, MEF2C is essential for early neurogenesis, neuronal migration and differentiation. To delineate the MEF2C syndrome phenotype and to determine genotype-phenotype correlations, we report on 4 patients presenting a 5q14.3-q15 deletion including the MEF2C gene and 6 patients presenting overlapping 5q14.3-q15 deletions that did not span MEF2C. Comparison of these new patients together with the 25 previously reported MEF2C cases revealed that MEF2C haploinsuficiency leads to severe ID, seizures and hypotonia often associated with poor visual contact, failure to thrive or to walk independantly. In constrast, when we compared the latter patients with those presenting a 5q14.3-q15 deletion that did not include MEF2C, we found a less severe phenotype, although comprising hypotonia, ID and seizures. In conclusion, our data support the role of MEF2C in brain development and motility and suggest that another gene in the 5q15 region might as well be responsible for ID and epilepsy.

890T

Full-blown expression of paternally inherited Fabry disease in a young woman due to complete skewed X-inactivation of a der(X)t(X;Y)(p22.13;p11.2)mat.ishYp11.3(SRY+). *F. VIALARD[1], K. BENISTAN[2], C. BELDJORD[3], S. HEIDE[2], D. MOLINA-GOMES[1], F. JABBOUR[4], P. DE MAZANCOURT[5], D.P. GERMAIN[2, 5]*. 1) Division of Cell Biology, University of Versailles, CHG Poissy, France; 2) Division of Medical Genetics, University of Versailles, CHU Raymond Poincaré, Garches, France; 3) Laboratory of Biochemistry, CHU Cochin, University Paris V Descartes, Paris, France; 4) Laboratory of Biochemistry, CHU Raymond Poincaré, Garches, France; 5) UPCG, EA 2493, University of Versailles, Montigny-le-Bretonneux, France.

Background : Fabry disease (FD, OMIM 301500) is an X-linked lysosomal storage disorder caused by the deficiency of the enzyme alpha-galactosidase A (α-Gal A). Females with FD often develop clinical manifestations. However, the disease generally follows a milder course in heterozygotes. Patient and Methods : We report on a young female patient with full-blown expression of FD including angiokeratoma, kidney failure (measured glomerular filtration rate (mGFR) = 37 mL/min/1,73 m2), cardiac manifestations (hypertrophic cardiomyopathy) and CNS involvement (periventricular white matter hyperintensities) from the age of 22 years. The proband was the patient's father. Alphagalactosidase A enzymatic assay, GLA gene sequencing, standard caryotype, comparative genomic hybridization (CGH) and X-inactivation studies were performed. Results : Alpha-Gal A activity, measured in leukocytes, was totally absent (0.0 ng/mmol/h). Gb3 urinary excretion was markedly elevated (603.1 mmol/mmol creatinine). The familial mutation in the GLA gene (c.154 T>C, p.C52R) was detected. Sequencing of all exons and flanking intron/exon boundaries revealed no other mutations in the girl. Standard caryotype revealed a 46,X,der(X)t(X;Y)(p22.3?;p11) formula. Fluorescent in situ hybridization (FISH) showed a der(X)t(X;Y)(p22.3?;p11).ishYp11.3(SRY+). To characterize the derivative chromosome, a CGH was done which confirmed the t(X;Y) translocation and identified a large 16.6 Mb deletion spanning from Xp22.13 to Xp22.32. As FD was inherited from her father and feminine phenotype was present on physical examination, we hypothesized a maternal inheritance of the derivative chromosome together with skewed X-chromosome inactivation. Using polymorphic markers at the androgen receptor locus, a 100% skewed X-inactivation of the der(X)t(X;Y)(SRY+) was observed in leukocytes. Furthermore, the maternal inheritance of the derivative was confirmed through its identification in the patient's mother. Discussion : This is, to our knowledge, the first report of a 16.6 Mb deletion of a maternally inherited der(X)t(X;Y)(p22.13;p11.2).ishYp11.3(SRY+) leading to (i) its completely skewed inactivation with consequent feminine external and internal genitalia in the patient despite the presence of a SRY gene (ii) expression of the cytogenetically normal X chromosome inherited from the father with a GLA gene point mutation leading to absence of residual enzyme activity and full-blown phenotype of FD.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

891W

Whole-genome array-CGH screening in patients with autosomal dominant sensorineural hearing loss points to novel susceptibility loci. *E.L. Freitas[1], D.T. Uehara[1], V.G.L. Dantas[1], A.G. Silva[2], A.C. Krepischi[2], R.C. Mingroni-Netto[1], C. Rosenberg[1]*. 1) Department of Genetics and Evolutionary Biology, University of São Paulo, São Paulo, SP., Brazil; 2) Hospital A.C. Camargo, São Paulo, Brazil.

Hereditary hearing loss is a clinically and genetically heterogeneous sensory disorder. In 20% of the cases it is inherited in an autosomal dominant pattern with no additional clinical signs (ADNSHL). To date, 59 loci for ADNSHL have been mapped to chromosomal regions. Due to tremendous genetic heterogeneity, the identification of genes that affect the process of hearing has been challenging; in order to identify novel DFNA locus, we investigated 50 patients with ADNSHL by oligonucleotide array-CGH. Copy number alterations were detected in ten ADNSHL patients (20%) at 2p23.1, 4p13, 5q32, 6q12, 6q23.2, 7q31.1, 8q13.3, 10q21.1, 17q24.3-q25.1 and 19q13.42-q13.43 chromosome regions. Some of regions were previously described as ADNSHL locus (DFNA10, DFNA15, DFNB23 and BOR1). The chromosome regions shown to be in altered copy number encompassed some biologically plausible candidate genes, including CHRNA9, SLC39A11, EYS and DOCK4. The first one, a 261 Kb duplication at 4p13 region, encompasses part of the CHRNA9 gene. CHRNA9 is an ionotropic receptor with a probable role in the modulation of auditory stimuli. This gene also interacts with SLC26A5, GRXCR1, MYO7A and POU4F3 genes, which have been described as hearing loss locus. The second one was a 87 Kb duplication at 17q24.3-q25.1 chromosome region, affecting the SLC39A11 gene. This gene belongs to the SLC family, which includes approximately 300 genes that transport various molecules across the membranes surrounding the cell and its component parts. Two SLC genes were previously described as hearing loss locus, namely SLC17A8 gene (DFNA25) and SLC26A5 gene (DFNB61). The third alteration was a 56 Kb deletion on chromosome 6q12 region, affecting part of the EYS gene. This gene is a eyes shut homolog (Drosophila) and is required to maintain the integrity of photoreceptor cells. When mutated can lead to Retinitis Pigmentosa. Many genes related to hearing loss are also involved in the development of ocular disorders. The last alteration was a 438 Kb duplication located on 7q31.1 region affecting two genes: IMMP2L and DOCK4. In this case the DOCK4 is a strong candidate gene for causing the ADNSHL since previous studies indicate that a DOCK4 harmonin-activated signaling pathway regulates actin cytoskeleton organization in stereocilia. Our results provide support for a role for rare CNVs in ADNSHL risk and reinforce evidence for the existence of novel susceptibility genes for this condition.

892T

Comparison of HER2/neu gene amplification assessment by Fluorescent In Situ hybridization (FISH) and Immunohistochemistry (IHC) in Iranian breast cancer patients. *M. Moradi chaleshtori[1], Z. Hojati[1], H. Teimori[2], SH. Ramazi[1]*. 1) Genetics division, Biology department,Science faculty, Esfahan university, Isfahan, Isfahan, Iran; 2) Genetics-Biochemistry division, medicine faculty, Shahrekord medicine science university, Shahrekord,Iran.

The Human Epidermal Growth Factor Receptor 2 (HER2) gene encodes a 185 kd protein that acts as a cell surface receptor with tyrosine kinase activity . About 20–30% of breast cancer tumors are positive for amplification and/or overexpression of HER2.HER2 positive tumors are more aggressive than HER2 negative tumors and are associated with poor prognosis. This study designed to evaluate HER2/neu gene status in breast cancer specimens by Fluorescent In Situ hybridization (FISH) and Immunohistochemistry (IHC). In this study formalin fixed paraffin embedded (FFPE) breast cancer tissue specimens with known IHC results were evaluated by FISH technique. The copy number of HER2 gene locus at 17q11.2-q12 and alpha satellite DNA located at band region 17p11.1-q11.1 (CEP17) was estimated by FISH in interphase cells of FFPE tissue sections(4–6 μm). HER2 gene amplification status was determined as the ratio between the total number of HER2 gene signals and the number of chromosome 17 centromere signals. HER2 gene amplification was reported in cases with HER2:chromosome 17 signal ratio≥2.2 then the results of FISH and IHC were compared. The results of this study confirm FISH provides more accurate scoring method for determining HER2 amplification status than IHC and FISH analysis should be applied in IHC equivocal (IHC 2+) cases of breast cancer.

893W

Array CGH of Normal Karyotype Plasma Cell Neoplasms Reveals Hidden Recurrent and Individual Genomic Copy Number Alternations. *G. Tang[1], H. Zhang[2], C. Shaw[2], P. Hixson[2], W. Bi[1], C. Borgan[2], M. Coyle[2], D. Freppon[2], D. Vo[2], J. O'Hare[2], J. Gonzalez Berjon[2], C. Chang[3], S. Cheung[1], A. Patel[1]*. 1) Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA; 2) Methodist Hospital, Houston, TX, USA; 3) Florida Hospital, Orlando, FL, USA.

BACKGROUND: The plasma cell neoplasms (PCNs) are a heterogeneous group of diseases characterized by clonal expansion of terminally differentiated B cells that typically secrete monoclonal immunoglobulin. Genomic aberration remains one of the most important prognostic factors in PCNs. We aimed to identify genomic aberrations in karyotypically normal PCNs. METHODS: 23 patients with normal karyotype PCNs (myeloma = 13, MGUS = 7, and amyloidosis = 3) were included in the study. Plasma cells enrichment was performed in 17 cases by using CD138 antibody. The median purification of plasma cells was 60% (range: 32%-93%). Array CGH was performed on DNA amplified from the enriched plasma cells using GenomePlex single cell whole genome amplification kit (17 cases)or on genomic DNA extracted from bone marrow using the Puregene DNA extraction kit (7 cases). A custom genome-wide oligo microarray was designed using the Agilent 44K platform containing 493 genes involved in carcinogenesis. DNA digestion, labeling, and hybridization were performed as previously described [PMID: 20724749]. RESULTS: 16 out of 23 (70%) patients showed at least one genomic copy number aberration. Among them, 2 patients had 1 abnormality, 2 patients had 2 abnormalities, and 12 patients had ≥3 abnormalities. Among the 7 patients with normal array CGH results, 5 patients were diagnosed with MGUS, and 2 patients with amyloidosis. All myeloma patients showed genomic imbalance. Several recurrent abnormalities were detected, including Loss in 13q12.11-q34 (x7, involving RB gene); 14q32 (x7, involving IgH gene); 16p11.2 (x5); 9p24.1 (x4); 17p11.2-13.3 (x2, involving p53); 20p13-p11.22 (x2); and 19p13.3 (x2); Gains in 5p15.33 (x4) and 1q21.1 (x3). Loss of 20p13-p11.22 was accompanied with Loss of 17p11.2–13.3. CONCLUSIONS: This study showed that 70% of PCNS with normal karyotype harbored genomic imbalances, and majority of cases showed complex aberrations. Several recurrent abnormalities were identified, some with known prognostic significance. We conclude that array CGH provides a more accurate biological or prognostic profile for PCNs than conventional cytogenetics and FISH. Future functional studies of the genes in the recurrent aberrant regions may help uncover common genetic events leading to the development of PCNs.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

894T

Interstitial duplication/triplication 20q13.2 : clinical, cytogenetical and genetical characterization. *A. Briand-Suleau[1,2], J. Martinovic[3], L. Tosca[4,5], B. Tou[6,7], S. Brisset[4], J. Bouligand[6,7], V. Delattre[1,2], I. Giurgea[1,2], C. Goumy[9], C. Francannet[10], A. Guiochon-Mantel[6,7], A. Benachi[8], J. Vermeesch[11], G. Tachdjian[4,5], P. Vago[9], M. Goossens[1,2], C. Métay[1,2].* 1) Service de Biochimie-Génétique, Plateforme de Génétique Constitutionnelle, AP-HP, Hôpital H. Mondor, Créteil, France; 2) INSERM U955, Plateforme de Génétique Constitutionnelle, Hôpital H. Mondor, Créteil, France; 3) Unité fonctionnelle de Fœtopathologie, Hôpital Béclère, Clamart, France; 4) Service d'Histologie, Embryologie et Cytogénétique, Hôpital Béclère, Université Paris Sud, Clamart, France; 5) INSERM U935, Villejuif, France; 6) Service de Génétique moléculaire, Pharmacogénétique et Hormonologie, Hôpital Bicêtre, Le Kremlin-Bicêtre, France; 7) INSERM U693, IFR93, Faculté de Médecine Paris-Sud, Univ Paris-Sud, Le Kremlin-Bicêtre, France; 8) Service de Gynécologie Obstétrique, Hôpital Béclère, Clamart, France; 9) Cytogénétique Médicale, Univ Clermont1, UFR Médecine, CHU Clermont-Ferrand, CHU Estaing, F-63003, France ; ERTICa, EA 4677, Univ Clermont1, UFR Médecine, F-63000, France; 10) Génétique Médicale, CHU Estaing, Clermont-Ferrand, France; 11) Centre for Human Genetics, Leuven, Belgium.

We report *de novo* 20q13 duplications/triplication ranging from 1,1 to 11,5Mb in three cases with multiple congenital anomalies. Patient 1 (P1) was a fetus exhibiting severe congenital heart defect, dysmorphy and brachymesophalangia. Patient 2 (P2) was a boy with developmental delay, macrosomia, dysmorphic facies, squeletal anomalies, cerebral malformations and congenital heart disease. Patient 3 (P3) was a girl with developmental delay, dysmorphy and scoliosis. Partial trisomy 20q is rare, and most reported cases result from unbalanced segregation of a balanced reciprocal translocation. So far, only four cases of *de novo* pure partial trisomies 20q have been described, and among them two have been more precisely characterized by Agilent CGHarray and reported here (duplication for P2 and triplication detected for P3 respectively initially described in Blanc P et al, AJMG 2008 and Menten B et al, JMG 2006). The third one is defined by an overlapping *de novo* duplication of 20q13.12q13.31 and thus is included in discussion (Iglesias A et al, Clin Dysmor 2006 and personal communication). The last described case doesn't overlap and thus is not included here (Wanderley H.Y.C. et al, Gene Couns 2005). Among common clinical features, all children exhibit developmental delay and some dysmorphic features. Skeletal and cardiac malformations were present in 3\4 of the patients. Among genes included in the overlapping duplications, two were of particular interest for cardiac abnormalities. *SALL4* is shared by cases 1, 2 and Iglesias et al (2006). It is associated with Duane-radial ray/Okihiro syndrome, characterized by upper limb, ocular and renal anomalies and congenital heart defects. P1 and P2 presented ventricular septal defects. Altered *SALL4* expression in mice suggested its' role in heart patterning and morphogenesis. *NFATC2* was encompassed in all duplications. In mice model and zebrafish, Nfatc2 was involved in heart valve morphogenesis. P1 and P2 shared mitral and aortic valves atresia and Iglesias et al (2006)'s patient displayed aortic coarctation. Cardiac malformations could result from abnormal expression of *SALL4* combined with *NFATC2*, as no cardiac malformations were described in P3. Nfatc2, implicated also in skeletal and cartilage development could be related to the skeletal features observed in 3\4 patients. These cases further highlight genotype-phenotype correlation in 20q13.2 interstitial duplication/triplication.

895W

Array-Based Copy-Number Analysis in Patients Associating Obesity and Developmental Delay/Learning Disabilities and Additional Features. *C. D'Angelo[1], I. Kohl[1], C. de Castro[1], C. Kim[2], D. Bertola[2], C. Lourenço[3], A-B. Perez[4], C. Koiffmann[1].* 1) Dept Genetics, Univ Sao Paulo, Sao Paulo, Brazil; 2) Genetics Unit, Department of Pediatrics, Children Institute, School of Medicine, University of Sao Paulo, Sao Paulo, Brazil; 3) Neurogenetics Unit, Department of Medical Genetics, School of Medicine, University of Sao Paulo, Ribeirao Preto, Brazil; 4) Department of Morphology, Medical Genetics Center, Federal University of Sao Paulo, Sao Paulo, Brazil.

There are numerous reports of obesity present as part of a syndromic chromosome abnormality. Many have an overlapping phenotype that poses a diagnostic challenge. The most prevalent of these conditions is Prader-Willi syndrome (PWS), which is due to lack of paternal expression of imprinted genes within chromosome 15q11q13. Other examples include deletions of chromosomes 1p36, 2q37, 6q16, 9q34, and 17p11.2. The recent introduction of genomic arrays in clinical practice has led to the identification of new genomic regions exhibiting a dosage dependent influence on body weight regulation and also learning and development. In this study, 93 syndromic patients with obesity (defined as a BMI at or above the 95th percentile) as a shared clinical feature were investigated by single nucleotide polymorphism (SNP) array platforms (Affymetrix 100K and 500K) and/or oligonucleotide array-based comparative genomic hybridization (CGH) platforms (Agilent SurePrint G3 Human CGH 8 × 60K and CytoSure ISCA 4 × 180K produced for OGT (Oxford Gene Technology, Oxford, UK) Identified deletions/duplications of at least a 300 kb region were confirmed using fluorescence in situ hybridization (FISH) or multiplex ligation-dependent probe amplification (MLPA). Parental testing was performed when parental DNA was available. We identified 21 individuals with pathogenic and potentially pathogenic CNVs as well as four patients with CNVs that are initially of unknown significance. Of the CNVs identified in this cohort, some have been previously reported in cases with obesity (e.g. deletions of chromosomes 2q37, 6q16 and 9q34). Haploinsufficiency of HDAC4, SIM1 and EHMT1 are likely to be responsible for the obesity phenotype in these cases. In addition, obesity and intellectual disability were recently reported in patients with 2p25.3 deletions overlapping minimally within MYT1L. Interestingly, the obesity-associated TMEM18 gene is located in band 2p25.3. In a number of regions, there are quite a few genes to be examined for a possible relationship to the patient's phenotype. For instance, we found a de novo deletion of 14q12 which contained only one obesity-associated gene, PRKD1. In conclusion, microarray analysis in clinical practice has facilitated the identification of new obesity-associated syndromes, while genotype-phenotype correlations are critical to determining the effects of such novel CNVs. Financial Support: CEPID-FAPESP, CNPq.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

896T

Complex genomic reorganization in microdeletions/microduplications syndromes and molecular mechanisms for formation. *L. Kulikowski[1,2], F. Piazzon[1], E. Zanardo[1,2], R. Dutra[1,2], A. Dias[1], M. Oliveira[3], M. Moreira[1,2], G. Novo-Filho[1,2], M. Basso[1,2], C. Negretto[2], A. Zandoná[1,2], M. Melaragno[3], C. Kim[2]*. 1) Department of Pathology, Cytogenomics Laboratory, LIM 03, Universidade de São Paulo, SP, Brazil; 2) Genetics Unit of the Children's Institute, University of São Paulo, São Paulo, Brazil; 3) Department of Morphology and Genetics, Universidade Federal de São Paulo, São Paulo, Brazil.

Different mutational mechanisms shape the dynamic architecture of human genomes and occasionally result in congenital defects and disease. Detailed delineation of genomic rearrangements requires the application of advance cytogenomic techniques, such as MLPA and array-based copy number analysis. Some mechanisms have been proposed to explain the complexity of genomic rearrangements in microdeletions/microduplications syndromes:Non-homologous end joining (NHEJ) associated to microhomology and inserted sequences; replication-based model called fork stalling and template switching (FoSTeS) and a novel model called microhomology-mediated break-induced replication (MMBIR). We report on two individuals with features suggestive of microdel/microdup syndromes but lacking associated distinctive genomic anomalies: a 2.8-year-old girl with features of monossomy 1p36 syndrome and other congenital defects and a boy with atypical deletion at Williams-Beuren syndrome region and a chromosome Y supernumerary. First patient: MLPA (MRC-Holland) assay using the P070 and P036 kits revealed the deletions of the *TNFRSF4* and *TNFRSF18* genes whereas the P064 kit showed a normal copy of the *TP73* gene. Analysis by SNP-array (Affymetrix Genome-Wide Human-6.0) showed multiple copy number changes. We observed a normal copy (nml) sequence followed by a deletion (564,620–2,456,203), followed by a normal copy (nml) sequence, followed by another deletion (2,473,257-3,446,813), followed by a normal copy (nml) sequence and then by a duplication (3,474,630-3,641,681) at 1p36 region. Second patient: MLPA with P064 revealed the deletion of *CYLN2, STX1A, ELN, LIMK1* genes and the presence of *FZD9* gene. Array showed rearrangements with complex patterns including 7p14.3, 7q11.23, 13q31.3, Xp22.33, Xq28, Yp11.2 and Yq11.23 regions. Recently, the phenomenon of chromothripsis was described in cancer cells by Stephens et al (2011). Liu at al (2011) propose that both the constitutional genomic rearrangements and the chromothripsis processes reflect basic DNA metabolism and share a cellular DNA replication/repair mechanism and Kloosterman et al (2011) provided evidence that chromothripsis may be generating structural variation in the germline that results in congenital defects. The delineation of the complex genome architecture and the definition of the mechanisms for formation of the rearrangements offer an opportunity for the discovery of novel syndromes and the development of translational research.

897W

Gene Mapping in the Finnish National Collection of Balanced Translocations and Inversions. *T.M. Luukkonen[2,3], M. Pöyhönen[1,4], P. Ellonen[2], S. Lagström[2], J.H. Lee[5], K.O.J. Simola[6], K. Aittomäki[4], J. Ignatius[7], R. Salonen[8], A. Palotie[1,2,9], J. Terwilliger[3,10,11], T. Varilo[1,3]*. 1) Dept. of Medical Genetics, Haartman Institute, University of Helsinki, Helsinki, Finland; 2) Institute for Molecular Medicine Finland FIMM, Helsinki, Finland; 3) National Institute for Health and Welfare, Helsinki, Finland; 4) Dept. of Clinical Genetics, Helsinki University Central Hospital, Helsinki, Finland; 5) Sergievsky Center/Taub Institute, Columbia University, New York, NY, USA; 6) Dept. of Pediatrics, Tampere University Hospital, Tampere, Finland; 7) Dept. of Clinical Genetics, Turku University Central Hospital, Turku, Finland; 8) Dept. of Medical Genetics, Väestöliitto, Helsinki, Finland; 9) Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK; 10) Dept. of Genetics and Development, Dept. of Psychiatry, Columbia Genome Center, Columbia University, New York, NY, USA; 11) Division of Medical Genetics, New York State Psychiatric Institute, New York, NY, USA.

We have assembled a nationwide collection of all known carriers of reciprocal balanced translocations and inversions from every Medical Genetics Department and Clinical Genetics Laboratory in Finland. Our database contains relevant medical records of all available carriers in each family to include all relatives carrying the same rearrangement identical by descent. The data facilitate diagnostic purposes, genetic counseling, and subsequent follow-up. We also use this collection as a valuable research tool, focusing on identification of the precise molecular location of the chromosomal breakpoints, and correlating these to phenotypes as a natural experiment for gene mapping. To date we have systematically examined 3016 carriers and their relatives, and are in the process of collecting biological samples from them, and linking our data to national comprehensive disease registers. As a pilot study, we have focused on families identified from this resource, which appear to have a striking correlation of a balanced translocation and some phenotype. The initial investigation includes three families with (a) intracranial and aortic aneurysm (n=5), (b) short stature (n=4), and (c) tetraphocomelia (de novo) likely caused by the translocation. In the first family, the proband has a thoracic aortic aneurysm and a translocation involving a region which was previously linked to aneurysms; the proband's brother and sister both died from cerebral hemorrhage; and the proband's father and aunt each died from aortic rupture. In the second family, a 10-year-old girl presented with a growth disturbance and a translocation involving a region previously implicated affecting the height of females; the proband's mother was 164 cm, the maternal grandmother 152 cm, and the maternal great-grandmother 147 cm. In individuals from each family, we have been performing standard cytogenetic analysis, copy number analysis (genotyping and aCGH), genome-wide paired-end sequencing, and capillary sequencing with the objective of identifying the specific breakpoints for each translocation and ruling out other potentially causative genetic factors. We are in the final stages of analyzing the data from the aortic aneurysm family, and data from the other projects are forthcoming.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

898T

Combined exon array CGH plus SNP genome analyses for optimized clinical diagnostics. *A. Patel[1], W. Bi[1], C. Shaw[1], S-H. Kang[3], A.N. Pursley[1], S. Lalani[1,2], P. Hixson[1], T. Gambin[4], A.C-H. Tsai[5], H-G. Bock[6], M. Descartes[7], F. Probst[1,2], F. Sacaglia[1,2], A.L. Beaudet[1,2], J.R. Lupski[1,2], C. Eng[1,2], SW. Cheung[1], C. Bacino[1,2], J. Wiszniewska[1]*. 1) Dept Molecular & Human Gen, Baylor Col Medicine, Houston, TX; 2) Pediatrics Baylor College of Medicine, Houston, TX; 3) Allina Cytogenetics Laboratory, Abbott Northwestern Hospital, Minneapolis, MN; 4) Institute of Computer Science, Warsaw University of Technology, Warsaw, Poland; 5) Department of Molecular and Medical Genetics, OHSU, Portland,OR; 6) Department of Pediatrics, University of Mississippi Medical Center, Jackson, MS; 7) Department of Genetics, University of Alabama, Birmingham, AL.

In clinical diagnostics, both array comparative genomic hybridization (array CGH) and single nucleotide polymorphism (SNP) genotyping have proven to be powerful technologies utilized for the evaluation of developmental delay, congenital anomalies, and neuropsychiatric disorders. Many clinical laboratories offer testing using commercially available standard designs or custom-designed arrays based on either of these technologies. Differences in these arrays may constitute a challenge for clinicians when it comes to choosing the right type of array to investigate the underlying genetic cause for the disease in their patient. While both approaches are effective in detecting copy number variants (CNVs) greater than several hundred kilobases in size, only SNP arrays show higher sensitivity for detection of low level mosaic aneuploidies and chimerism and offer the ability to detect copy number neutral regions of absence of heterozygosity (AOH) which are relevant clinically to identify uniparental disomy(UPD), consanguinity and potential recessive loci. However, SNP arrays are unable to detect single exon CNVs due to the distribution of SNPs across the genome. In an effort to provide a comprehensive clinical testing service for both CNVs and copy neutral AOH, we have enhanced our custom designed high resolution oligonucleotide array (manufactured by Agilent) with exon targeted coverage of 1714 genes and 60k SNP probes referred to as CMA Comprehensive. Of the 3240 cases evaluated by this array, CNVs were detected in 452 cases (14%) including 28 clinically significant exonic CNVs. 166 cases (5.1%) showed AOH giving a combined diagnostic yield of 19.1%. This demonstrates that even though this array has a lower density of SNP probes than other commercially available SNP arrays, it was able to reliably detect AOH events >10 Mb as well as intragenic copy number changes beyond the detection limitations of SNP genotyping. Thus, combining SNP probes and exon targeted array CGH into one platform increases diagnostic yield and provides clinically useful genetic screening for identifying small exonic copy number changes, UPD and consanguinity in a cost effective manner.

899W

Facing current diagnostic challenges in array copy number analysis. *F.B. Piazzon[1,2], R. Dutra[1,2], E. Zanardo[1,2], A. Dias[1,2], M. Moreira[1,2], G. Novo-Filho[1,2], D. Christofolini[3], V. Meloni[3], M. Oliveira[3], A. Zandoná-Teixeira[2], M. Basso[1,2], C. Negretto[1,2], C. Moreira-Filho[2], M. Melaragno[3], C. Kim[2], L. Kulikowski[1,2]*. 1) Department of Pathology, Cytogenomics Laboratory, LIM 03, University of São Paulo, São Paulo, Brazil; 2) Unit of Genetics of Children's Institute, University of São Paulo, São Paulo, Brazil; 3) Department of Morphology and Genetics, Federal University of São Paulo, São Paulo, Brazil.

The recent technological advances in array-based genomics have established the opportunity to detect several different genomic structural rearrangements in patients with congenital malformations. However, there is still no consensus for interpretation of the results and the DNA alterations responsible for phenotypic variation. We report on five patients with atypical results after array analysis: del(7q)/del(13q31.3)/dup(Xpter) and one extra Y chromosome; complex del/dup(1p36); del(Xq22.1)/del(17q23.3); del(2q33.1/del(3q) LOH in Xq21.1; dup(17q)/del(17p); add(7) and mar(15q/16p). The definition of the results is not clear in these patients part due to little understanding about the natural history and clinical variability associated with chromosomal aberrations detected by array. The presence of a deletion or duplication alone does not necessarily mean that the copy number alteration causes the observed phenotype and we cannot assure to consider a copy number imbalance as to be pathogenic. Also after confirming the diagnosis of the patients the array was able to show even more unexpected genomic aberrations whose correct interpretation became a new challenge. Ascribing clarification of recurrence risk and prognosis for atypical rearrangements remain a challenge for diagnosis and improvements in clinical ascertainment.

900T

Genetic analyses of Korean patients with unexplained mental retardation and developmental delay using the multiple ligation-dependent probe amplification (MLPA). *S. Shim[1], S. Sung[1], S. Park[1], K. Kang[1], M. Lee[1], S. Lyu[2], D. Cha[1,2], M. Kim[3]*. 1) Genetic Lab, CHA Gangnam Medical Center, Seoul, South Korea; 2) Department of Obstetrics and Gynecology, CHA Gangnam Medical Center, CHA University, Seoul 135-081, Korea; 3) Department of Rehabilitation Medicine, Bundang CHA Medical Center, CHA University, Sungnam 463–712, Korea.

Subtelomeric regions of each chromosome contain many genes so called gene-rich region and deletions and duplications of the regions result in deleterious consequences such as multiple congenital anomalies, moderate to severe mental retardation, dysmorphism, and/or abnormal growth or behaviors. Many studies have shown that unbalanced rearrangements of these regions are a cause of unexplained mental retardation and, or developmental delay (MR/DD) and approximately 5% of patients with MR/DD had the abnormalities. MLPA can detect various chromosomal imbalances, such as a whole chromosome aneuploidies, partial deletions and duplications, depending on the probe set used. As a primary screening method, it has many benefits such as high-throughput, cheap and easy to interpret. Total 249 of Korean patients with unexplained MR/DD were examined. The samples were initially analyzed using both MLPA and routine chromosome analysis. For MLPA analyses, two probe mixes for subtelomeric regions, SALSA P070 and P036B probe mix, were used. Each MLPA mix contains probes for subtelomeric regions of all 46 human chromosomes. All procedures including multiplex PCR and capillary electrophoresis were performed by manufacturer's protocol. The data were analyzed by the GeneMarker® software. MLPA results were obtained all 249 cases. The subtelomeric aberrations were identified in 13 cases (5.2%) using both P070 and P036 probe Mix. Five patients (2.0%) had partial deletion, two (0.8%) had partial duplications and six (3.2%) were complex rearrangements. The deletions were detected five different subtelomeric regions (1p, 3p, 4p, 11q, 22q) and duplications were detected in two subtelomeric regions (11p, 15p). Seven cases with partial deletion or duplication were de novo, not identified in parents. The abnormalities found were confirmed by FISH and/or array-based CGH and parental analyses. On the other hand, 18 cases (7.2%) had a partial deletion or duplication which was inherited from one of parents. They were considered as a normal variation, so called copy number variations (CNVs). In conclusion, a combined analysis, MLPA followed by FISH and array-based CGH is very useful method for screening of submicroscopic chromosomal imbalances. Especially, the MLPA, as a primary screening method, is less labour-intensive and more cost-effective than FISH and array-based CGH, though some probes might be located in the regions of CNVs in Korean population.

901W

Complex nature of apparently balanced chromosomal rearrangements in patients with autism spectrum disorder. *A.C. Tabet[1, 2], A. Verloes[1, 3], M. Pilorge[2], B. Benzacken[1, 3, 4], C. Betancur[2]*. 1) Cytogenetics Unit, Department of Genetics, Robert Debre Hospital, AP-HP, Paris, France; 2) Inserm U952, CNRS UMR 7224, Pierre & Marie Curie University, Paris, France; 3) Inserm U676, Paris, France; 4) Cytogenetics Unit, Jean-Verdier Hospital, AP-HP, Bondy, France.

Apparently balanced chromosomal rearrangements can be associated with an abnormal phenotype, including intellectual disability and autism spectrum disorders (ASD). Recent studies using microarrays have revealed cryptic genomic imbalances related or not to the initial chromosomal rearrangement in 25%-61% of patients with an abnormal phenotype carrying an apparently balanced chromosomal rearrangement. Here, we report a SNP-array study of 18 patients with ASD carrying an apparently balanced chromosomal rearrangement. We identified clinically significant de novo copy number variations (CNV) in 4/18 patients (22%); three were related to the initial breakpoints of the rearrangement and one occurred elsewhere in the genome. These CNVs ranged from 3.5 to 4.3 Mb. In two cases, the CNV encompassed regions that have already been involved in ASD, 2q37 and 7q21q22, while in other cases, the CNVs affected novel regions, pointing to potential candidate genes and loci. Of note, we report the first patient with autism carrying a duplication of the Wolf-Hirschhorn syndrome critical region, contributing to delineate the phenotype of this rare genomic disorder. These findings underscore the utility of SNP arrays for investigating apparently balanced chromosomal abnormalities in subjects with ASD or related neurodevelopmental disorders.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

902T

The identification of submicroscopic gains and losses at the breakpoints of apparently balanced rearrangements. *B. Torchia, J.A. Rosenfeld, L.G. Shaffer.* Signature Genomic Laboratories, PerkinElmer Inc, Spokane, WA.

The finding of an apparently balanced translocation or inversion in a child with developmental disabilities or in a prenatal specimen with or without abnormal ultrasound findings raises questions, including whether the rearrangement is truly balanced. During the course of routine microarray investigation, we have tested 106 postnatal cases and 174 prenatal cases with known, apparently balanced translocations referred for further molecular characterization of the rearrangements. Among these, 116 were de novo, 52 were inherited from a carrier parent, and in 112 cases, the inheritance was unknown because parents were not tested. Of the postnatal cases, 7.5% showed an imbalance at one of the translocation breakpoints. In addition, 9.4% of cases showed a clinically significant copy number alteration (CNA) at another genomic location, unrelated to the translocation. Among the prenatal cases, 6.9% showed an imbalance at one of the translocation breakpoints and 2.9% showed a clinically significant finding elsewhere in the genome. Among the 84 balanced inversion cases, 12.5% of the postnatal and 4.5% of the prenatal cases showed an imbalance at one of the inversion breakpoints while 5% of the postnatal and 2.3% of the prenatal cases showed a clinically relevant CNA elsewhere in the genome. In total, including 15 cases of complex, apparently balanced rearrangements, 8.4% showed a CNA at one of the breakpoints of the rearrangement while 5.3% showed a clinically relevant CNA elsewhere in the genome. Overall, these results on a total of 379 cases demonstrate the usefulness of genomic microarrays in uncovering genomic alterations at apparently balanced chromosomal rearrangements. Moreover, a substantial number of cases had CNA elsewhere in the genome that may explain the developmental disabilities or prenatal ultrasound findings in these cases.

903W

Complex phenotype associated with 17q21.31 microdeletion. *A. Pic-Taylor[1], H. Dornelles-Wawruk[2], C. Rosenberg[3], A.C. Krepischi[3], H.P.N. Safatle[2], I. Ferrari[2], J.F. Mazzeu[2, 4].* 1) Department of Genetics and Morphology, Institute of Biologic Sciences, University of Brasilia (UnB), Distrito Federal, Brazil; 2) Laboratory of Clinical Genetics, Faculty of Medicine and Health (FS), University of Brasilia (UnB), Distrito Federal, Brazil; 3) Human Genome Study Center, Department of Genetics and Evolutionary Biology, Institute of Bioscience, University of São Paulo (USP), São Paulo, Brazil; 4) Post-graduation in Genomic Sciences and Biotechnology, Catholic University of Brasilia (UCB), Distrito Federal, Brazil.

The 17q21.31 microdeletion syndrome was recently described and present various known symptoms such as developmental delay, neonatal hypotonia, low birth-weight and a pear-shaped nose. Microdeletion size can vary, however, a minimal chromosomal region has been defined to a segment of 424 kb genomic DNA (crh17: 41046729–41470954, hg 17) and includes the MAPT, STH, CRHR1, KIAA1267, and IMP5 genes. Recently, this critical region was narrowed to a 160.8 kb segment (crh17: 41310123–41470954, hg 18) that comprises only the MAPT and STH genes. The STH gene is located within one of the introns of the MAPT gene and its transcription follows the same sense of the MAPT gene transcription. Here we describe a patient with 17q21.31 microdeletion detected by array-CGH, presenting a complex phenotype. The patient is an 18-year old girl with friendly behavior, nasal speech and moderate mental retardation and seizures. Clinical examination revealed abundant thick straight hair, a triangular face with a high and flat forehead, ptosis, up-slanting palpebral fissures, epicanthic folds, a bulbous nasal tip, anteverted nostrils, a long and deep nasal filter, preauricular fistula, a high palate, a discreet microretrognatia, kyphosis, an umbilical hernia and a clubfoot. Previous clinical exams revealed: early cataract, inter-atrial communication (both surgically corrected), chronic anemia, conductive hearing loss in the left ear, hydronephrosis, ureterocele and agenesis corpus callosum. Chromosome analysis of peripheral blood lymphocytes by G-banding at the 550 or higher banding level showed a normal 46,XX karyotype. However, the results of the array-CGH exam identified an interstitial deletion in the long arm of chromosome 17, affecting the band q21.31 between the chr17:41012276–41707073 (hg 18) genomic positions. This submicroscopic chromosomal aberration consists of approximately 695 kb. The deleted region contains the MAPT, STH, CRHR1, KIAA1267, IMP5 genes and a hypothetical ORF (LOC644246). This deletion is comparatively larger than the critical region suggested. Comparison of our patient with other cases reported in the literature, who had a deletion size of similar length, showed that the presence or absence of a symptom is not dependent on the size of the deletion. Our patient presented most of the symptoms previously described as well as chronic anemia,a condition not yet described in this syndrome.

904T

Association of a Deletion Downstream of *NKX2-1* With Benign Hereditary Chorea: Transfer of a Benign CNV to Likely Pathogenic Category. *U. Surti, J. Liao, K. Coffman, J. Hu, M. Sathanoori, S. Yatsenko, S. Madan-Khetarpal, M. McGuire, A. Rajkovic.* University of Pittsburgh Medical Center, Pittsburgh, PA.

First described in 1966, benign hereditary chorea (BHC, MIM 118700) is a rare autosomal dominant movement disorder with onset in infancy or childhood. It is characterized by chorea with little or no progression and normal cognitive function. The phenotype is highly variable, and sometimes includes hypothyroidism and neonatal respiratory distress which is also named as brain-lung-thyroid syndrome. Since 2002, more than 20 heterozygous mutations and deletions of *NKX2-1* (also known as *TTF1* or *TITF1*), a gene located on chromosome 14q13, have been identified in a number of BHC families. NKX2-1 is a member of the homeodomain-containing NKX2 family of transcription factors. It is expressed in the developing thyroid, respiratory epithelium, and specific areas of the forebrain in which it is involved in interneuron specification and migration. Although recent evidence suggests that haploinsufficiency of this gene is responsible for the pathogenesis of BHC, there are several clinically diagnosed BHC families that did not have detectable *NKX2-1* mutations, indicating other genes or regulatory elements on chromosome 14 or elsewhere may also play a role. Here we report three BHC families where no *NKX2-1* gene mutation was identified. By using high-density array CGH analysis, we detected a 117 kb deletion on chromosome 14q13.3 region, which is located 175 kb downstream of *NKX2-1* in all affected members but not in unaffected ones of all tested families. Interspecies genomic comparative analysis of the deleted area revealed multiple evolutionary highly conserved non-coding sequences, indicating the existence of potential cis-acting regulatory elements in the region. We hypothesize that this deletion downstream of *NKX2-1* disrupts long-range regulatory elements necessary for its expression, causing haploinsufficiency of this gene and BHC phenotype in these families. Our results provide the first evidence supporting the role of non-coding regulatory elements in the pathogenesis of BHC, and also indicate that this deletion is likely to be a clinically significant pathogenic CNV instead of a benign one.

905W

Chromosomal Microarray Analysis (CMA) in 22q11.2 deletion syndrome-New insight to LCR-mediated rearrangement. *A. Tsai[1,2], D. Klepacka[2], M. Dodge[2].* 1) Molecular and Medical Genetics, OHSU, Portland, OR; 2) Childrens Hospital Colorado, Aurora, CO 80045.

Introduction: Small, crucial/complex chromosome imbalances have been reported in increasing numbers since the implementation of array comparative genomic hybridization (aCGH). Traditionally, individuals with the 22q11.2 deletion syndrome have been diagnosed by FISH along with high resolution chromosomes. However, with this method, concurrent chromosome imbalances could go undetected which may contribute to the complexity of the 22q11.2 deletion syndrome. In order to test the hypotheses that those who are clinically more affected can have a larger deletion or comorbid with other chromosomal anomalies, we performed microarray assay in patients who did not a microarray upon diagnosis. Methods: From 09/2010-03/2012, 52 patients with 22q11.2 del were seen at CHC 22q clinic. Our cohort consists of families from CO and WY. 48 patients were diagnosed initially by FISH, and 4 were diagnosed by microarray. 30 patients had microarray assay performed after the visit. The platform of microarray was oligoarray (105K-180K) with the exception of one patient who had a BAC array upon evaluation. There are 23 males and 11 females with ages ranging from 3 months to 18 years (median age of 6). Ethnicity includes Caucasians (54%), Hispanic (32%), African America (4%), Asian (2%) and 2% other. Results: The CMA results showed a variable size of deletion: 2 patients had deletion less than 0.6 MB; 3 patients had a deletion size about 1.7 Mb; 20 patients had a deletion ranging from 2.5 to 3Mb and 3 had a greater than 3Mb deletion. Of note, 8 patients had additional findings in addition to the 22q11 deletion: Translocation t(6;21)(q24.2;q22.1), gain of 3 BACs in 2p25.3, a 187 kb loss at 15q26.3, a 509kb gain at 15q13.3, a 400 kb gain at 17p13.1, 371kb gain in 10q21.1, and an 843 kb gain at 18q23.3. More profound delay,anxiety, autism and malformations were noted in those with greater than 3.0 MB deletion and additional chromosomal anomalies. Conclusion: 25% of the patients who had CMA were found to have additional chromosomal anomalies and a few presented with more severe phenotypes. This study also revealed the size of deletion seemed to vary in size instead of the predicted fragment based on common LCR. The co-existing CNVs identified in a significant portion of patients may also suggest a predisposition for rearrangement or lack of self-correction/repair system. aCGH should be considered for first line testing for children with 22q11.2 deletion syndrome.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

906T

Array CGH as a first line diagnostic test in place of karyotyping for postnatal referrals - results from four years clinical application for over 7,800 patients. *J.W. Ahn[1], S. Bint[2], A. Bergbaum[2], K. Mann[2], R.P. Hall[2], C. Mackie Ogilvie[1].* 1) Cytogenetics, Guy's Hospital, London, United Kingdom; 2) GSTS Pathology, Guy's Hospital, London, United Kingdom.

Array CGH is widely used in cytogenetics centres for postnatal constitutional genome analysis, and is now recommended as a first line test in place of G-banded chromosome analysis. At our centre, first line testing by oligonucleotide array CGH for all constitutional referrals for genome imbalance has been in place since June 2008. Since then, 7,890 patients have been tested, with referral indications ranging from neonatal congenital anomalies through to adult neurodisabilites; 25% of these patients have CNVs either in known pathogenic regions or in other regions where imbalances have not been reported in the normal population. Deletions and nullisomy account for 48% of these CNVs, duplications and triplications account for 51%, and mosaic imbalances make up the remainder. The vast majority of these CNVs, 88%, are <3Mb and would not be detected by traditional G-banding analysis. All our diagnostic findings can be viewed at http://bbgre.org. Clinical interpretation of array CGH findings is complex, and simple rules cannot be used; for instance, of our de novo findings (21% of completed inheritance studies), some are for small, gene-poor regions, and may not therefore be significant, whilst of inherited imbalances, at least 20% are likely to be clinically significant. By employing a patient vs patient (phenotype mismatched) hybridization strategy, we have been able to reduce costs to that of a traditional karyotype, and provide array CGH as a first line test within the financial constraints of a state-funded health system. First line testing by array CGH provides a higher diagnostic detection rate than traditional methodology, and adds to the sum of information and understanding of the role of genomic imbalance in disease. Experience of the challenges in interpretation and reporting of array CGH results will inform the implementation into clinical diagnostic service of higher resolution technologies such as whole exome and whole genome sequencing.

907W

Identification of the critical region for 7q32q34 deletion syndrome in a 1.5 Mb region at 7q32.3q33. *T.-J. Chen, J. Hoffman, K. Phelan, H.C. Andersson.* Hayward Genetics Center, Tulane University School of Medicine, New Orleans, LA.

Deletion of 7q32q34 is a rare genetic condition with clinical features including developmental delay, early feeding difficulties and failure to thrive, short stature, behavior problems, language delay, and learning difficulties. However, there is variable expression among affected patients and many features resolve with age. Most of the described cases carry deletions in the range of several million bases located in the 7q32q34 region. Herein we report an affected female and describe a critical region contributing to this phenotype. The two months old patient was referred to genetics clinic due to the history of poor feeding and hypotonia. The patient was delivered full term by C-section with reported history of decreased fetal movement. Apgar scores were 8 and 9 at 1 and 5 minutes, respectively. The patient spent the first 23 days of life in the NICU due to feeding problems and hypotonia, and she underwent placement of a G-tube at one week of age. Her birth weight was 3.15 kg (11–25%), and the length was 49 cm (11–25%). At the age of two months, she was fed by mouth and G-tube and hypotonia was still obvious. Her weight was 3.7 kg (<3%) and her length was 52 cm (<3%). No dysmorphic facial features were noted. Laboratory studies included karyotype, methylation and FISH assays for PWS/AS, mutation analysis for DM and SMA, and biochemical profiles for amino acids and organic acids, which were all normal. The patient had a normal brain MRI and cardiovascular evaluation by ECG, chest x-ray, and echocardiography. Chromosomal array CGH was performed and revealed a 1.5 Mb deletion at 7q32.3q33. There are only three genes, CHCHD3, EXCO4, and LRGUK, located within this deleted region. EXCO4 was showed to be haploinsufficient in a previous report, suggesting deletions including this gene might play a pathogenic role contributing to this clinical phenotype. Considering this patient presented most of the typical clinical features of 7q32q34 deletion syndrome as well as possessed a characterized deletion at 7q32.3q33, we propose that the 1.5 Mb deleted region is the critical region for this genetic condition. Parental FISH analysis showed that the patient's father, who was reportedly phenotypically normal, carries a deletion at the same location. Thus, in addition to variable expressivity, incomplete penetrance might be a genetic feature for this deletion syndrome.

908T

Clinical implementation of whole-genome array CGH in Korean patients with pediatric neurologic diseases and multiple congenital anomalies. *E. Shin[1], K. Lee[1], B. Eun[2].* 1) Genome Res Ctr, NeoDin Med Inst, Seoul, Korea; 2) Department of Pediatrics, Korea University Guro Hospital, Seoul, Korea.

Chromosomal aberrations are a common cause of multiple anomaly syndromes that include developmental delay, mental retardation, dysmorphic features and multiple congenital anomalies. Microarray based comparative genomic hybridization (array-CGH) is considered to be superior for the investigation of chromosomal aberrations in children with multiple anomaly syndromes and has been demonstrated to improve the diagnostic detection rate of these small chromosomal abnormalities. We performed a high resolution genome-wide aCGH analysis for submicroscopic chromosomal rearrangements using the Roche NimbleGen 135K oligonucleotide array on 249 children with pediatric disease and congenital anomalies. Using array-CGH we detected chromosomal imbalances in 55 of 249 patients (22.1%) ranging in size from 0.3 to 61 Mb and 33 of 55 patients with pathogenic copy number variants (60%) had normal karyotype and the others (40%) had abnormality in conventional chromosome analysis. We found 20 known syndromes such as 1p36 deletion, Mowat-Wilson, Williams-Beuren, Kleefstra and Jacobsen syndrome and CNVs containing pathogenic genes and recurrent CNVs containing the same region among different patients were detected in 35 of 55 patients (63.6%). Two cases with low level mosaicism were detected by array CGH and one of them had mosaicism for rearranged chromosome 18 in 7.9% of cell and another case showed an isochromosome 12p (associated with Pallister-Killian syndrome) using genomic DNA extracted from peripheral blood. Recent advances in the analysis of patients with congenital abnormalities using whole-genome array CGH has uncovered large scale copy number variations were frequently observed and become an important tool for the detection of chromosome aberrations and has the potential to identify genes involved in pediatric neurologic disease and multiple congenital anomalies. The detection of genomic imbalances of clinical significance using whole-genome array CGH will increase knowledge of the human genome by performing genotype phenotype correlation.

909W

Chromosome and segmental imbalances detected by array CGH in samples of product of conception: a pilot study. *S. Xu, X. Wang, W. Xu, F. Gong, J. Lee, X. Lu, S. Li.* Department of Pediatrics, the University of Oklahoma Health Sciences Center, Oklahoma City, OK 73104.

Approximately 15 % of pregnancy complaints are pregnancy loss, including spontaneous abortion (SB), and still birth. Routine cytogenetics studies have demonstrated that only half of SB and less than one third of still birth have been detected with chromosome abnormalities, left the remaining up to 50 to 70% of cases with unknown causes. Microarray analysis, capable of detecting cryptic chromosome changes less than 100 kb, will unravel many submicroscopic chromosome abnormalities missed by conventional karyotype. In this study, we carried out microarray analysis of 53 samples of products of conception (POC), which were cytogenetically normal or failed to grow. DNAs were extracted from POC specimens and then hybridized on oligo microarrays (NimbleGen 720K v.3.1). All data were analyzed using SignalMap software. Of 16 (30.2 %) positive POC cases, 8 showed deletions or duplications, with all sizes between 47.5 kb to 1.94 Mb except for one with 32 Mb. The remaining 8 positive cases presented a gain of chromosomes Y, 9, 15 and 21 and a loss of chromosome X. In these aneuploidy-like cases, 6 of them may cytogenetically manifest as 45,X (2 cases), 47,XYY and trisomy 21(3 cases). The majority of routine karyotype data for the 16 positive chromosomal imbalance cases were not available due to culture failure resulted from macerated tissues received. Our preliminary data in this study provides evidence that array CGH should be considered when the routine cytogenetic analysis results were normal or failed to yield the results. Accumulated study in the future may help to identify genes or chromosomal regions leading fetal death. Microarray data from our lab and combination of such data from other labs, as well as further supplementary experiments and in depth parental/family surveys (whenever possible) are underway.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

910T

Hemophilia A and trigonocephaly in a male with a complex chromosome rearrangement involving chromosomes X and 1. *J. Liu, J. Chernos, R. Lamont, R. Perrier, B. Argiropoulos.* Dept of Medical Genetics, University of Calgary, 3330 Hospital Dr. NW, Calgary, Alberta, Canada.

We report on a 2½ year old male with hemophilia A and trigonocephaly. Array CGH and FISH analyses demonstrated a complex chromosome rearrangement involving two gains on the X chromosome within band q28 and one gain on chromosome 1 within band p13.1. Interestingly, the chromosome 1 gain was inserted into the X chromosome at band q28. The proximal 53 kb X chromosome gain overlaps exons 14 to 25 of the Factor 8 (F8) gene while the distal 219 kb deletion overlaps 5 RefSeq genes. These gains are separated by a 69 kb region of normal copy number. The chromosome 1 gain is 450 kb and overlaps 5 RefSeq genes, none of which are associated with disease. In light of the patient's clinical diagnosis of hemophilia A, this complex rearrangement is most likely perturbing F8 gene structure and function. Several published reports have demonstrated a strong association with X-linked trigonocephaly to chromosome band Xq28. Therefore, it is possible that at least one of the three duplications has inserted and disrupted a trigonocephaly locus at Xq28. DNA sequencing studies are underway to elucidate the genomic architecture of this complex rearrangement with the aim of identifying the trigonocephaly locus at Xq28.

911W

Identification of cryptic genomic imbalance in congenital and development abnormalities: An institutional experience. *F. J. Sheth[1], J. Andrieux[2], M. Desai[1], J. Sheth[1], T. Mampilly[3], G. Mampilly[3], C. Neeradha[3], J. Vijayalakshmy[3], K. Godbole[4], S. Desai[5].* 1) Biochemical and Molecular Genetics, Institute of Human Genetics, Ahmedabad, Gujarat, India; 2) Laboratory of Medical Genetics, Jeanne de Flandre Hospital CHRU de Lille, Lille Cedex, France; 3) National Institute for Empowerment of Persons with Multiple Disabilities (NIEPMD) East Coast Road, Muttukadu, Kovalam Post Chennai - 603112, India; 4) Deenanath Mangeshkar Hospital, Near Mhatre Bridge, Erandawne, Pune-411004. India; 5) Shree Krishna Hospital, Karamsad, Anand. India.

Submicroscopic genomic alterations are a major cause of congenital and developmental abnormalities including dysmorphic features, mental retardation, developmental delay, and multiple congenital anomalies. Present study was carried out in 21 cases that were initially investigated for conventional cytogenetic G-Banding analysis; amongst which 8 had congenital anomalies with/without mental retardation and 13 had with developmental delay, learning disability and seizures. No structural or numerical anomalies were detected at 450-band resolution. Hence, all 21 cases were further analyzed using 44k oligonucleotide array-CGH study using Agilent platform. Cryptic genomic alteration was detected in 11 of the 21 cases; these include 62% [5/8] amongst individuals with congenital anomalies and 46% [6/13] amongst individuals with learning disability and seizures. The chromosomes involved in submicroscopic alterations were 2q, 5q, 6q, 7q, 10q 12q, 19q, 22q and Xp. Single copy number variation [CNV] was detected in 8 cases while 3 had multiple CNVs. De novo alteration was confirmed in 3 cases, parental inheritance was established in 1 case, inheritance is under evaluation in other 5 cases while in 2 cases it could not be studied for lack of consent. Array-CGH can be significantly helpful in knowing the specific cause of abnormality and in counseling parents for future pregnancies, amongst patients with congenital anomalies or developmental delay or learning disabilities with an apparently normal karyotype.

912T

Lack of Evidence for Biparental Inheritance of Copy Number Variation Contributing to Pathogenic Features in a Series of Clinical Microarray Cases. *C.E. Cottrell[1], S. Kulkarni[1,2,3], H. Al-Kateb[1], M. Evenson[1], S. Anderson[1], D. Lamb Thrush[4,5], S. Hashimoto[4], A.K. Bailes[4], L. Erdman[4], A. McKinney[4], C. Weber[4], J. Weslow-Schmidt[4], S. Ramsey[4], C. Astbury[4,6], S. Reshmi[4,6], J.M. Gastier-Foster[4,5,6], R.E. Pyatt[4,6].* 1) Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, MO; 2) Department of Pediatrics, Division of Genetics and Genomic Medicine, Washington University School of Medicine, St. Louis, MO; 3) Department of Genetics, Washington University School of Medicine, St. Louis, MO; 4) Department of Pathology and Laboratory Medicine, Nationwide Children's Hospital, Columbus, OH; 5) Department of Pediatrics, The Ohio State University, Columbus, OH; 6) Department of Pathology, The Ohio State University, Columbus, OH.

Several models evaluating the effect of copy number variation (CNV) have been proposed to account for the clinical features in patients bearing multiple copy number alterations. For example, a two-hit model has been suggested for cases of developmental delay in probands inheriting a microdeletion at 16p12.1 along with a second large alteration (Girirajan et al., 2010). Whenever possible, studies include CNV origin; however, few data focusing on the effect of biparental inheritance exist, which may be attributed to the low frequency of such events. One recent study examining CNV in 1,227 probands with neurodevelopmental phenotypes failed to identify a single instance of biparental inheritance (Girirajan et al., 2011). Consequently, we undertook a retrospective study to examine instances of biparental inheritance of CNV interpreted as variants of unknown significance (VUS) in a series of 6,322 clinical cases referred for microarray analysis at Nationwide Children's Hospital or Washington University; 47 probands were identified with 2 or more VUS with at least 1 variant inherited from each parent. No common combinations of CNV were observed in more than 1 family. Paternally inherited duplications involving the *PARK2* gene (6q26) were observed in two unrelated probands, as were paternal duplications of the *CHRNA7* gene (15q13.3), maternal deletions involving the *CTNNA3* gene (10q21.3) and maternal alteration of the *OPCML* gene (11q25). Analysis of the second biparentally inherited CNV in each of these 8 unrelated proband pairs failed to identify genes with common pathways or expression patterns suggesting that their co-inheritance may be a chance event. Interesting combinations of VUS with biparental inheritance were observed within families when multiple members underwent array analysis. One sibling-pair shared a maternally inherited 7p22.2p22.1 duplication and a paternally inherited 6q24.3 duplication, while a third sibling inherited only the paternal 6q24.3 gain. Three unrelated probands were identified with 3 inherited VUS and a fourth with 4 variants (2 paternal and 2 maternal) which may represent large CNV burdens in these cases. While this study is hindered by the limited availability of both parental samples in additional cases, these results do not support biparental inheritance as a mechanism of pathogenesis. It is possible that total CNV burden in relation to genetic content may prove a more likely mechanism for pathogenicity.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

913W

7q33 deletion in a family with intellectual disability, dysmorphic features and behavioral changes. *A. Fortuna[1], F. Lopes[2,3], S. Sousa[2,3], J. Silva[1], C. Gomes[2,3], P. Maciel[2,3].* 1) Genetics, Inst.Ricardo Jorge/Centro Geneti, Porto-Portugal, Porto, Portugal; 2) Life and Health Sciences Research Institute (ICVS), School of Health Sciences, University of Minho, Braga, Portugal; 3) ICVS/3B's - PT Government Associate Laboratory, Braga/Guimarães, Portugal.

Intellectual disability (ID) is one of the most frequent and disabling neurological impairments in school-age children. Among the genetic causes of ID, genomic imbalances are one of the most frequent. With the introduction of array comparative genomic hybridization in the routine diagnostic allows the discovery of new and previously undetermined imbalances, as well as the discovery of new ID associated genes. In this work we report a maternally inherited 7q33 deletion in a 19 year old boy with mild ID (IQ= 53), familial history of ID (mother and sister affected; a second-grade cousin with ID, dysmorphic features and epilepsy), seizures, reduced vision, hypotony, dysmorphisms, aggressiveness and hyperactivity. The same alteration is present in his sister that displays mild ID (IQ=62), short stature, facial dysmorphisms (similar to the ones present in the brother) and behavioral changes (disinhibition). The mother cannot read or write and has facial dysmorphisms similar to the ones present in the daughter. The alteration was determined by aCGH analysis using a human genome CGH Agilent 180K custom array with a mean resolution of 17 Kb (Agilent, Santa Clara, CA) and the confirmation studies and inheritance analysis were carried out by qPCR using a fragment designed inside the altered region. aCGH revealed a 2 Mb deletion at chromosome region 7q33 containing 15 genes (AGBL3, AKR1B1, AKR1B10, AKR1B15, BPGM, C7orf49, CALD1, CNOT4, EXOC4, LRGUK, NUP205, SLC35B4, STRA8, TMEM140, WDR91). One of the most interesting genes covered by the deletion is the CNOT4 (CCR4-NOT transcription factor complex, subunit 4) gene. The CNOT4 protein plays both positive and negative roles in transcriptional regulation and a positive role in transcriptional elongation. In yeast, the ortholog of CNOT4 (Not4) regulates the expression of Jhd2 (the yeast ortholog of JARID1C, for which mutations have been described in patient with X-linked ID). Comparison with similar cases, expression and functional studies should help us clarify the relevance of the deleted genes for ID.

914T

Investigation of the Parental Origin and Genomic Mechanisms Involved in *de novo* Pathogenic CNVs in Congenital Disorders. *S. Hayashi[1,2], M. Naganawa[1], D.T. Uehara[1], J. Inazawa[1,2,3].* 1) Dept Molec Cytogenetics, Med. Res. Inst, Tokyo Med & Dent Univ, Tokyo, Japan; 2) Hard Tissue Genome Research Center, Tokyo Med & Dent Univ, Tokyo, Japan; 3) Bioresource Center, Tokyo Medical and Dental University, Tokyo, Japan.

We have analyzed more than 700 patients presenting with intellectual disability and/or multiple congenital anomalies of unknown etiology and normal karyotype by using several types of microarrays, and we detected pathogenic copy number variants (pCNV) in more than 20% of the cases. In most of them the pCNVs were *de novo*. Based on these results, we attempted to clarify the mechanisms responsible for the formation of these rearrengements and their parental origin. For this purpose, we determined the precise size of each pCNV using a high resolution oligonucleotide array (Human CGH 2.1M Whole-Genome Tilling Arrays, Roche-Nimblegem) and read the sequence at the breakpoint region to classify each pCNV into the three major mechanisms underlying genomic rearrengements, namely non-allelic homologous recombination (NAHR), as cause of recurrent rearrangements, and non-homologous end joining (NHEJ) or microhomology-mediated break-induced replication (MMBIR), as causes of non-recurrent rearrengements. We analyzed the breakpoints of 16 microdeletions and found that four cases (25.0%) were produced by NAHR. On the other hand, eight cases (50.0%) were originated by MMBIR, with 1–7 bp of microhomology, and two cases (12.5%), by NHEJ. The other two cases were produced by complex rearrengements. We also analyzed parental samples of nine cases and found that in eight (88.9%) the event leading to the rearrengement occurred in the paternal allele. These results were quite contrasting to our analysis in another cohort of healthy population, in which almost half (13/28=46.4%) of *de novo* CNVs had paternal origin. Our study demonstrates that *de novo* pCNVs were frequently produced by non-recurrent mechanisms rather than by the recurrent mechanism, with most of the events occurring in the paternal allele. This also suggests that *de novo* pCNVs may be generated by problems related to spermatogenesis, which is concordant with previous reports. We are currently sequencing the breakpoint regions of benign *de novo* CNVs in healthy individuals in order to elucidate the mechanisms producing these rearrangements.

915W

Our experience of clinical implementation of whole-genome array CGH as a first-line test in postnatal cases. *V. Malan[1,2], JM. Lapierrre[1], MC. de Blois[1,2], C. Ozilou[1], S. Nusbaum[1], O. Raoul[1,2], M. Le Merrer[1,2], A. Harroche[1,2], R. Caumes[1,2], G. Baujat[1,2], F. Cartault[3], M. Rio[1,2], J. Amiel[1,2], S. Lyonnet[1,2], V. Cormier-Daire[1,2], A. Munnich[1,2], M. Vekemans[1,2], S. Romana[1,2].* 1) Département de Génétique, Hopital Necker-Enfants Malades, Paris, France; 2) Université Paris Descartes; 3) CHU de la Réunion, France.

Today, the array CGH has supplanted the karyotype for the global study of the genome with a resolution from 10 to 500 times better and with providing access to the genes content of the chromosomal rearrangement. The study of patients with intellectual disabilities and/or congenital malformations allowed to detect thanks to this technique 10 to 15% of abnormalities not visible on the karyotype. In addition, its use extended to other indications revealed anomalies associated with neuropsychiatric diseases (such as autism, schizophrenia, bipolar disorder or isolated epilepsy) and cardiac, bone, kidney isolated defects. The array CGH freeing a morphological analysis of chromosomes of the karyotype in contrast, offers the advantage of being fully automated allowing improved reliability and reproducibility of the results and a study of more patients with a constant medical and technical effective. Since April 2011, all patients have referred to Medical Genetics Center of Necker Hospital with a syndromic or isolated intellectual disability and / or autism spectrum disorders and / or birth defects benefit from evaluation of array CGH as a first line test. To date, among the 1400 patients tested by whole genome array CGH, we detected 12% of genomic imbalances considered as pathogenic (with about 2% of abnormalities that could have been visualized on the karyotype). The limit of detection of an abnormality was set at 400 kb (international standards) to reduce the detection of benign CNVs and variants of uncertain significance. Besides the expected increase in the efficiency of cytogenetic diagnosis, this change in practice has led to a real breakthrough in the organization of our cytogenetic lab. Complementary techniques of the karyotype, molecular cytogenetic techniques for the study of chromosomal anomalies are now the main part of the cytogenetic department. The analysis of genes in the unbalanced chromosome segment to establish the causal link between the genomic imbalance identified and the pathology observed is at the core of the cytogeneticist's work. This requires to redefine the training of technicians as well as those physicians who become today “internists of genomic pathology”.

916T

Copy number variation in hereditary colorectal cancer. *A.L. Martin[1,2,4,5], B.A. Talseth-Palmer[1,2], D.M. Grice[2,5], G.N. Hannan[5], R.J. Scott[1,2,3].* 1) University of Newcastle, Newcastle, NSW, Australia; 2) Hunter Medical Research Institute, Newcastle, NSW, Australia; 3) Hunter Area Pathology Service, Newcastle, NSW, Australia; 4) Australian Rotary Health-Rotary District 9650, Parramatta, NSW, Australia; 5) Commonwealth Scientific Industrial Research Orgaization, North Ryde, NSW, Australia.

Rationale: Colorectal cancer (CRC) is one of the most common cancers in the industrialised world. Hereditary non-polyposis colorectal cancer (HNPCC) is the most common form of familial CRC which describes the clustering of epithelial malignancies. Mutations in DNA mismatch repair (MMR) genes (MLH1, MSH2, MSH6 and PMS2) are clearly associated with as subset of HNPCC, termed Lynch syndrome (LS). Despite the distinct similarities of HNPCC to LS, anywhere from 30–50% of clinically tested patients fail to show germline mutations in these genes. Likewise, familial adenomatous polyposis (FAP) is the second most common form of inherited colorectal cancer characterized by the development of 100's-1000's adenomas in the colon and rectum. In 80–90% of FAP cases, disease is attributed to mutations in the adenomatous polyposis coli (APC) gene involved in regulating the WNT signalling pathway. In 10–20% FAP cases, the genetic cause of the disease remains elusive. Copy number variation (CNV) is the most recently described form of genetic variation and is yet to be fully characterized in hereditary CRC. Objective: The aim of this study is to identify if CNVs can account for a proportion of patients without a genetic diagnosis of LS or FAP. Methods and Results: 198 HNPCC, 72 FAP and 50 healthy control samples were screened for CNVs using the Cytogenetic Whole Genome 2.7M and CytoHD arrays from Affymetrix. All CRC cases had been deemed mutation negative after routine genetic testing (MLPA and Sanger sequencing). Golden Helix SVS7 CNAM was used to analyse the results. The cases were compared to controls as a means to eliminate non-pathogenic CNVs. Genome-wide analysis has identified hundreds of CNVs, some likely to be associated with HNPCC and FAP. This includes CNVs which encompass coding, non-coding and regulatory regions of genes. Single sample analysis also identified numerous cases exhibiting unique CNV profiles. This includes a patient with multiple chromosome p-arm deletions and a patient with a large deletion in the vicinity of the MLH1 gene, which is likely to be associated with disease risk. Conclusions: This study has generated a catalogue of CNVs likely to be associated with HNPCC and FAP even though the significance of any one of these remains unclear. This study supports the need for genome-wide screening for patients as a mean of more accurately delineating the genetic diagnosis in inherited cancer predispositions.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

917W

Using macrocephaly as a quantitative endophenotype in autism patients to enhance the identification of novel candidate genes. *J. Reiner[1], D. Moreno-De-Luca[1], E.B. Kaminsky[1], D.H. Ledbetter[2], C.L. Martin[1]*. 1) Department of Human Genetics, Emory University, Decatur, GA; 2) Genomic Medicine Institute, Geisinger Health System, Danville, PA.

Autism spectrum disorders (ASD) are a heterogeneous group of neurodevelopmental disorders characterized by deficits in social reciprocity, impaired language skills, and restricted interests. Although the degree of social and cognitive impairments in autism varies widely, individuals with analogous copy number variants (CNVs) often exhibit physical similarities. As such, selective genotyping of probands with specific endophenotypes may enhance the discovery of novel autism genes. We previously reported an increased incidence of autism and macrocephaly in individuals harboring a recurrent 17q12 deletion. Other recurrent CNVs have also been shown to impact head circumference in a dosage sensitive manner. For example, the 16p11.2 deletion and reciprocal duplication have been reported to predispose individuals to macrocephaly and microcephaly, respectively. Therefore, the objective of this study is to use autism plus macrocephaly as a quantitative trait to identify causative genes associated with these phenotypes. Since head circumference can be influenced by heritable factors and pathogenic mutations, we restricted the scope of our study to macrocephalic probands whose parents have a normal head circumference. We used the Simons Simplex Collection (SSC), a cohort of 2,629 families with a single individual with autism, and identified 180 probands (6.8%) with autism plus de novo macrocephaly. The SSC was recently genotyped by SNP arrays and two probands harboring large pathogenic CNVs were excluded from our experimental group, leaving 178 probands for our study. Interestingly, one of the excluded probands carries a 16p11.2 deletion, demonstrating the validity of our approach. In accordance with reports of a higher incidence of autism in males, the ratio of males to females with de novo macrocephaly is 4.8:1. However, the proportion of males and females exhibiting de novo macrocephaly relative to the total number of probands in each group is comparable, suggesting that other factors, including de novo mutations, may play a greater role in determining thisendophenotype than gender. We are interrogating the genotype of these probands using a custom-designed, exon-focused array with coverage of 1,282 genes including neurodevelopmental candidate genes and genes within 13 recurrent autism CNV regions. In light of our enrichment strategy, our approach should facilitate the identification of exon-level CNVs in genes associated with autism and de novo macrocephaly.

918T

Study of Genomic Changes in Children with Intellectual and Developmental Disabilities Using Chromosomal Microarray. *P. Sharma, N. Gupta, M.R. Chowdhury, R. Shukla, M. Ghosh, M. Kabra*. Genetic unit, Department of Pediatrics, All India Institute of Medical Sciences, New Delhi, India, 110029.

Intellectual disability (ID) / developmental delay (DD) is a highly diverse group of disorder with multiple etiologies. Despite extensive investigations, the cause remains unknown in approximately 50% of the patients. Cryptic chromosomal imbalances have been identified as a common cause of disabilities in these children which are too small to be detected by conventional karyotyping due to its limited resolution. Chromosomal microarray (CMA) technology has emerged as a major tool to identify genomic imbalances in children with ID/DD due to whole genome coverage at higher resolution. In this study we analyzed 78 patients (50 male and 28 female) with unexplained ID/DD aged six months to eighteen years using CMA. These children were evaluated by Clinical Geneticists and were tested using CMA after no known cause for ID/DD was established by routine work up such as syndromic database search, G-banded karyotyping, MRI (brain) and metabolic testing. We used Agilent CGH Microarray Kit 8×60K and IIlumina Cyto-SNP -12 chip for our study. Data was analysed using Agilent Genomic Workbench and Genome Studio software. Regions with abnormal genotype were cross-matched with available literature and databases such as ISCA, DECIPHER, DGV to establish the pathogenicity of the changes observed. We identified significant copy number variations (CNVs) in 12 (15.3%) patients. In seven cases CNVs were in the subtelomeric region and in five they were in the interstitial region. Eleven of these CNVs were confirmed by other methods such as Multiplex Ligation Dependent Probe Amplification (MLPA) and Fluoroscence In Situ Hybridization (FISH). We were able to provide genetic counseling to all families and prenatal diagnosis in one family. In three other cases, causal relationship of CNV with ID/DD was not clear. CMA helped in increasing the diagnostic yield for cases with unknown genetic cause which is important for patient care in several ways: i) It decreases the number of diagnostic procedures that a patient is subjected to. ii) For families a diagnosis may diminish anxiety. iii) It allows for proper risk recurrence estimate and prenatal diagnosis in future pregnancies. Online databases and studies published on children with ID/DD using CMA has helped in interpreting the data and correlating the phenotype of patient with the genotype observed. Interpreting the clinical significance of novel or extremely rare CNVs when limited clinical literature is available is still challenging.

919W

Association of rare copy number variations and risk for Alzheimer's disease. *D. Villela[1], D. Schlesinger[2], C.K. Suemoto[3], L.T. Grinberg[3], A.C. Krepischi[4], C. Rosenberg[1]*. 1) Department of Genetics and Evolutionary Biology, Institute of Biosciences, University of Sao Paulo, São Paulo, SP, Brazil; 2) Israel Institute for Teaching and Research Albert Einstein, São Paulo, SP, Brazil; 3) Brazilian Aging Brain Study Group, Department of Pathology, University of Sao Paulo, São Paulo, SP, Brazil; 4) National Institute of Science and Technology in Oncogenomics, São Paulo, SP, Brazil.

Alzheimer's disease (AD) is the most common cause of dementia in the elderly. Neuropathologically, the disease is characterized by extracellular senile plaques containing β-amyloid (Aβ) and intracellular neurofibrillary tangles containing hyperphosphorylated tau protein. So far, four genes have been definitively implicated in the etiology of AD. Mutations of the genes encoding amyloid precursor protein (APP) and presenilin 1 and 2 (PSEN1, PSEN2) cause the rare hereditary early-onset form of the disease. However, despite several genome-wide association studies using single nucleotide polymorphism (SNP) arrays have been performed over the last years, APOE is still the only confirmed susceptibility locus for late-onset of AD, the more prevalent form of the disorder. Currently, DNA copy number variations (CNVs) are recognized to be a frequent form of common genetic variation and represent a substantial proportion of total genetic variability in human population. Moreover, since the recognition of the existence of CNVs it has increased the number of studies that demonstrate its critical role in the phenotypes of complex diseases. Recently, two independent investigations reported the association of CNVs and risk for AD, but the CNVs were not selected by relevant gene content and they are common on the population. Here we used a whole-genome array-CGH to identify rare constitutive CNVs that possibly contribute to the development of late-onset AD. The Brain Bank of the Brazilian Aging Brain Study Group provided DNA from blood samples of 110 subjects, including 55 individuals with AD and 55 normal controls, matched for age and gender. The partial analysis of array-CGH revealed 11 rare CNVs (15% less frequent or never been documented as variable on the Database of Genomic Variant) in 27 subjects with AD and none of them was recurrent. Because of its relevant gene content for the investigated phenotype, we selected 5 genes that have never been associated to the disease. In summary, our results show the existence of novel susceptibility genes for the late-onset AD and reinforce the role of rare CNVs as disease-susceptibility variants for the investigated phenotype.

920T

Coronary artery disease genomics - CNV findings in the group of patients from Lithuania. *V. Kucinskas[1], I. Pepalyte[1], A. Pranculis[1], I. Bagdonaite[2], V. Dzenkeviciute[3], Z. Kucinskiene[2]*. 1) Department of Human and Medical Genetics, Faculty of Medicine, Vilnius University, Vilnius, Lithuania; 2) Department of Physiology, Biochemistry and Laboratory Medicine, Faculty of Medicine, Vilnius University, Vilnius, Lithuania; 3) Clinic of Cardiology and Angiology, Faculty of Medicine, Vilnius University, Vilnius, Lithuania.

We aimed to evaluate quantitative and qualitative characteristics of structural variation in the group of patients with coronary artery disease (CAD). The strategy was to gather trios to identify *de novo* variants in the probands' group (32 trios were recruited). The probands' group included individuals with angiografically determined premature CAD or myocardial infarction and evaluated biochemical markers in blood samples. Genotyping was performed using Illumina 770K HumanOmniExpress-12 v1.0 array. Data was analyzed with QuantiSNP v2.1.

We presumed that parents of the probands were healthy individuals and focused only on the probands' CNV analysis, especially the *de novo* variants. In total there were 90 autosomal CNVs identified in the probands' group. According to the origin 60 CNVs were inherited and 32 - *de novo*. Out of 60 inherited variants 33 were deletions (25 heterozygous, 8 homozygous) and 27 - duplications (25 heterozygous, 2 homozygous). Out of 32 *de novo* variants 20 were deletions (4 heterozygous, 16 homozygous) and 12 - duplications (11 heterozygous, 1 homozygous). The most intriguing characteristic is the zygosity ratio of the *de novo* deletions compared to the *de novo* duplications and all inherited variants (homozygous and heterozygous duplications and deletions). *De novo* CNVs were classified using Database of Genomic Variants: out of 32 variants, 13 were classified as polymorphisms, 19 were not in the database - classified as unknown. We have also searched databases for CNVs' gene content. More than 50% of *de novo* CNVs did not include genes. Nevertheless, there were CNVs that included genes, which may provide a potential phenotype explanation.

According to our findings the zygosity's pattern in *de novo* variation could be one of the disease characteristics. *De novo* homozygous deletions, which are the dominating type of CNVs among all *de novo* variants, did not include relevant genes. Moreover not every gene spanning CNV could provide even a hypothetical reason for the pathogenesis. Thus, the link between pathogenesis of CAD and genomic findings on structural variation still remains unsolved. All in all our results improve knowledge of structural variation and provide the basis for better understanding of CAD genomics.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

921W

An unbalanced translocation causing a 21q11.2 deletion in an infertile male. *I. Holm[1], K. Eiklid[1], C. Hauge[2], V. Rønning[2], H. Wold[2]*. 1) Medical Genetics, Oslo University Hospital, Oslo, Norway; 2) Oslo and Akershus University College, Oslo, Norway.

Infertility is thought to affect about 10–15 percent of all couples in Western countries, where male infertility is the cause in about 50 percent(1). About 15 percent of male infertility is caused by genetic abnormalities(2). Many studies have shown that chromosomal aberration causing meiotic interruption can lead to oligozoospermia or azoospermia, which again can lead to infertility(3). Reciprocal translocations carriers do often produce more sperms with unbalanced chromosomes than non carriers(4). Up to 1500 genes are thought to be involved in spermatogenesis and 300–600 genes are specifically expressed in male germline. Deletions and duplications in genes that affect spermatogenesis increase the risk of infertility (1). Here we present a case report of an infertile man with a presumed balanced reciprocal translocation of 46,XY,t(12;21)(q24.33;q10). By performing array comparative genomic hybridization (Agilent Technology, 105K) we found a deletion starting at 21q11.2 going at least 0.7 Mb up to the telomeric part of chromosome 21 (arr 21q11.2(14,795,268-15,558,340)x1). We speculate if this deletion also could include the whole telomeric region. This deletion included three OMIM genes *LIPI* (OMIM *609252), *POTED* (OMIM *607549) and possibly *TPTE* (OMIM *604336) if it includes the entire short arm. These genes are all expressed in testis and are candidate genes for infertility. *TPTE* and *POTEB*, which in addition to *POTED* is a member of the ankyrin domain family, are previously described in relation to infertility(5). The deletion found in this patient will be further characterized with quantative PCR and FISH in order to confirm it and to characterize the breakpoint and the size of it. References: 1 Tüttelmann et al.(2011), 2 Bianco et al.(2009), 3 Juchniuk de Vozzi et al.(2009), 4 Cifuentes et al.(1998), 5 Guediche et al.(2012).

922T

Single cell copy number profiling in a 24-hour workflow with oligo aCGH. *P. Costa, S. Fulmer-Smentek, A. De Witte*. Agilent Technologies Genomics, Santa Clara, CA.

Since its development, array-based Comparative Genomic Hybridization (aCGH) has become a widely used tool in cytogenetic and cancer applications for researchers to study genome-wide copy number (CN) changes. However, not all cells in a tissue share the same genomic background and so the ability to characterize individual genomes is becoming an area of focus in cancer, reproductive and stem cell research. Taking advantage of the high-resolution and sensitivity of the Agilent SurePrint G3 Microarrays technology, we applied the power of the aCGH platform to CN profiling in single cells. The use of Agilent's aCGH overcomes the limitations posed by the low-resolution of traditional FISH and PCR-based research techniques and, more recently, the low-resolution and poor reproducibility associated with BAC arrays. Agilent's aCGH application relies on a two-color assay to measure CN changes in a test sample relative to a reference. To increase the amount of DNA while maintaining the genomic representation, whole genome amplification (WGA) was performed. Single cells biopsied from embryos and genomic DNA (gDNA) with a known aberration diluted to single cell levels were amplified. To minimize variation associated with the extreme dilution of gDNA and the amplification process, the reference sample, which was also diluted to single cell levels, was amplified in multiple reactions and pooled prior to labeling. Test and reference amplified DNA samples were differentially labeled, combined and hybridized for 16 hours to SurePrint G3 Human Catalog 8×60K CGH Microarrays, with uniform backbone coverage and denser coverage in genes. Following hybridization, the data were extracted and analyzed for CN alterations using algorithms implemented in Agilent CytoGenomics and Genomic Workbench. With an optimized workflow for sample and microarray processing, CN aberrations were assessed in the single cells and confirmed in the gDNA, bearing a known amplification on chromosome 9 and diluted to single cell levels. A reduction in the noise level and improvement in the data was observed for samples hybridized to a pooled reference as compared to a single reference. Using an approximately 24-hour single cell workflow from sample preparation to analysis, researchers can now obtain reliable results at a high resolution while remaining cost-effective.

923W

Four new polymorphic CNVs in cohort of 100 Bulgarian patients with intelectual disability/congenital malformations and autisme. *S.P. Hadjidekova[1], D.M. Avdjieva-Tzavella[2], B.B. Rukova[1], D.V. Nesheva[1], R.S. Tincheva[2], D.I. Toncheva[1]*. 1) Medical Genetics, Medical University-Sofia, Sofia, Sofia, Bulgaria; 2) State University Pediatrics Hospital “Queen Evdokia”, Medical Faculty, Medical University, Sofia, Bulgaria.

The increasing resolution of DNA-microarrays and the techniques optimization allow detection of a large number of small CNVs, whose clinical significance in some cases is unknown. Oligo array-CGH was applied in 100 patients, divided in two groups - with intellectual disability/multiple congenital anomalies and autism in order to unravel the underlying genetic abnormalities. We have used BlueGnome CytoChip oligo 2×105K microarray, v1.1, with 35 kbp backbone resolution. This is the first whole genome high resolution screening in a larger group of Bulgarian patients. Except causative CNVs we found 299 benign CNVs and 219 CNVs with unknown clinical significance. The interpretation of CNVs with unknown clinical significance remains still a challenge. Due to insufficient data for the Bulgarian population we made an individual assessment of each unknown variant. In our study four types of unknown CNVs in chromosomal loci 2q37.3, 10q11.22, Xp22.33 and 8p23.1 were found in higher percentage (≥5%). This gave us a reason to suppose that they were probably not pathogenic and are specific for the Bulgarian population. There is an obvious need for large population studies and the creation of detailed maps of CNV frequencies in the Bulgarian population. This would facilitate the precise interpretation of genomic imbalances of unknown nature in clinical aspect and would allow the widespread introduction of microarray diagnosis in practice. The protocol and informed consent documents were reviewed and approved by the local ethics committee. The consent forms were prepared in conformance with the Declaration of Helsinki and local country laws. All participating volunteers were asked to provide written consent after receiving an explanation of the study.

924T

First Detection Of Clonal Trisomy 4 Cells From 3 Cases Of Ossifying Renal Tumor Of Infancy. *J. Liu[1], M. Guzman[1], B. Pawel[2], D. Pezanowski[1], J. Roth[3], G. Halligan[4], J. de Chadarévian1[1]*. 1) Department of Pathology and Laboratory Medicine, Drexel University College of Medicine and St. Christopher's Hospital for Children, Philadelphia, PA; 2) Department of Pathology and Laboratory Medicine, University of Pennsylvania School of Medicine and Children's Hospital of Philadelphia, Philadelphia, PA,; 3) Department of Surgery, Section of Pediatric Urology, Temple University School of Medicine and St. Christopher's Hospital for Children, Philadelphia, PA; 4) Department of Pediatrics, Section of Oncology, Drexel University College of Medicine and St. Christopher's Hospital for Children, Philadelphia, PA.

The ossifying renal tumor of infancy is a rare neoplasm diagnosed in the first two years of life, predominantly in boys. The neoplasm is primarily characterized by the presence of a large ossifying component. Its most common mode of presentation is hematuria, and it has a uniformly benign behavior. The karyotypic make-up of the process has not been reported. Thus a study was undertaken and it allowed demonstration of clonal trisomy 4, which was confirmed by the fluorescent in-situ hybridization (FISH)-probing of two additional archival formalin-fixed, paraffin-imbedded similar tumors. Based on the findings in these 3 cases, it seems that clonal trisomy 4 may be considered as a characteristic of the tumor, which makes it distinct from any other infantile renal tumor.

925W

17p13.3 microduplication syndrome: report of additional findings. *L. Mutesa[1], E. Chabchoub[3], JH. Caberg[2], AC. Hellin[2], M. Jamar[2], H. Peeters[3], G. van Buggenhout[3], G. Pierquin[2], J. Andrieux[4]*. 1) Medical Genetics, National University of Rwanda, Kigali, Kigali, Rwanda; 2) Center for Human Genetics, University of Liège, Liège, Belgium,; 3) Department of Human Genetics, University Hospitals of Leuven, Catholic University of Leuven, Belgium,; 4) Laboratory of Medical Genetics, Jeanne-de-Flandre Hospital CHRU de Lille, Lille cedex, France.

Chromosomal duplication including 17p13 is rare and only a few patients have been reported to date. Here, we report on six additional patients with submicroscopic duplications of 17p13.3 identified by microarray-based comparative genomic hybridization. The common clinical features of all those five patients were marked hypotonic long face, downslanting fissures, low-set ears, high stature, development delay and mental retardation. Apart from the presence of other several genes, the duplicated regions contain two putatives functional genes, the PAFAH1B1 (encoding LIS1) and YWHAE genes which are involved in directing the movement of nerve cells (neuronal migration) and normal brain development, respectively. We suggest that 17p13.3 duplication is responsible for brain development delay and distinctive craniofacial dysmorphic features. The report additional allows to definitely substantiate the significance of 17p13.3 duplication and contributes to delineation of the clinical spectrum.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

926T

Chromosome aberrations among 4617 chromosomal studies at a pediatric Mexican hospital in 19-year period of time. *J. M. Aparicio[1, 13], M. L. Hurtado[2], M. P. Barrientos[3], W. B. San Martin[4], A. F. Nuñez[5], R. G. Ruiz[6], H. T. Gomez[7], S. P. Rodriguez[8], R. M. Zamudio[9], F. L. Cuellar[10], M. A. Cubillo[11], F. P. Sierra[12], M. G. Palma[13], H. O. Chavez[13], S. M. Chatelain[14]*. 1) Dept Gen; 2) Cytogenetics; 3) Endocrinology; 4) Estomatology; 5) Neumology; 6) Internal Medicine; 7) Alergology; 8) Neurosurgery; 9) Cardiology; 10) Urology; 11) Rehabilitation terapy, Hosp para el Nino Poblano, Puebla; 12) Genetics, Hosp de la Mujer, Puebla; 13) Estomatology, Benemerita Universidad de Puebla; 14) Biotechnology, Universidad Autonoma Metropolitana, Mexico.

Mutations or chromosome aberrations are considered alterations in the chromosome number or structure. They are mainly considered due to gametogenesis inborn error (meiosis) or during the zygote first cellular divisions. All these alterations might be observed during metaphase from the cellular cycle, where DNA loses are seen (clastogenic processes) due to DNA repair processes deficiency o total absence, among others. 4617 chromosomal studies were performed at Hospital Para El Niño Poblano (Pediatric Hospital) in Mexico. During 19 years period of time (from 1992 to 2011) were 34.6% (1596 patients) showed different chromosomal alterations. Among the studies population, male and female pediatric patients with different genetic diseases were chosen. These chromosome changes are classified as numeric or structural alterations, respectively. Another group of genetic alterations are known as mutations and can be inherited among generations. A wide variety of pediatric patients with genetic diseases due to chromosome aberrations are described in this study analyzing their clinical characteristics, medical or surgical treatments and their medical evolution according to the genetic disease.

927W

Study of sodium pertechnetate on bone marrow cells of mice, in vivo: Micronucleus assay Aranha IP Dept. Genética, IBRAG, Univ. do Estado do Rio de Janeiro, Brazil. *I. Aranha*. Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil.

Technetium-99m (^{99m}Tc) has become the most widely used radioisotope in the detection of inflamatory sites as well as in the diagnosis of transplanted tissues. The goal of the present work is to study the effect of sodium pertechnetate on chromosomes of mice bone marrow in vivo, using the micronucleus assay. Animals (ICR mice) were separated into three groups. In the first group, 6 animals received sodium pertechnetate intraperitoneally in a concentration equivalent to 25% of the LD50, once. In the 2nd day, animals were sacrificed by cervical dislocation, their femurs removed, the bone marrows collected and smears were made for slides preparation. After 24h cells were stained with Giemsa Gurr (2%) and analyzed under optical microscope. As positive control, 6 animals received cyclophosphamide (50mg/mL) also once. Six animals not exposed to any drug served as negative control for the experiment. In the test group, 12130 cells were observed and 143 showed micronuclei. In the positive control group, of 12000 cells analyzed, 163 had micronuclei and in the negative control group, 12173 cells were studied and 07 had micronuclei. The chi-square test for independence showed that our results were extremely significant ($P<0.0001$). They suggest that sodium pertechnetate is responsible for the micronuclei observed.

928T

Cryptic translocactions of acrocentrics chromosomes in normal couples and couples with aneuploid miscarriages. *S. Ramos[1], O. Castro[1], B. Molina[1], M. Angeles[1], P. Grether[2], D. Mayen[3], S. Frias[1,4]*. 1) Laboratorio de Citogenetica, Instituto Nacional de Pediatria, Mexico city, DF, Mexico; 2) Departamento de Genetica, Instituto Nacional de Perinatologia, Mexico city, DF, Mexico; 3) Servicio de Genetica, Hospital Angeles -Lomas, Mexico city, DF, Mexico; 4) Instituto de Investigaciones Biomedicas, Universidad Nacional Autonoma de Mexico, Mexico city, DF, Mexico.

Aneuploidy is the principal genetic cause of spontaneous abortions, approximately 90% of cases die in utero, and congenital anomalies are produced in live births. Little is known about the etiology of aneuploidy, although the main mechanism is meiotic nondisjunction, especially in the oogenesis of older women. The structural chromosomal rearrangement called cryptic translocations of acrocentric chromosomes (TCA) could generate aneuploidies by influencing meiosis and producing non-disjunction, and aneuploid conceptions. The aim of this study was to determine the frequency of TCA in couples with abortions with aneuploidies of acrocentric chromosomes. We included 59 healthy couples with at least one normal child and no history of spontaneous abortions; 12 couples with at least one abortion or offspring with aneuploidies of acrocentric chromosomes and 58 samples of aneuploid abortions. All couples agreed and signed informed consent form. All samples were subjected to GTG banding and FISH with centromeric probes for chromosome 13/21, 14/22 and 15. For both methods, at least 15 cells were analyzed per sample. Using FISH, we found a frequency of 18.6% TCA (22/118) in the individuals from the healthy couples group; 29.1% (7/24) in individuals from couples with abortions or aneuploid offspring, and 12.06% (7/58) in the abortions. The translocation more frequently found was: dic(13;15), followed by dic(14;15). The results found in the normal couples suggest that TCA may be a chromosomal polymporhism, however the high frequency of TCA we found in the couples with miscarriages with aneuploidy of acrocentric chromosomes, suggest that this TCA may interfer with the meiotic chromosome pairing favoring aneuploid conceptions. It is necessary to increase the population of study to confirm whether TCA are related to the generation of aneuploidy.

929W

A comparative cytogenetic resolution analysis between classic chromosome aberrations and Cytokinesis-Block Micronucleus Cytome assay. *R. Saraswathy.* VIT University, Vellore, India.

Until recently the genetic instability of the syndrome was estimated by the application of classic chromosome aberrations analysis. However at present after the discovery of Cytokinesis-Block Micronucleus Cytome assay by Fenech, the situation has changed. In this study a comparative cytogenetic resolution analysis between classic chromosome aberrations and Cytokinesis-Block Micronucleus Cytome assay was undertaken. For this study the following syndromes were utilized, such as, Dilated Cardiomyopathy (n=50), Diabetes (n=25), Head and Neck Cancer (n=15), congenital anomalies (n=15), Gonadal dysgenesis (n=15), Down syndrome (n=20), Aplastic anemia (n=15). Age and sex matched controls (n=50) from Vellore region of South India were analysed for CA and CBMN. Our results indicate that in all the syndromes a significantly higher frequencies of CA, micronuclei (MNi) and genomic damages such as nuclear buds (NBUDs) and nucleoplasmic bridges (NPBs) were observed in the patient group in comparison to the control group ($p<0.001$). Even though CBMN assay is having several advantages over the other, yet it is advisable to use both the type of assays to get a better picture of genomic instability.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

930T

In vitro assessment of genotoxicity and cytotoxicity of methyl mercury chloride. *S. Tayel[1,2], M. Naim bushra[2], A. Abd Ruboh[1], S. Hilal[2], A. El Agwany[2]*. 1) Alexandria Regional Centre for Women's Health & Development, Alexandria, Egypt; 2) Anatomy & Embryology Deptartment, Faculty of Medicine, Alexandria University, Alexandria, Egypt<.

Methyl mercury chloride (CH3HgCl) exists in soil, air, fishes, shellfish and algae, rejected water system, from mining operations and lixiviation of soils after deforestation. It has become one of the most hazardous environmental pollutants reaching levels of potential toxicity especially in our aquatic ecosystems. The objective of the present study was to assess the cytotoxicity and genotoxicity of CH3HgCl in cultured human lymphocytes. Short-term human leukocyte cultures from 30 healthy donors (15 females and 15 males) were set-up by adding drops of whole blood in a complete medium. Cultures were incubated at 37°C for 72 h and doses O, 50, 500 and 5000 μg/ L of CH3HgCl were added 24 h before harvest. One hundred metaphases were analyzed by solid stain and GTG banding for each culture with a total of 3,000 metaphases for each concentration. Genotoxicity of CH3HgCl was expressed by chromosome aberrations in the form of aneuploidy, hyperploidy, chromatid endoreduplication, breaks, satellite associations, and double minutes formation in a concentration related manner where a significant increase ($P < 0.05$) in the relative frequency of chromosome aberrations was observed for all concentrations of CH3HgCl when compared to control. Cytotoxicity as assessed by the micronucleus test showed binucleated cells with one or more micronuclei and cytoplasmic bridges in a concentration related manner. Mitotic index showed an inverse relation with increase in the concentration of CH3HgCl indicating its cytotoxicity. In conclusion, CH3HgCl is geno- and cytotoxic even in low doses. The changes it induces in lymphocyte culture may be an early marker of this toxicity in humans and should be taken into account in the preliminary evaluation of the risks to populations exposed in vivo.

931W

The deletion 5q in myelodysplasia : correlation of cytogenetic, peripheral blood and bone marrow findings. *V.M. Srivastava, S. Yuvarani, U. Sitaram, A. Nancy, R. Ahmed, A. Abraham, A. Vishwabandya, B. George, V. Mathews, A. Srivastava.* Christian Medical College, Vellore, Tamil Nadu, India.

Background: Deletion of the long (q) arm of chromosome 5 (del 5q) is seen in about 15% of myelodysplasia (MDS). MDS with a solitary del 5q have a favourable outcome. When additional chromosomal abnormalities are present, prognosis is poor with early progression to leukemia, resistance to treatment and short survival. Patients and methods:The study group consisted of patients with a morphologic diagnosis of MDS and del 5q seen in the Department of Haematology between January 2003 and December 2010. At least 15–20 G-banded metaphases were studied from each patient. Cytogenetic findings were correlated with peripheral blood and bone marrow findings. Results: The del 5q was seen in 60 (11%) of the 537 patients with MDS. The median age was 56 years (range 3–76).There were 34 males (57%). The median haemoglobin was 7.2g/dl (range 2.5–14.2 g/dl), median WBC 4.6×109/L(range 0.5–48 ×109/L), and median platelet count, 118×109/L (range 5 -794 109/L).The morphologic diagnoses were: 5q- syndrome -12 ;refractory anaemia (RA) - 3;refractory anaemia with ringed sideroblasts (RARS)-1;refractory anaemia with excess blasts(RAEB)-14; refractory cytopaenia with multilineage dysplasia (RCMD) - 9;RCMD with ringed sideroblasts (RCMD-RS) - 4;hypoplastic MDS - 4;MDS unclassified(MDS-U) -13. Hypolobate megakaryocytes were noted in 14 patients (23%),dyserythropoiesis and dysmyelopoiesis in 42 each (70%)and trilineage dysplasia in 29(48%).The del 5q was solitary in 50% of patients. A single additional abnormality was seen in 13% of patients while 37% had complex karyotypes(more than two abnormalities). Most (57%) patients with a solitary del 5q were adult females (age range:27–76) and almost half (48%) had either refractory anaemia or the 5q- syndrome. RAEB and MDS-U accounted for the majority of patients with complex karyotypes (68%) or a single additional abnormality (63%).The additional abnormalities seen were monosomies 7,17,20 and 13, trisomy 21 and deletions 7q and 20q (< 9 % each). Conclusions: The overall incidence of the del 5q in MDS and its association with additional abnormalities is similar to the literature. Our findings differ from the literature with respect to the median age of our patients (lower than in Western countries but similar to reports from other Asian countries) and the frequency of the 5q- syndrome (similar to Western countries but higher than in reports from other Asian countries).This is the first detailed analysis of the del 5q from India.

932T

The quality of the bone marrow sample for successful conventional cytogenetic analysis is important. *E. Tegg[1, 2], E. Raik[2]*. 1) Haematology, University of Tasmania, Hobart, Australia; 2) Haematology, Royal North Shore Hospital, Sydney, Australia.

The role of cytogenetics in the diagnosis and prognosis of a number of haematological malignancies is well recognised. The contribution made by the quality of the sample given to a cytogenetics laboratory to successfully provide diagnostic results is not universally recognised. This was a retrospective control study made after a change to the bone marrow aspiration technique. The test group consisted of the first 65 patients who underwent a bone marrow biopsy requiring cytogenetic analysis after this change. The control group consisted of 65 patients who had undergone a bone marrow biopsy for cytogenetic analysis prior to this change. The parameters measured were disease type, morphological cellularity of the aspirated marrow, cell count in the cytogenetic sample, cytogenetic culture failure rate, number of metaphases obtained and the banding resolution achieved. Results were analysed using Microsoft Excel student t-test. The cellularity of the cytogenetic sample was significantly increased (1.23 × 109/L compared to 0.59 × 109/L; $p<0.001$). The culture failure rate fell significantly from 15% in the control group to 2%. ($p<0.01$). The optimal number of metaphases (20 metaphases per patient) was achieved in 75% of cases compared to 50% in the control group ($p<0.04$). The banding resolution, probably reflecting the increase in the number of metaphases from which to choose, was significantly increased in the test group compared to the control group ($p<0.04$). Information in the literature is sparse when assessing the effect of BM cellularity on the ability to obtain a successful conventional cytogenetic result. This study showed the positive effect of increasing the cellularity of the bone marrow aspirate sample for the successful cytogenetic investigation of haematological malignancies.

933W

Parental interstitial translocation underlie deletion in Langer-Giedion syndrome. *B. Min[1], M. Seo[1], T. Cho[2], W. Park[1]*. 1) Biomedical Science, Seoul National University, Seoul, South Korea; 2) Orthopaedic Surgery, Seoul National University College of Medicine, Seoul, Korea.

Langer-Giedion syndrome (LGS, MIM 150230), also called trichorhinophalangeal syndrome type II (TRPS2) is a contiguous gene syndrome caused by simultaneous deletion of TRPS1 and EXT, manifesting phenotypes of both trichorhinophalangeal syndrome type I and hereditary multiple exostoses. Using array CGH, we investigated genotypes of a family consisting of two siblings affected with LGS, two unaffected siblings, and unaffected parents, in order to explain the recurrence of LGS from unaffected parents. The affected siblings were found to have the exactly same 7.3 Mb interstitial microdeletion in 8q23-q24 containing TRPS1 and EXT1 genes in two patients in an exactly same way. The mother and one of the unaffected siblings were found to have a 1.3 Mb microdeletion in 8q24.1, which shared distal breakpoint with the deleted segment of the affected. In addition, 6.0 Mb microduplication spanning from proximal breakpoint was detected in the other unaffected sibling. Karyotyping and FISH revealed an interstitial translocation t(8;13)(q23–24;q33). Flanking sequences of breakpoint and junctions for 7.3Mb interstitial deletion contained 200bp-long unique sequences in 8q and 13q. Here we reports a family in which parental silent interstitial translocation caused recurrence of LGS, and this phenomenon should be taken into consideration in genetic counseling.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

934T

Concurrent deletions and duplications at 1p36. *M. Gajecka[1,2], J. Karolak[1], J. Shen[3], C. Glotzbach[4], L.G. Shaffer[4]*. 1) Institute of Human Genetics, Polish Academy of Sciences, Poznan, Wlkp, Poland; 2) Department of Pharmaceutical Bacteriology, Poznan University of Medical Sciences, Poland; 3) Children's Hospital Central California, Madera, CA, USA; 4) Signature Genomic Laboratories, PerkinElmer, Inc., Spokane, WA, USA.

Deletions of 1p36 occur in approximately 1 in 5,000 newborns. To date, we have ascertained more than 150 cases with monosomy 1p36, representing four possible classes of rearrangements: pure terminal deletions, interstitial deletions, unbalanced translocations, and complex rearrangements. Here we present five complex rearrangements involving duplications in monosomy 1p36. For each individual, deletion and duplication sizes, and parental origin of the rearrangements were determined using array CGH and genotyping. Deletion and duplication sizes as well the breakpoints locations were different in each patient. To further understand the mechanisms for concurrent deletions and duplications, the rearrangements breakpoints were investigated using molecular cytogenetics and molecular biology methods. Fiber FISH analysis demonstrated either tandem or direct duplications in the cases. Both duplications and deletions were interrupted with insertions of material from 1p. Although numerous experiments were performed and different methodology applied, we were not able to identify all junctions in the rearrangements. Our results show higher than accepted complexity of the studied rearrangements and indicate involvement of multiple mechanisms in the DNA breakage and repair process during rearrangement formation. Support: Polish Ministry of Science and Higher Education, Grant NN301238836.

935W

Multiple structural abnormalities involving 4q33 in a spontaneous abortion. *M.B. Sheridan[1], C. Wigner[2], D.A.S. Batista[2,3], K. Turner[4], C. DeScipio[2,5]*. 1) Institute of Genetic Medicine, Johns Hopkins University, Baltimore, MD; 2) Department of Pathology, Johns Hopkins Medical Institutions, Baltimore, MD; 3) Cytogenetics Laboratory, Kennedy Krieger Institute, Baltimore, MD; 4) Department of Obstetrics and Gynecology, Johns Hopkins Community Physicians, Columbia, MD; 5) Department of Gynecology and Obstetrics, Johns Hopkins University School of Medicine, Baltimore, MD.

We describe a complex karyotype observed in cultured placental villi from a first trimester (8 weeks) loss from a 24 year-old woman with no significant reported drug or environmental exposures. An abnormal male karyotype was seen with multiple cell lines, each with a structurally aberrant chromosome 4 with an apparent breakpoint in band 4q33. Abnormal cell lines, each observed in multiple cultures, include: (1) addition of material of unknown origin to 4q33 in 9 of 20 cells; (2) a derivative chromosome 4 that is the result of an apparently unbalanced translocation between the long arms of chromosomes 1 and 4 with breakpoints at bands q21 and q33, respectively, in 3 of 20 cells; (3) an apparently terminal deletion with a breakpoint at 4q33 in 2 of 20 cells. Additionally, single cell and single culture abnormalities also involving 4q were observed in three cells. Three cells, all present in the same culture, appeared to have two structurally normal chromosome 4 homologs (46,XY). Fetal tissue was not available for analysis; therefore, confined placental mosaicism cannot be ruled out. Pregnancy history includes four previous first trimester losses and a healthy 2 year old daughter. Parental karyotypes are normal. Chromosomal microarray studies to assess copy number and delineate chromosome 4q breakpoints are pending. Collectively, this analysis revealed a complex, unbalanced karyotype with apparent loss of material from 4q33 to 4qter plus gain of various other chromosomal regions. Potential mechanisms to explain these findings include a jumping translocation or the presence of a fragile site at 4q33. Jumping translocations typically involve one donor and multiple recipient chromosomes; however, in this case, the converse is seen. Breakage at 4q33, due to a fragile site, with subsequent repair with multiple chromosomal partners appears to be a possible mechanism. To our knowledge, however, this type of repair at a fragile site has not been reported. In summary, we have described a complex karyotype that includes distinct abnormalities with similar breakpoints that may be the result of a yet undescribed mechanism of chromosomal rearrangement.

936T

When transmission modifies the complexity of familial chromosome rearrangements. *V. GATINOIS[1], G. LEFORT[1], C. COUBES[2], J. PUECHBERTY[1,2], A. SCHNEIDER[1], S. TAVIAUX[1], M. TOURNAIRE[1], M. DI NICOLA[1], M. GIRARD[1], P. SARDA[2], F. PELLESTOR[1]*. 1) Department of Cytogenetics, Hospital Arnaud de Villeneuve, CHRU Montpellier, France; 2) Department of Medical Genetics, Hospital Arnaud de Villeneuve, CHRU Montpellier, France.

Complex Chromosome Rearrangements (CCR) are rare events, defined by the occurrence of at least 3 chromosome breakpoints. The degree of complexity of CCRs can be very high and the related meiotic disturbances result in important chromosomal imbalance in gametes. Only 30% of reported cases are familial and transmission is mainly through males. Conversely 70% of CCRs are de novo, caused by an error during spermatogenesis and frequently lead to male infertility. During meiosis, crossing-over between involved chromosomes can modify the complexity of the rearrangement with numerous possibilities of chromosomal imbalance in gametes. We report 2 familial cases in which meiotic events result in a new chromosome rearrangement. In our first case, karyotyping was done following the discovery of a multiple malformation syndrome on ultrasound examination. The fetal karyotype showed an unbalanced CCR with 3 breakpoints. Subsequent analysis of the mother's chromosomes revealed an apparently balanced CCR with 4 breakpoints. For the second case, a male carrier of a simple reciprocal translocation inherited through his mother fathered a girl with a more complex rearrangement involving a third chromosome. For both cases, breakpoint studies led to an understanding of the involved mechanisms underlying the simplification or complexification of the parental rearrangement.For the breakpoint studies, in addition to the technics used routinely in our lab (metaphase karyotype, chromosome painting and array-CGH) we developed 2 additional methodological approaches : “chromosome walking” and Array Painting.

937W

Evidence of a Trisomy 9 Rescue Event in Amniotic Fluid. *L.E. Northrop, M.J. Macera, V. Johbanputra, B. Levy*. Pathology, Columbia University, New York, NY.

Trisomy 9 is rarely seen in a live birth as the natural outcome is usually a spontaneous abortion within the first trimester. Those cases that do survive are primarily mosaics with a high risk for fetal abnormalities, diminished life expectancy and variable phenotype, depending on the level of mosaicism. The phenotype described in the small number of reported live-birth cases includes: facial dysmorphism, skeletal abnormalities, hypotonia, congential heart defects, skin lesions and severe pyschomotor retardation. At our clinic, a 44 year old, G4P1woman, underwent a transcervical chorionic villus sampling at 13 weeks gestational age because of advanced maternal age. Cytogenetic analysis revealed level III mosaicism for trisomy 9 with the remaining cells showing a 46,XY chromosome constitution. Amniocentesis at 16 weeks of gestation was performed to distinguish between confined placental mosaicism (CPM) and true fetal mosaicism. FISH using a centromere-specific probe for chromosome 9 (CEP 9) was performed on uncultured cells and revealed an extra centromeric signal, indicative of trisomy 9, in 88% on interphase nuclei. Routine chromosome studies (G-banding) on cultured cells revealed a minute supernumerary marker chromosome (SMC) in 85% of cells analyzed. The SMC was identified as being derived from chromosome 9 using the CEP 9 probe. The final karyotype was: 47,XY,+mar[17]/46,XY[3].ish der(9)(D9Z1+). The fetal ultrasound was normal with no indication of malformations. The finding of trisomy 9 mosaicism at the time of CVS coupled with the results from the amniotic fluid analysis indicate that a rescue event likely occurred with one of the chromosomes 9 reducing to a minute, centric fragment. Microarray analysis was recommended to determine the extent of euchromatic material present on the chromosome 9 marker; however the patient elected to terminate the pregnancy. Without analyzing additional fetal tissue, it is uncertain whether the amniocentesis results reflect fetal wide rescue.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

938T

Large-Scale Sequence Analysis of Translocation Breakpoint Junctions. *B. Weckselblatt, M.K. Rudd.* Human Genetics, Emory University, Atlanta, GA.

Translocation is one of the most common structural chromosome abnormalities observed in humans. Germline unbalanced chromosome translocations result in partial monosomy and partial trisomy of many genes, leading to intellectual disability and birth defects. We have analyzed chromosome breakpoints from 54 human subjects with unique unbalanced translocations using a combination of array comparative genome hybridization (CGH), targeted sequence capture, and next-generation sequencing. Included in this group of 54 rearrangements is a subset of translocations between three, four, or five nonhomologous chromosomes. Breakpoints with complex junctions involving multiple chromosomes may be indicative of chromosome shattering, a recently described process that plays a role in germline and somatic chromosome rearrangements. We fine-mapped breakpoint locations in 47/54 translocations with high-resolution array-CGH, and then captured breakpoint junction regions with SureSelect target enrichment and next-generation sequencing. Sequence analysis of breakpoint junctions allows inference of the type of DNA repair at translocation breakpoints and can reveal a more complex rearrangement structure than expected from array CGH alone. Consistent with recent publications that demonstrate the variability in breakpoint characteristics, we have sequenced four junctions that have either no microhomology, a few basepairs of microhomology, or insertions. Sequence analysis of the remaining translocation junctions is ongoing. Results from this large-scale study of patient breakpoint sequences will pave the way to more complete models for chromosome rearrangement by translocation.

939W

Analysis of the nucleotide sequence diversity within the *SUZ12* gene and its pseudogene *SUZ12P* as a means to investigate the signature of nonallelic homologous gene conversion. *T. Muβotter[1], J. Vogt[1], K. Bengesser[1], J. Högel[1], D.N. Cooper[2], H. Kehrer-Sawatzki[1].* 1) Institute of Human Genetics, University of Ulm, Ulm, Germany; 2) Institute of Medical Genetics, School of Medicine, Cardiff University, Cardiff, United Kingdom.

Double Holliday junctions (dHj) formed between paralogous sequences are recombination intermediates that are either resolved by crossover in a process termed *nonallelic homologous recombination* (NAHR) or by the non-crossover pathway associated with *nonallelic homologous gene conversion* (NAHGC). These pathways of resolution constitute alternative processes. However, since NAHR gives rise to putatively deleterious deletions and duplications, the products of the non-crossover pathway are observed more frequently. The molecular signature of meiotic NAHGC detected within the recombining paralogs is an increased SNP density as compared to the genomic average. Following this line of reason, several low-copy repeats (LCRs) exhibiting meiotic NAHR activity were found to be involved in frequent NAHGC-mediated sequence exchange as reflected by an increased SNP density. In addition, the LCRs located within the *NF1* gene region (termed NF1-REPs), which mediate meiotic NAHR causing type-1 *NF1* microdeletions, manifest an increased SNP frequency suggestive of frequent NAHGC resulting from meiotic nonallelic homologous recombination. Whereas nonallelic homologous recombination with or without crossover during meiosis has been studied in greater detail, the features of these processes during mitotic cell divisions are not well delineated. In this study, we have investigated whether NAHGC might also operate during mitotic nonallelic homologous recombination. To this end, we determined the SNP density within the paralogous *SUZ12* gene and *SUZ12P* pseudogene that undergo predominantly mitotic (postzygotic) NAHR giving rise to mosaic type-2 *NF1* microdeletions. We noted an increased SNP density (4.4 SNPs/kb) within the mitotic NAHR hotspot PRS4 recently identified in patients with type-2 *NF1* microdeletions as compared with the genome average of 1 SNP/kb ($p<0.0001$). The increased SNP frequency within the mitotic NAHR hotspot PRS4 is suggestive of sequence exchange by NAHGC operating between *SUZ12* and *SUZ12P*. Since PRS4 is a hotspot of mitotic (but not of meiotic) NAHR, we surmise that the increased SNP density is indicative of enhanced NAHGC activity operating during pre-meiotic mitoses. Our analysis suggests that not only meiotic but also mitotic dHj intermediates of nonallelic homologous recombination are resolved by non-crossover pathways leading to gene conversion in regions flanking the recombination-initiating DNA double-strand breaks.

940T

An Examination of the Origin of the Excess of Males Liveborn with Trisomy 21. *C. Walker[1], S. Gandy[1], J. Jones[1], A. Harden[1], T. Naluai-Cecchini[3], E. Cheng[3], S. Sherman[2], T. Oliver[1, 2].* 1) Department of Biology, Spelman College, Atlanta, GA; 2) Department of Human Genetics, Emory University, Atlanta, GA; 3) Department of Obstetrics and Gynecology, University of Washington, WA.

The best estimate for the sex ratio among live-born individuals with Down Syndrome is 1.15 whereas the sex ratio among normal individuals is 1.05, thus there is an excess of males with trisomy 21 that survive gestation. What could lead to an excess of males among live-born individuals with Trisomy 21 (T21)? Research shows there is an excess of live-born (LB) females with a heart defect among individuals with T21. Thus it is possible that the excess of LB males with T21 could be due to an excess of females being aborted with T21 and a heart defect. In order to explore this hypothesis, we examined the frequency of heart defects among aborted individuals with a prenatal diagnosis of T21. The presence or absence and type of heart defect were detected by prenatal ultrasound and our final population consisted of 245 cases. 122 heart defects were detected in our population. These included atrial ventricular septal defects (61/122), atrial septal defects (16/122), ventricular septal defects (44/122) and Tetralogy of Fallot (1/122). In our approach, logistic regression was used to determine if there was an excess of females with a heart defect among aborted fetuses with T21. Results from our analysis showed that among aborted individuals with T21, the odds of having a heart defect did not differ between males and females (OR: 1.50, $p = 0.14$), thus aborted females with T21 were not more likely to have a heart defect than aborted males. In addition, upon looking by type of heart defect, we did not find that the odds of having an atrial ventricular septal defect (OR: 0.16, $p=0.98$), an atrial septal defect (OR: 0.44, $p=0.36$), or a ventricular septal defect (OR: 0.18, $p=0.10$) differed between sexes. Thus, neither type of heart defect was more common in aborted females compared to aborted males. Thus our findings do not support the hypothesis that the excess of males LB with T21 could be due to an excess of females being aborted with T21 and a heart defect.

941W

Down - Turner mosaicism. *L. Martelli[1,2], C.G. Picanço[2], C.H.P. Grangeiro[2], C.S. Pereira[1], R.M. Scaparo[2], S.A. Santos[1], J. Huber[2], E.S. Ramos[1,2].* 1) University of Sao Paulo, Dept Genetics, Ribeirao Preto, SP, Brazil; 2) Medical Genetics Division HCRP, Ribeirao Preto, SP, Brazil.

The occurrence of double aneuploidy in the same patient, although rare, has been described in the literature. The most common involves trisomy 21 and monosomy X, with estimated prevalence of 1 to 2 million live births. We describe the presence of Down-Turner mosaicism, diagnosed in a patient with syndromic facies and developmental delay. The cytogenetic diagnosis by GTG banding after 100 metaphase analysis was 45, X [79] / 47, XX, + 21 [21]. The patient, 30yo, was born at 36 weeks of gestation, with 2150g and 47cm, lymphedema of hands and feet, and hypotonia with weak suckling. Hypothyroidism was diagnosed at 2yo. She presented secondary sexual characteristics and spontaneous menarche at 14 years old. Physical examination showed moderate intellectual deficit, brachycephaly, sparse hair, flat facial profile with hypoplasic midface, upward slanted palpebral fissures and ptosis, apparent macroglossia, dysplastic small ears, short neck, increased interval between 1 and 2 toes and ligamentous laxity. The double aneuploidy could be caused by non-disjunction pre-zygotic resulting in zygote 47, XX, +21, followed by delayed post-zygotic anaphase of the chromosomes X and 21, resulting in 45,X cell line. Harada et al (1998) suggested that the origin of mosaicism 45, X/47, XX, +21 could be the non-disjunction in maternal first meiotic division, followed by loss of chromosomes X and 21 derived from the mother during the first mitotic division. Supported by CNPq and FAEPA-HCRP.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

942T

Accurate, Precise, and Tunable Cytogenetics with Next-Generation Sequencing. *M.A. Eberle, T. Royce, F. Kaper, J. Cottrell, J-B. Fan.* Dept Bioinformatics, Illumina, Inc, San Diego, CA.

Technological advances in whole-genome shotgun sequencing have led to the successful diagnoses of once-enigmatic illnesses resulting from rare, de novo point mutations. Application of this same technology to the discovery of cytogenetic abnormalities holds similar promise, with the full range of capabilities yet to be completely described and quantified. To explore the effectiveness of sequencing for cytogenetics applications, we analyzed sequence data from several cytogenetically interesting samples with documented chromosomal abnormalities each paired with a cytogenetically normal cell line. Each sample was sequenced and we informatically down sampled the data to simulate various amounts of total sequencing. We used this down sampled data to simulate copy number gains and losses of varying size. To explore the limit of resolution for detecting copy number gains and losses, we informatically combined sequence data from different samples to simulate losses and gains of varying sizes. This method of creating simulated copy number events allows us to generate a large number of events of relatively small size to accurately assess the sensitivity of detecting these features with sequence data. After sampling the sequence data down to the throughput of the MiSeq personal sequencer, we find that the technique has the ability to detect 95% of copy number losses of 200 kb and 95% of copy number gains of 400 kb. Additionally, at these limits of resolution, we find virtually no false positive calls. The technique is extendable to genetically mosaic events and uniparental disomy, albeit at a higher limit of resolution. Increasing the total amount of reads improved all capabilities. With appropriate study design, the same shotgun sequencing technologies routinely used to identify simple point mutations can be employed for sensitive and precise detection of cytogenetic abnormalities. At current sequencing throughputs, measures of sensitivity and specificity are on par with array technologies. As shotgun sequencing throughputs improve, sequencing will become an increasingly compelling alternative to existing cytogenetics technologies.

943W

Two new cases of interstitial 6q deletion, associated to microcephaly, corpus callosum hypoplasia and further brain and cerebellar malformations. *V. Parisi[1], S. Loddo[1], L. Travaglini[2], A. Ferraris[1], G. Vitiello[3], G. Zanni[2], L. Bernardini[1], M.L. Di Sabato[4], M.C. Digilio[2], A. Novelli[1], E. Del Giudice[5], A. Rossi[6], E. Bertini[2], B. Dallapiccola[2], E.M. Valente[1].* 1) Cytogenetics Lab, CSS-Mendel Inst, Rome, Italy; 2) Bambino Gesù Children Hospital Rome-Italy; 3) Department of Biopathology and Diagnostic Imaging, Tor Vergata University, Rome, Italy; 4) Sapienza University of Rome Department of Child Neurology and Psychiatry Rome-Italy; 5) Department of Paediatrics, “Federico II” University, Naples, Italy; 6) Departments of Pediatric Neuroradiology, IRCCS G. Gaslini, University of Genoa, Genoa, Italy.

Rare distal 6q deletions show a broad phenotypic spectrum depending on the size and position of deleted segments. The chromosome region 6q25.2q25.3 has been described as critical for microcephaly, growth delay, hearing loss and corpus callosum agenesis/hypoplasia (ACC). In this region, two genes (TIAM2 and SYNJ2) have been indicated as candidate for microcephaly and ACC. We report on two new distal 6q deletions, detected by microarray platforms: patient 1 showed a deletion, extended about 13 Mb, at 6q25.1q27 (4×44K; Agilent); in patients 2 a deletion was disclosed at 6q25.3q27 (~11.3 Mb) (Human GeneChip 6.0; Affymetrix). The two patients showed brain malformations, in particular ACC, and shared other clinical features, such as psychomotor delay, facial dysmorphisms, microcephaly. Haploinsufficiency of TIAM2 and SYNJ2 could be the direct cause of ACC in patient 1, as the deleted region included both genes. In patient 2 the deletion was more distal and did not encompass these genes, although preliminary analysis indicated a reduced expression of TIAM2, suggesting the presence of a regulatory element in the hemizygous distal 6q25.3q27 region. Our data confirm and expand neuroradiological phenotype associated to distal 6q deletion and suggest the hypothesis that the reduced expression of TIAM2 gene, due to deletion or other mechanism, is causative of ACC.

944T

Ring chromosome 13: an eventual risk for intellectual disability, surdity, congenital malformations and leukemia. *I. BEN-ABDALLAH-BOUHJAR[1,2], H. HANNACHI[1,2], S. MOUGOU-ZERELLI[1,2], H. BEN-KHELIFA[1,2], A. LABALME[3], D. SANLAVILLE[3], H. ELGHEZAL[1,2], A. SAAD[1,2].* 1) Cytogenetics and Reproductive Biology Department, Farhat Hached University Teaching Hospital,, SOUSSE, sousse, Tunisia; 2) Cytogenetics and Reproductive Biology Common Service Units for Research in Genetics, Faculty of Medicine of Sousse, Avenue Mohamed Karoui, University of Sousse, Tunisia; 3) Hospices Civils de Lyon, Service de Cytogénétique Constitutionnelle, Lyon, France.

The 13q deletions syndrome are currently known to share several common semiologic features, including, in particular intellectual disability ranging from moderate to severe, marked short stature, neural tube, brain and heart defects, microcephaly, genital malformations in males including undescended testes and hypospadias and characteristic facies. However, the hearing and speech delay in our cases are uncommonly reported in either the deletions of chromosome 13, so hearing and speech impairments are suggested as another new clinical feature to 13q deletion Syndrome, in this study we report three cases of ring chromosome 13 and we compare the growth and clinical features of these patients with previously reported cases, with a similar deletion on the long arm of chromosome 13. The patient karyotypes were 46,XY,r(13)(p11;q34) dn for the patient 1, 46,XX,r(13)(p11;q34) dn for the patient 2 and 46,XY,r(13)(p11:q14) dn for the patient 3, as a result of the deletion in the telomeric regions of chromosome 13. They were, therefore, monosomic for the segment 13q34◊13qter, in addition for the patient 3 the deletion was more large encompassing the segment 13q14◊13qter. Fluorescence in situ hybridization analysis showed loss of a specific subtelomeric 13q region in r(13) in three cases and loss of a specific locus 13q14/RB1 in the patient 3, who died from Multiple Myeloma during the cytogenetic analysis, FISH on fixed Bone Marrow cells of the patient 3 diagnosed with MM, using the probe RB1gene showed deletion in 13q14.2. Array CGH exploration of DNA's of patient 2 revealed a loss of at least 2.9 Mb on the short arm of chromosome 13 and a deletion of 4.7 Mb on the long arm of the same chromosome 13. Initially FISH defined, the chromosome breakpoints and were precisely localized by array CGH on 13q11 [arr 13q11q11 (18601703- 21593561)×1 dn] instead of 13p11, and 13q34 [arr 13q34 (109599699- 113964366)×1 dn]. In our patients the deletion included the gene EFNB2, GJB6, ARHGEF and MYO16 are probably involved respectively in genital development, hearing loss, microcephaly and growth retardation. To the best of our knowledge, our patient 3 with Multiple Myeloma is the first report in which constitutional deletion of RB1 gene is also detected in Bone Marrow cells. We conclude that the Ring chromosome 13 is an eventual risk for intellectual disability, surdity, congenital malformations and leukemia.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

945W

Use of Oligo-SNP array for the detection of abnormalities in CLL. *N.C. Christacos[1], M.L. Slovak[1], M.A. Sanidad[1], Y. Hsu[1], J.C. Kelly[1], P.N. Mowrey[1], D.M. Jones[2]*. 1) Dept. Cytogenetics, Quest Diagnostics, Chantilly, VA; 2) Dept. Pathology, Quest Diagnostics, Chantilly, VA.

B-cell chronic lymphocytic leukemia (CLL) is a hematologic malignancy characterized by the clonal expansion and significant accumulation of immunophenotypically similar (usually co-expressing CD5 and CD23) mature B-lymphocytes. Conventional cytogenetics and fluorescence in situ hybridization (FISH) studies are used worldwide to subdivide CLL into different pathogenetic and prognostic subgroups. Newer molecular cytogenetic studies using microarray technology promise to expand our knowledge of the genetic aberrations in CLL. To address this supposition, we used the high resolution genome-wide Affymetrix Cytoscan HD array which combines oligonucleotides and single nucleotides polymorphisms (SNP) (OSA) as a tool to detect both copy number changes (CNAs) and acquired isodisomy (aka, uniparental disomy). We performed OSAs on samples from 21 patients with a clinical indication of CLL and compared the microarray results with either the conventional cytogenetics and/or FISH analyses. All samples had karyotypic or FISH detectable abnormalities. OSA showed a total of 127 CNAs (average CNAs per sample,6; range, 2–22 CNAs). Deletion 13q was missed in two patients with 6% and 15% mosaicism and one patient with deletion 11q in two mitotic cells after culturing in CpG-oligonucleotide DSP30 plus IL-2 for 72 hours; however, the other two clonal aberrations in that sample were detected. Trisomy 12 and TP53 deletions showed 100% OSA/FISH concordance. Overall, clonal mosaicism detected by OSA correlated well with the percentage of interphase cells showing the FISH abnormalities. Seven samples (33%) showed acquired isodisomy >5 Mb (range, 5.4–36 Mb). Two samples showed non-CLL aberrations consistent with either t-AML or hairy cell leukemia, indicating the ability to obtain valuable genetic information for those samples that would be missed using the targeted CLL FISH probe panel. Furthermore, at relapse and in secondary leukemia, the karyotype often shows a high degree of genetic complexity with multiple tumor subclones. Such clonal diversity reflects the underlying biologically selected alterations that can, if carefully mapped, highlight the locations of oncogenes and tumor suppressors underlying disease progression. Genetic aberrations including acquired isodisomy are common in CLL and are easily detected by OSA. This relatively new technology may be able to reliably support or succeed current cytogenetic methods.

946T

Array Based Comparative Genomic Hybridization Identifies Unique Copy Number Alterations in Diffuse Large B-cell Lymphoma Subtypes: A Review of Published Microarray Data. *R. Garcia*. The UT Southwestern Medical Center at Dallas-Department of Pathology, Clinical Cytogenetics.

Diffuse large B-cell lymphoma (DLBCL) accounts for more than 30% of adult non-Hodgkin lymphoma cases and it is characterized by a heterogeneous set of lymphomas. Many conventional and molecular cytogenetic studies along with morphological, biological, clinical and microarray studies have made an effort to better classify these sets of lymphomas. However, in spite of these, many remained unclassifiable. In fact, most of these lymphomas remain undistinguishable and fall into the not otherwise specified (DLBCL-NOS) category. The goal of this current study was to review the literature for currently published microarray data; particularly array based comparative genomic hybridization (aCGH), classical CGH and gene expression profiles for DLBCL. Copy Number Alterations (CNA) with greater than 20% on any given subset were catalogued in a table form. Based on this, two separate chromosome ideograms were generated depicting similarities and differences between subsets. Consequently, this review identified unique CNA specific for DLBCL subtypes. These included: primary mediastinal large B-cell lymphoma (PMLBCL), gain at 2p15, 9p24, 9q34, Xp11.4–21, Xq24–26; LBCL of the bone, gain at 2p16, 6p21 and loss at 15q15-q26; DLBCL of the CNS, loss at 6p21–25,17p12–13; Leg type DLBCL, gain at 2p22, 2p12, 3p21–25, 3q28–29, 9p12–21,16q23,22q11and loss at 4q, 8p11,14q, Xq13–25; CD5+ DLBCL, gain at 13q21–34; Plasmablastic lymphoma, gain at 10p12, 14q32, 16q24,17p12–13 and primary effusion lymphoma with gains at 4q28–35, 8q11.2–23.1, 11p, 17q23–24, 19p13 and loss at 11q24–25. However, a number of unclassifiable DLBCL subtypes post-microarray studies remains. These include intravascular LBCL, Epstein-Barr virus DLBCL of the elderly, LBCL arising in HHV8+ multicentric Castleman disease, Lymphomatoid granulomatosis, LBCL with features intermediate between DLBCL and Hodgkin lymphoma and pyothorax associated lymphoma. In brief, although this review re-classified some of the DLBCL subsets by distinct CNA, better classification models are clearly needed for DLBCL. Keywords: DLBCL, aCGH, Copy number alterations.

947W

Subtype of ETV6/ABL1 fusion transcript is associated with different leukemia phenotype. *A. Kwon[2], J. Park[1,2], J. Lim[1,2], Y. Kim[1,2], K. Han[1,2], S. Lee[3], H. Kim[3], W. Min[3], J. Kim[2], M. Kim[1,2]*. 1) Department of Laboratory Medicine, The Catholic University of Korea, Seoul, South Korea; 2) Catholic Genetic Laboratory Center, The Catholic University of Korea, Seoul, South Korea; 3) Division of Hematology, Department of Internal Medicine, Catholic Blood and Marrow Transplantation Center, The Catholic University of Korea, Seoul, South Korea.

The ETV6/ABL1 fusion transcript is thought to be a very rare aberration in hematopoietic malignancies and only 22 cases were published; 9 atypical chronic myeloid leukemia, BCR/ABL1 negative (aCML), 3 myeloproliferative neoplasm (MPN), 4 acute myeloid leukemia (AML), and 6 acute lymphoblastic leukemia (ALL). There are two types of the fusion transcript at the mRNA level; one transcript, type A composed of the first 4 exons of ETV6 fused to the second exon of ABL1, and the other, type B where the ETV6 exon 5 fuses to the ABL1 exon 2. The biological consequences of ETV6/ABL1 show a significant similarity with BCR/ABL1. Herein, we described two new cases of acute leukemia with the ETV6/ABL1 fusion, one with AML with eosinophilia (Case 1) and the other with B cell ALL (Case 2), which was screened by multiplex RT-PCR. Cytogenetics of case 1 showed a normal karyotype, but RT-PCR revealed positive to ETV6/ABL1 type B fusion transcript. In case 2, karyotying represented 46,XX,del(9)(p22),der(10)t(9;10)(q22;p15)[12]/46,XX[8], but RT-PCR revealed both type A (595 bp) and type B (1141 bp) fusion transcripts. Fluorescence in site hybridization using mixture of BCR/ABL1 and ETV6/RUNX1 probes revealed that the ETV6/ABL1 fusion were located on 12p13 which meant the crucial event resulted by the insertion of a part of C-terminus region from ABL1 gene into the intron 1 of ETV6 gene. This fusion signals were also detected in eosinophils of case 1, so that we considered the clonal eosinophilia as the characteristics of AML with ETV6/ABL1. This eosinophilia has been known to accompany with MPN and AML but not always with ALL. A thorough review of all published cases about ETV6/ABL1, including our new cases, all ALL patient possessed type A fusion transcript. Otherwise, type B fusion transcript accompanied with MPN, aCML, and AML. These findings are very similar that BCR/ABL1 fusion transcript with minor BCR breakpoint is associated with ALL phenotype. Reported patients including present study, 6 out of 7 ALL, 4 out of 5 AML were died. Therefore, ETV6/ABL1 is probably a poor prognostic factor in acute leukemia. Transient response to inhibit ABL1 kinase in such patients suggested that those drugs maybe in ETV6/ABL1 positive acute leukemia therapeutically useful but not sufficient to induce remission. It is helpful to develop a modified treatment modality including ABL1 kinase inhibitor.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

948T

Detecting Chromosomal Inversions Using Chromatid Painting Strategies- Applications. *F.A. Ray[1,2], E. Zimmerman[2], M.N. Cornforth[2,3], E.H. Goodwin[2], J.S. Bedford[1,2], S.M. Bailey[1,2]*. 1) Env Rad Hlth Sci, Colorado State Univ, Fort Collins,CO 80523-1618; 2) Kromatid Inc., 515 East Laurel St., Fort Collins, CO 80524-3151; 3) 3Department of Radiation Oncology, University of Texas Medical Branch, Galveston, TX 77555-0884, USA.

Chromosomal inversions are symmetrical exchanges resulting from two DNA double strand breaks within a single chromosome that are incorrectly rejoined in an inverted orientation. Using banding methodologies, large inversions, (i.e. >10 Mb) are often difficult to detect and inversions smaller than a typical metaphase chromosome band are rarely detected. We developed a new methodology to detect chromosomal inversions with a 10 fold improvement in resolution over existing techniques. Chromatid painting, like chromosome painting requires chromosome-specific fluorescent probes, the difference being that all of the probes in a chromatid paint are designed to bind to only one of the two 5' to 3' DNA strands. Newly-replicated (nascent) strands are removed from metaphase chromosomes and genomics is used to develop probes directed to unique targets on one chromatid and not the other. The first chromatid paint was developed to human chromosome 3. The chromatid paint behaves as envisioned, i.e. chromosomes have one chromatid per targeted chromosome with fluorescent signal and the other without, inversions appear as a fluorescent signal switch to the corresponding position on the other chromatid Radiation of cells prior to chromatid painting resulted in a dose dependent increase in the inversion frequency. Dose responses using our chromatid 3 paint following γ-ray or HZE ion exposures were determined. Comparison of inversion vs. translocation frequencies involving chromosome 3, suggests that inversions are approximately 6 times more frequent than translocations. Intriguingly while preliminary, the slope of the dose response following HZE irradiation was much steeper than that for the γ-rays suggesting the possibility of using radiation-induced chromosome inversions to discriminate between and high and low LET radiations. We also developed an assay for for a thyroid cancer-specific, radiation-induced inversion based on chromatid painting. Using this assay, the PTC 1 inversion was detected at low frequency in human thyroid cells immediately after irradiation, and at an increased frequency after these irradiated cells were cultured for additional population doublings. Detection of the specific inversion after radiation demonstrates a new application for the assay and the increased inversion frequency after multiple cell divisions, indicates that this specific inversion was selected for in the population of cells. Other cytogenetic applications will be discussed.

949W

Efficacy and Implementation of a SNP Microarray for the Evaluation of Patients with Multiple Myeloma. *S. Schwartz, R.D. Burnside, J. McElligott, I. Gadi, J. Kesler, V. Jaswaney, K. Phillips, J. Tepperberg, B. Williford, P. Papenhausen.* Lab Corp of America, Research Triangle, NC.

The laboratory confirmation of multiple myeloma (MM) is problematic since chromosome analysis (CA) and FISH analysis are often normal due to indolent, low percentage disease. The utilization of magnetic CD138 cell sorting increases the frequency of abnormalities by FISH, but fails to detect the clone for many of these patients. Microarray analysis is common in constitutional studies, but for neoplasia in general it is still limited. In order to determine the efficacy of this technology in MM patients, a novel approach was undertaken involving 18 patients. All patient marrow aspirates were studied initially with CA and/or FISH after CD138 cell sort enrichment. This was followed by SNP microarray analysis utilizing the Cytoscan HD platform on both the CD138 positive and negative fractions. This approach revealed that 82% of patients with a normal karyotype had a detectable abnormality by the array analysis. It also showed that in CA abnormal patients, numerous additional abnormalities could be detected by the array analysis and that many of these had diagnostic implications. The results from these studies show the importance and efficacy of utilizing microarrays in the study of MM. The important findings from this work include:(1)The superior detection of abnormalities by microarray in 82% of cases that were initially diagnosed as normal by CA and in 50% of cases that were normal by FISH;(2)In cases determined to be abnormal by CA and/or FISH, additional abnormalities were detected by the microarray analysis in all of patients; (3)Prognostic/diagnostic information was obtained in over 93% of the patients studied; (4)Homozygosity associated with acquired uniparental disomy(aUPD) was detected in 36% of the patients with abnormal findings. Comparison of the positive and negative CD138 fractions provided clear resolution of constitutional homozygosity from acquired UPD; (5)Chromothripsis was identified in 21% of patients with abnormal findings. Specific allele ratios in the array permitted delineation of chromosomal structure and mechanism of formation of a number of complex abnormalities; (6)Homozygous deletions of tumor suppressor genes were identified in 22% of patients with abnormal findings; (7)Different clonal abnormalities could be found in the positive and negative CD138 fractions in two cases, indicating at least one where the patient had both MM and MDS; (8)These studies provide a new unique paradigm for the evaluation of patients with suspected MM.

950T

A de novo intragenic deletion of *AUST2* in a patient with autism spectrum disorder. *S. Kantarci[1], I.E. Amarillo[1], J. David[1], E. Vilain[2], X. Li[1]*. 1) Pathology and Laboratory Medicine, David Geffen School of Medicine at UCLA, Los Angeles, CA; 2) Pediatrics, Medical Genetics, David Geffen School of Medicine at UCLA, Los Angeles, CA.

Autism spectrum disorder (ASD) is a clinically heterogeneous developmental disorder with a strong genetic component. The prevalence of ASD is estimated to be 1 of 85 children and five times as many boys as girls. Although ASD is known to be highly heritable (about 90%), the genetics underlying ASD is complex and remains largely unknown. About 5–15% of ASD cases have an identifiable genetic etiology corresponding to known single-gene disorders and various chromosomal abnormalities. Microarray studies have revealed that rare de novo or transmitted copy number variations (CNVs) appear to contribute significantly to the pathogenesis of this complex disease. De novo CNV events have been observed in about 10% of ASD cases. We report a 4.5 year-old female with ASD who has been found to have a de novo copy number loss on chromosome 7q11.22. The proband was born by normal vaginal delivery at term following an uncomplicated pregnancy. She was noted to drool constantly at the age of 6 months. She was then noted to have developmental delays and started speech therapy at 18 months of age. She presented tonic-clonic movements, lasting for up to a minute and followed by a period of decreased alertness at 3 years old. She does not have dysmorphic features and her CT scan and neurological exams are normal. Chromosomal microarray analysis (CMA) using Affymetrix SNP array 6.0 revealed a potentially pathogenic de novo 62 Kb microdeletion on chromosome 7q11.22. This interval overlaps with a single exon (exon 6) of *AUTS2* (autism susceptibility 2 gene) and has a 5% overlap with benign CNVs. Affymetrix CytoScan HD array confirmed the deletion in proband. *AUTS2* consists of 19 exons and spans 1.2 Mb on the genomic DNA. Although the function of the protein encoded by *AUTS2* is still unknown in humans, it is highly expressed in the mice developing cerebral cortex and cerebellum. Disruptions of the *AUTS2* gene by de novo chromosomal rearrangements have been previously reported in seven cases with ASD and/or intellectual disability. In addition, a whole genome array comparative genomic hybridization study in a cohort of individuals with mixed types of idiopathic epilepsy identified rare overlapping deletions within *AUTS2* intronic region in two cases. To our knowledge, this study identified the first de novo intragenic deletion of *AUST2* in ASD. This study not only supports the association of *AUTS2* with autism but also highlights the significant role of CMA in genetic evaluation of ASD.

951W

22 q 11.2 syndrome in a group of patients who attended Operation Smile Foundation in Colombia, between 2005 and 2012. *I. Briceno[1,2], JC. Martinez[1], A. Venegas[2], O. Moreno[2], M. Sarmiento[2], S. Bohorquez[1], M. Montiel[1], A. patino[1]*. 1) Universidad de La Sabana, Chia, Colombia; 2) Universidad Javeriana. Bogota Colombia.

Velocardiofacial syndrome (VCFS) is a genetic condition characterized by abnormal pharyngeal arch development that results in defective development of the parathyroid glands, thymus, and conotruncal region of the heart. More than 180 different clinical features are associated with velocardiofacial syndrome, with no single anomaly present in every patient. Affected individuals may present with structural or functional palatal abnormalities, cardiac defects, unique facial characteristics, hypernasal speech, hypotonia, and defective thymic development. An estimated 75% of patients with velocardiofacial syndrome have cardiac anomalies. Palatal abnormalities predispose to speech and feeding difficulties. The defective thymic development is associated with impaired immune function. In addition, affected individuals may present with learning disabilities, overt developmental delay, psychiatric disorders, and renal and musculoskeletal defects. About 10% of patients with velocardiofacial syndrome have DiGeorge syndrome, as many as 15–20% of patients have Pierre Robin syndrome. Velocardiofacial syndrome is a specific syndrome that includes as part of its phenotypic spectrum the DiGeorge sequence, the Pierre Robin sequence, and disorders associated with CHARGE syndrome. Is considered Genetically, 90 percent of patients with the features of this syndrome are missing a small part of their chromosome 22 at the q11 region. This region encompasses about 30 individual genes and results in developmental defects in specific structures throughout the body. It is not known why this region of chromosome 22 is prone to become deleted, but this is one of the most frequent chromosome defects in newborns. Deletion 22q11.2 is estimated to occur in one in 3,000 to 4,000 live births. Most of the 22q11.2 deletion cases are new occurrences or sporadic (occurs by chance). However, in about 10 percent of families, the deletion is inherited and other family members are affected or at risk for passing this deletion to their children. En este trabajo nosotros reportamos el study among 2000 patients attending “Operation Smile” in Bogota, 2 cases of 22 q 11.2 syndrome were diagnosed, confirmed by cytogenetics, in situ fluorescence hybridization showed the 22q11 microdeletion. which corresponds to 0.1 of the population studied. This result indicates a frequency significantly higher than that reported in the literature for general population. The Clinical description is presented.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

952T

Incidence of mosaicism detected by exon-targeted, high-resolution array comparative genomic hybridization (aCGH) in 10,362 consecutive cases with special emphasis on detection of complex unbalanced structural rearrangements. *W. Bi[1], J. Pham[1], C. Shaw[1], P. Hixson[1], A. Ester[1], A. Pursley[1], K. Plunkett[1], P.L. Magoulas[1], S.-H.L. Kang[1,2], S.R. Lalani[1], C. Bacino[1], P. Stankiewicz[1], A. Patel[1], S.W. Cheung[1].* 1) Medical Genetics Laboratories, Department of Molecular and Human Genetics, Baylor College of Medicine, Houston TX, 77030; 2) Allina Medical Laboratories, Minneapolis, MN, 55440.

Chromosomal mosaicism, the presence of two or more different cell lines in an individual, is clinically important as it has been established to cause miscarriage, birth defects, developmental delay, or cancer and to contribute to phenotypic variation. However, somatic mosaicism is usually under-diagnosed due to various factors, which include subtle phenotypic abnormalities, technical limitations, and inherent tissue-specific mosaicism. Here, we studied 10,362 patients, most of whom had a clinical indication of cognitive impairment, using a custom-designed, exon-targeted, whole-genome oligonucleotide array. Somatic mosaicism ranging from 6.5% to 92% was detected in 59 cases. Mosaicism for numerical abnormalities including ring chromosomes was detected in 35 cases. In the remaining 24 cases with unbalanced structural rearrangements, the following abnormal cell lines were detected: [1] simple terminal or interstitial deletions/duplications (11 cases), [2] complex structural abnormalities, involving two or more copy-number changes (5 cases), [3] derivative chromosomes (3 cases), and [4] mosaic copy-number losses encompassing or disrupting a single gene (5 cases). For complete characterization, it was necessary to perform additional studies such as FISH analysis, chromosome analysis, and/or using other molecular methods. We present here two cases in which patients had two abnormal cell lines, one of which was beyond the resolution of chromosome analysis and thus, only detectable by array CGH. Patient 1 is a 1- year-old girl with global developmental delay, seizures, and short stature, who has mosaicism for a cell line with a der(17), resulting from an unbalanced translocation between 5p and 17p in 25% of cells whereas the remaining cells had an isolated 17p13.3 deletion. Patient 2 is an 8-year-old boy with developmental delay and seizure disorder. Array studies showed mosaicism for a copy-number loss at chromosome 17p13.3p13.1, suggesting two different-sized deletions of this region. FISH analysis confirmed a smaller terminal deletion at band 17p13.3 in 77.5% of cells whereas a larger terminal deletion spanning to band 17p13.1 was found in 22.5% of cells. In summary, increasing array resolution indicates that chromosomal mosaicism may be more frequent than previously thought and particularly that patients with mosaic chromosomal rearrangements may have an additional abnormal cell line that is only detectable by aCGH.

953W

A Cryptic Derivative Chromosome 12 Detected by Oligo-SNP Array. *Z. Dai, N.C. Christacos, S. Schonberg, J. Kelly, E. Wallenhorst, K. Sullivan, C. Adams, Y. Hsu, H. Walker, T. Simanivanh, P. Mowrey.* Cytogenetics, Quest Diagnostics, Chantilly, VA.

We report the cytogenetic and microarray results of a 1 year old male with failure to thrive. Conventional cytogenetic analysis failed to identify any apparent structural abnormality or rearrangement. Concurrent high resolution microarray analysis using an oligo-SNP array (Afffymetrix Cytoscan HD) indicated a derivative chromosome 12 resulting from an unbalanced translocation between chromosomes 8 and 12. Two variants of unknown significance were also detected. With respect to the identified imbalance, an approximately 3.2 MB terminal loss on the distal short arm of chromosome 12 at p13.32p13.33 was detected. Additionally, an approximately 7.0 Mb gain of terminal chromosome 8 long arm material at 8q24.23q24.3 was detected. Together, these findings are consistent with a derivative chromosome 12 from an unbalanced t(8;12)(q24.23;p13.32) which was confirmed by subtelomere FISH studies using probes for the subtelomeric regions of 8q and 12p. The rearrangement on the short arm of chromosome 12 (net gain of 3.8 Mb of 8q material) was “cryptic” by conventional cytogenetic G-band analysis. Multiple genes are contained in the regions partially monosomic on 12p and partially trisomic on 8q; therefore, this finding is considered to be causative in the abnormal phenotype seen in this child. Both of these rearrangements are rare and as such, a clearly recognized syndrome for either rearrangement has not emerged. Reported patients with 12 p deletion have in common a mild phenotype with moderate mental retardation and severe behavioral problems. The few reported cases with duplication of 8q had in common severe mental retardation, growth delay and epilepsy. This case illustrates how oligo-SNP array testing is crucial to identifying “cryptic” rearrangements in apparently karyotypically normal patients with multiple congenital anomalies and/or developmental delay.

954T

GLUT1 Deficiency and Chromosome Microarray Analysis. *A.J. Dawson[1,2,3], A. Mhanni[2,3], F. Booth[4], D. Bernier[1], M. Tomiuk[1], J. Hartley[3,5], M. Strecker[6], K. Hovanes[6].* 1) Dept Cytogenetics, Diagnostic Services Manitoba/Univ Manitoba, Winnipeg, MB, Canada; 2) Section of Genetics and Metabolism, Department of Pediatrics and Child Health, Health Sciences Centre, Winnipeg, Manitoba, Canada; 3) Department of Biochemistry and Medical Genetics, University of Manitoba, Winnipeg, Manitoba, Canada; 4) Section of Pediatric Neurology, Department of Pediatrics and Child Health, Health Sciences Centre, Winnipeg, Manitoba, Canada; 5) Genetics and Metabolism Program, Winnipeg Regional Health Authority, Winnipeg, Manitoba, Canada; 6) Combimatrix Diagnostics, Irvine, California, USA.

The patient, an 8 month old boy, presented with recurrent seizures with his first event of tonic, clonic afebrile seizure occurring at 5 months of age. He was the first born to a non-consanguineous healthy 35 year old mother and a 41 year old father. The prenatal and perinatal histories were unremarkable. An amniocentesis had been performed because of advanced maternal age. A de novo, apparently balanced, reciprocal translocation was detected: 46,XY,t(1;20)(p34.1;q13.1)dn. The child's seizures were difficult to control despite the use of multiple anti-epileptic medications. The child's physical examination was remarkable for microcephaly with an OFC below the 5th centile, with weight and height at the 25th centile. No other abnormalities were noted. An extensive biochemical work up on plasma and urine was undertaken with no specific etiology identified. The findings on neuroimaging, including an MRS, were nonspecific. An extensive metabolic work-up revealed a ratio of CSF to blood glucose of 0.378 (normal ratio: 0.65 +/ − 0.01), suggesting a biochemical diagnosis of glucose transporter type1 (GLUT1) deficiency (De Vivo disease). The results of chromosome microarray analysis, performed at the same time, showed a single copy number decrease of 1.2 Mb of the 1p34.2-p34.1 region: arr 1p34.2p34.1(42,871,391–44,074,069)x1. The genes involved in the loss of 1p34.2-p34.1 include SLC2A1. Mutations and deletions of the SCL2A1 gene are responsible for GLUT1 deficiency, an autosomal dominant disorder. Parental chromosomal microarray analyses were performed and did not detect the deletion. This and FISH results confirmed that the deletion was due to the translocation and was de novo. This deletion is likely the result of an unbalanced translocation between chromosomes 1p and 20q. The translocation breakpoint on 20q could not be determined by this microarray since it did not involve any copy number change. The patient has been on a ketogenic diet for four months. He has generally been well and has been making considerable developmental progress, with seizures remaining under control. This case clearly demonstrates the value of genomic arrays in the investigation of children with developmental delays by increased resolution of cytogenetic rearrangements, which leads to correlation of clinical manifestations with specific changes in gene copy number. Specifically, in our patient, chromosome microarray analysis detected a condition that was treatable.

955W

Clinically relevant mosaic findings in a total of 9,605 patients and parents by using genome wide high resolution SNP array analysis in constitutional diagnosis. *N. de Leeuw, J.Y. Hehir-Kwa, B.H.W. Faas, T.K. Rinne, M.J.E. Kempers, S.A. de Munnik, N.F.A. Leijsten, G.C. Machielsen, S.L.J. van Gessel, M. Wunderink, M.J.G. Banning, R. van Beek, M. del Rosario, B.B.A. de Vries, D.F.C.M. Smeets, R. Pfundt.* Dept Human Genetics, Radboud University Nijmegen MC, Nijmegen, Netherlands.

We routinely perform genome wide SNP array analysis as the first-line diagnostic test for patients with intellectual disability and/or congenital anomalies and in prenatal diagnosis in case of structural ultrasound anomalies.

In 28.8% of patients a potentially causative, rare copy number variant (CNV) was detected. We categorized 6.5% as de novo, 9.1% inherited, 0.8% X-linked, and in 12.4% the inheritance patterns still needs to be determined. We observed a significantly increased percentage of homozygosity in an additional 6.1% of our patients. These homozygous regions subsequently led to the identification of pathogenic mutations in recessive disease genes, uniparental disomies, or low-mosaic aneuploidies in more than 30 patients. In November 2011, we switched from the Affymetrix 250K SNP array to the CytoScan HD array platform (Affymetrix, Inc., Santa Clara, CA, USA) which further enhanced the CNV detection resolution, improved the detection of mosaic imbalances and also enabled us to detect clinically relevant, mosaic, copy neutral changes of homozygosity in three patients. A mosaic finding (CNV, aneuploidy or allelic imbalance) was detected in 24 of 7,294 patients, two of 247 unborn babies, and in eight of 2,064 parents.

Mosaic abnormalities appear to be relatively common not only in patients (1 in 304) and unborn babies with structural ultrasound anomalies (1 in 124), but in parents (1 in 258) as well. It is crucial to determine whether an apparently de novo imbalance in an affected child, is in fact due to a mosaic aberration in the (unaffected) parent, highlighting an (often dramatically) increased recurrence risk. The percentage of mosaicism often differed between tissues samples of mesodermal or ectodermal origin from each of these individuals. In two patients such tissue-dependent differences were shown to change over time, most likely due to a rescue mechanism.

Genome wide SNP array analysis is a suitable and highly effective method to apply in genome diagnostics for the detection of CNVs and homozygosity pre-screening in one single test.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

956T

Genomic differences between spontaneously aborted fetuses and live-born children with 45,X monosomy. *J. Domínguez Ortiz[1], L. Torres[1], S. Sanchez[1], V. del Castillo[1], L. Orosco[2], A. Carnevale[2], P. Grether[3], D. Mayen[4], S. Frias[1,5].* 1) Genetica Humana, Instituto Nacional de Pediatría, Ciudad de México, Mexico; 2) Instituto Nacional de Medicina Genomica, Ciudad de México, Mexico; 3) Instituto Nacional de Perinatologia, Ciudad de México, Mexico; 4) Diagen, Ciudad de México, UNAM; 5) Instituto de Investigaciones Biomedicas, Ciudad de México, UNAM.

It has been described that for all the pregnancies with monosomy 45,X only 1% survives. To find some explanation we analyzed the genomic differences between live-born and abortions. We studied 11 patients and 3 abortions with 45,X karyotype performed conventional cytogenetic. gDNA was extracted from peripheral blood or from tissues of the miscarriages. For the precise characterization of the genomic differences we performed individual arrays for 11 live born cases and 3 spontaneous abortions using Affymetrix Genome Wide Human SNP 6.0 arrays. We identified some common regions with loss or gain of genomic material; each region was analyzed using UCSC genome browser, Database of Genomic Variants, DECIPHER, NCBI, Gene Ontology, miRBASE and miRANDA. The genomic differences observed in the SNP 6.0 array were the following: in Live-Born the lost of the next regions: 1q31.2 (Promoter of RGS18), 6q23.3 (KIAA1244), 8p23.1 (XKR6), 15q24.3 (SCAPER), 21q22.13 (DYRK1A, KCNJ6, DSCR4, DSCR8). In Abortions we saw gain of 17q12.2 (CNV), and lost of 2q32.1 (CNV), 3q26.1 (CNV), 4q32.2 (FSTL5, CNV), 7q22.1 (CUX1), 13q14.13 (LRCH1, HTR, SUCLA2, MED4), 14q21.1 (CNV), 19p13.2 (INSR, TRAPPC5, CD209, CLEC4G, RETN, STXBP2, ZNF358, PNPLA6), 22q13.1 (APOBEC3A). We found one coincident region lost in miscarriages as compared with live-born genomes, 7q22.1 which includes CUX1 gene, involved in the transcriptional regulation of several genes implicated in embryonic development, the lost of this region may cause an imbalance in the fetus that could originated a transcriptional de-regulation during development triggering a development failure. The most evident difference between the groups is that abortions presents more regions deleted and in consequence this imbalance could compromise the normal development. Our study provides the first genomic evidence for possible reasons for 45,X embryo lethality. Funding FONCICyT 95419.

957W

Characterization of a rare 16.25 Mb duplication of Xp21.2p22.2. *M.J. Macera[1], E. Bornstein[2], V. Aggarwal[3], L. Northrop[3], M. Shirazi[3], A. Sorbrino[1], B. Levy[3], D. Warburton[4], V. Jobanputra[3].* 1) New York-Presbyterian Hospital, Columbia University Medical Center, New York, NY; 2) Lenox Hill Hospital North Shore-LIJ, New York, NY; 3) Departments of Pathology and; 4) Genetics and Development and Pediatrics, Cloumbia University, New York, NY.

Partial duplications associated with the short arm of the X chromosome remain rare. Found in both sexes, precise delineation of the individual breakpoints was not possible until the use of microarray analysis; to date, only five dup(X) cases have been analyzed by microarray. We report an Xp21.2-Xp22.1 duplication discovered in a 30 year old primi-gravida woman, who presented for chorionic villi sampling at 12 weeks of gestation due to an abnormal first trimester screen. Prior family history was unremarkable. Cytogenetic analysis revealed an abnormal 46,Y,dup(X)(p21.2p22.1).ish dup(X)(wcpX+) in all fetal cells examined. The father was normal 46,XY, while the mother was found to have a 46,X,dup(X)(p21.2p22.1) karyotype. SOMA Microarray analysis on fetal DNA identified a 16.25 Mb gain of chromosomal material in the p arm of the X chromosome from positions 13,458,031-29,710,794, corresponding to Xp21.2 to Xp22.2 The karyotype was revised to: 46,Y,dup(X)(p21.2p22.2).arr Xp22.2p21.2(13,458,031–29,710,794)x2. This duplicated region has 58 OMIM annotated genes; 18 of these are disease-associated genes (TRAPPC2, OFD1, FAAP95, PIGA, AP1S2, NHS, CDKL5, RS1, PHKA2, PDHA1, RPS6KA3, SMPX, MBTS2, SMS, PHEX, SAT1, ARX, and IL1RAPL1) Mental retardation and/or congenital abnormalities in males has been associated with duplication of this region. Analysis of DNA from the mother, using the HUMARA test for X-inactivation skewing, showed the same X to be active in 89% of maternal cells. It is likely that this is due to preferential inactivation of the X with the duplication, providing a normal gene dosage and explaining the mother's normal phenotype. Skewed inactivation has been shown in phenotypic normal females identified with different X chromosome duplications; although the Xp21.2-Xp22.1 duplication appears to be unique. The patient's unremarkable family history would suggest a de novo origin of the duplication in the mother. This information was provided to the parents who elected to terminate the pregnancy and pursue IVF with preimplantation diagnosis for future pregnancies. Precise delineation of genes involved in these rare Xp duplications should improve the ability to predict associated pathology.

958T

Single Nucleotide Polymorphism-Based Microarray Comparison for Detection of Constitutional Copy Number Variation and Absence of Heterozygosity. *H. Mei[1], S.A. Scott[1], H.M. Kearney[2], D. del Gaudio[3], J.D. Weisfeld-Adams[1], M.P. Wasserstein[1], S. Das[3], N. Cohen[1], L. Edelmann[1].* 1) Department of Genetics and Genomic Sciences, Mount Sinai School of Medicine, New York, NY 10029; 2) Fullerton Genetics Center, Mission Health System, Asheville, NC 28803; 3) Department of Human Genetics, University of Chicago, Chicago, IL 60637.

Although microarray-based comparative genomic hybridization is a widely accepted method for detection of copy number variants (CNVs), single nucleotide polymorphism (SNP)-based microarrays can detect genotype and absence of heterozygosity (AOH) in addition to copy number. Three major SNP-based microarray platforms are commercially available (180K ISCA CGH+SNP, Agilent Technologies; CytoScan® HD, Affymetrix; HumanCytoSNP-12, Illumina) and this study aimed to compare them using 24 DNA samples representing a range of constitutional cytogenetic abnormalities. Using platform-specific software, CNVs over 200 kb were also compared among the diploid samples, and the 180K CGH+SNP, CytoScan® HD, and HumanCytoSNP-12 arrays identified an average of seven, one, and three CNVs per sample, respectively. Among these, an average of only one CNV per sample was identified by all three platforms and one additional CNV identified by two platforms. Samples with UPD for chromosomes 2, 6, 7, 8, 14, and 15 were detected by all platforms with strong concordance between identified regions of AOH. Parental samples were available for one UPD15, which indicated a maternal origin by all three platforms. Probands from two consanguineous families were tested, which suggested second and fifth generation parental relationships based on their ~10% and 1.8% of genomic regions identical by descent, respectively. Triplicate testing of one proband sample indicated excellent CNV and AOH reproducibility for all platforms, with slightly greater concordance for the CytoScan® HD and HumanCytoSNP-12 arrays. Triploid samples (69,XXX and 69,XXY) were identified using the HumanCytoSNP-12 array by both aberrant genotype and increased genome-wide copy number, whereas the 180K CGH+SNP and CytoScan HD® arrays identified these samples as abnormal by aberrant genotype and X and Y copy number. Of relevance to prenatal SNP-based microarray testing, artificially derived 46,XX MCC specimens were detected at ≥50% contamination by mosaicism for genome-wide AOH using the CytoScan HD® and HumanCytoSNP-12 arrays. In conclusion, all three arrays reproducibly detected constitutional regions of AOH and parent of origin; however, differences in resolution resulted in differences in the number of detected CNVs. Additionally, the interpretation of genotype information by the native software of the CytoScan® HD and HumanCytoSNP-12 platforms provided an advantage when assessing triploidy and other rare mosaic abnormalities.

959W

Atypical copy number abnormalities in patients with suggestive phenotype of 22q11.2 Deletion Syndrome. *M.C. Molck[1], T.P. Vieira[1], I.C. Sgardioli[1], M. Simioni[1], A.C. Xavier[2], J. Souza[3], V.L. Gil-da-Silva-Lopes[1].* 1) Department of Medical Genetics, University of Campinas (UNICAMP), Campinas, São Paulo, Brazil; 2) Center for Research and Rehabilitation of Lip and Palate Lesions (CRRLPL) Centrinho Prefeito Luiz Gomes, Joinville, SC, Brazil; 3) Center for Intregral Assistance of Cleft Lip and Palate (CIAC), Curitiba, PR, Brasil.

22q11.2 Deletion Syndrome (22q11.2 Del S.) is the most common microdeletion syndrome in humans, affecting 1 in 4,000 live births and with a highly variable phenotype. Chromosome 22q11.2 region contains low copy repeat (LCR) sequences that mediate non-allelic homologous recombination and predispose to copy number abnormalities (i.e. deletions and duplications) at this locus. The most common one is the proximal microdeletion of 3 Mb spanning LCRs A to D. Atypical copy number abnormalities spanning other LCRs have been described, associated or not with the phenotype of 22q11.2 Del S. Three patients with suggestive phenotype of 22q11.2 Del S. with atypical deletion or duplication in 22q11.2 region are described. The chromosomal aberrations were detected by Multiplex Ligation-dependent Probe Amplification (Kit P250, MRC-Holland MLPA®) technique. Patient one showed the ~3 Mb typical deletion and an atypical duplication at distal region (22q11.23) between LCRs F and H. Array analyses by Genome-Wide Human SNP Array 6.0 (Affymetrix) showed a 2.5 Mb deletion at 22q11.21 and a 1.3 Mb duplication at 22q11.23. Patient two demonstrated a ~1.6 Mb atypical deletion at distal 22q11.2 region located between LCRs D and E. Patient three exhibited a ~4.2 Mb atypical deletion overlapping proximal and distal 22q11.2 region and spans the genomic region between LCRs B and F. The breakpoints reported in patients one and three have not been described previously. These findings exemplify the complexity and genetic heterogeneity observed in 22q11.2 Del S. and corroborate the idea that genetic modifiers contribute to the phenotypic variability found in this syndrome. Financial support: FAPESP, CAPES and CNPQ.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

960T

Mosaicism in live born and spontaneous abortions with trisomy 13, 18, 21 and monosomy X. *S. Sánchez[1, 4], P. Grether[2], D.G. Mayén[3], B. Molina[1], M.J. Zavaleta[2], J. Domínguez[1], R. Meléndez[3], S. Frias[1,5].* 1) INSTITUTO NACIONAL DE PEDIATRIA, MEXICO CITY, Mexico; 2) INSTITUTO NACIONAL DE PERINATOLOGIA, Mexico; 3) HOSPITAL ANGELES LOMAS, Mexico; 4) POSGRADO EN CIENCIAS BIOLOGICAS, UNAM, Mexico; 5) INSTITUTO DE INVESTIGACIONES BIOMEDICAS, UNAM, Mexico.

In humans, the study of first trimester abortions by conventional and molecular cytogenetics, have shown that the frequency of chromosomal abnormalities is 60%, mainly trisomies and monosomy X. Over 90% of aneuploid pregnancies are lost as early abortions, 5.8% as still births and only 4% born alive. Aneuploidies that survive are trisomies (T) 13, 18 and 21 and monosomy X. There are several studies on aneuploidy as a cause of spontaneous abortion, but remains unknown why a fetus with trisomy or monosomy comes to birth? The evidence shows that newborns with T13, 18, 21 and 45,X are exceptions to the rule of lethality and the reason for this has remained unexplored. Mosaicism with a normal cell line in embryos 45,X has been proposed as a survival factor, however there is no information regarding the autosomal trisomies. The aim of this study is to determine if there exist differences in the presence of mosaicism and the amount of disomic cells between born alive and abortions with the same aneuploidy. We analyzed peripheral blood samples (n=35) and desquamation bladder cells (n=20) of born alive and spontaneous abortions tissues (n=33). We searched for the presence of a normal cell line, in 500 to 1000 cells by fluorescence in situ hybridization (FISH) with probes for the detection of loci: 21q22.13-q22.2 (LSI21), 13q14.13 (LSI13-RB), α-centromeric 18 (CEP 18), α-centromeric X (CEP X) and α-centromeric Y (CEP Y). Samples with more than 3% of normal cells were considered mosaics. We detected mosaicism in live born in 1/3 cases with T18, 3/15 with T21 and 3/12 with 45,X; in T21 normal cells where found only in desquamation bladder cells; T13 cases did not show mosaicism. We analyzed 33 spontaneous abortions samples, and found mosaicism in 3/6 with T13, 1/7 T18, 2/8 T21 and 3/10 45,X. In one case with T13 and other with monosomy X the mosaicism was found in chorionic villi and amniotic membrane from de same miscarriage. In conclusion, we detected mosaicism with normal cells in 20% of born alive samples and in 27% of miscarriages. Most of them were found in cases with 45,X karyotype. These results suggests that the mosaicism is a frequent finding in live birth and abortions with trisomies 13,18,21 and monosomy X and due to the similar frequency, it is possible that mosaicism does not confer a biological advantage that explain the survival of live births with aneuploidy. This work was funded by FONCICYT-95419. SSanchez is a fellow of CONACyT-100254.

961W

Array-CGH analysis in patients with Goldenhar Syndrome. *P.AC. Santos[1], E.L. Freitas[2], H.P.N. Safatle[3], C. Rosenberg[2], I. Ferrari[4], S.F. Oliveira[4], J.F. Mazzeu[1,5].* 1) Programa de Pós-graduação em Ciências da Saúde, Universidade de Brasília, Brasília, Distrito Federal, Brazil; 2) Departamento de Genética e Biologia Evolutiva, Instituto de Biociências, Universidade de São Paulo, São Paulo, Brazil; 3) Hospital Universitário, Universidade de Brasília, Distrito Federal, Brazil; 4) Laboratório de Genética, Departamento de Genética e Morfologia, Instituto de Ciências Biológicas. Universidade de Brasília, Brasília, Distrito Federal, Brazil; 5) Programa de Pós-Graduação em Ciências Genômicas e Biotecnologia, Universidade Católica de Brasília, Brasília, Distrito Federal, Brazil.

Goldenhar syndrome (GS), also known as oculo-auriculo-vertebral syndrome, is a congenital defect from anomalous development of first and second branchial arches. GS presents a spectrum ranging from mild to severe forms and the patients may exhibit anomalies including ear malformations, facial asymmetry, cleft lip/palate, eye, vertebral and congenital heart defects. GS is heterogeneous and its etiology includes environmental factors, unknown genetic factors and different chromosome aberrations. Deletions involving 2p11, 2q11, 5p15, 12p13, 13q21, 14q32 and 22q11 as well as duplications at 4q35, 8q11, 9q34, 11q21, 13q13, 14q22, 18p11, 20p12 and a translocation at t(4;8)(p15.3;q24.1) disrupting *BAPX1* have been previously reported in GS patients . A candidate region at 14q32 was proposed as by linkage studies. We investigated eight patients with GS for chromosome rearrangements by G-banding and array-CGH (44K, Agilent technologies). G-banding did not disclose any rearrangements. In one patient, array-CGH revealed a ~2.5 Mb deletion at genomic position chr22: 17,276,772-19,770,655 (Build36/hg18) at 22q11.21. This patient presents clinical features of both Goldenhar and DiGeorge syndromes, namely: hemifacial microsomia, ptosis, dysplastic right ear, retrognathia, aphasia and hearing loss. She also presents with joint laxity and myopia. ECG showed a mitral valve prolapse. 22q11.2 microdeletions have been described in patients with GS, however they do not overlap with the deletion here reported. In another patient array-CGH revealed a ~550 Kb duplication at genomic position chr20:8,794,395-9,336,676 (Build36/hg18) at 20p12.3-p12.2. He presents macrocephaly, facial asymmetry, microtia on the left and cryptorchidism. The duplication was inherited from the phenotypically normal father and have not been reported in other GS patients or normal controls. Only one gene maps to this segment *PLCB4*, that is part of a gene family which operates in intracellular transduction of extracellular signals. Further studies are needed to infer the relation of this microduplication and the occurrence of GS clinical signs. No abnormalities were identified in the six remaining cases. This study corroborates previous reports indicating that patients with GS should be screened for chromosomal imbalances. Financial support: CAPES, FAPDF/PPSUS.

962T

Complex rearrangements involving five chromosomes and at least fourteen breaks. *J. Wang, B. Huang, R. Habibian, A. Hajianpour.* Integrated Genetics, Monrovia, CA.

Array CGH (aCGH) was performed on a 20-month old child with psychomotor delay, facial dysmorphism, cryptorchidism, and dermal sinus. Chromosome analysis performed previously in another laboratory revealed a de novo apparently balanced three-break reciprocal translocation: 46,XY,t(4;6;10)(q31.1;q25.1;q22.1). aCGH revealed interstitial deletions of chromosomes 4, 6, and 20 [arr 4q31.22(145,181,574-145,568,241)x1,6q25.3(155,961,736-156,674,789)x1, 20p12.1(13,142,859-13,806,764)x1,20p12.1(14,314,316-14,622,432)x1]. Multiple FISH analyses were performed. Combining cytogenetic, FISH, and aCGH analyses, a complex rearrangement involving five chromosomes (4, 6, 9, 10, 20) was detected. There was an unbalanced three-way translocation between chromosomes 4, 6, and 10, with deletion of DNA sequences around the breakpoints at 4q and 6q. In addition, further rearrangements involving chromosomes 9 and 20 were also observed. These included (1) a small portion of the segment deleted from 6q25.3 was inserted into the proximal short arm of chromosome 20 at p11.2; (2) two non-contiguous interstitial deletions in the short arm of chromosome 20 at 20p12.1 were detected by array. FISH analyses indicated the intervening sequence and the sequence proximal to the two non-contiguous interstitial deletions were inserted to the proximal long arm of chromosome 9 at 9q13; (3) another segment of chromosome 20, possibly also 20p12.1, was inserted to the long arm of chromosome 4 possibly at 4q31.22. There were two breaks on der(4), three on der(6), at least one on der(9), one on der(10), at least seven on der(20), with a minimum of fourteen breaks. Complex chromosome rearrangements, both familial and de novo, balanced and unbalanced, have been reported repeatedly. Massive chromosomal shattering and reorganization, a phenomenon termed chromothripsis, has recently been reported and suggested to occur in 2%-3% of all cancers. Chromothripsis has also been shown to occur in constitutional de novo structural rearrangement. The rearrangement in this child is potentially more complex than described. Further study of this and other similar cases may provide additional insight to the mechanism of genomic rearrangements.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

963W

Trisomy Xp and partial tetrasomy Xq resulted from gain of a rearranged X chromosome in a female fetus: pathogenic or not? *M. Yiu[1], Z. Qi[1], A. Ki[3], K. Hashimoto[3], M. Ernster[2], J. Yu[1]*. 1) Laboratory Medicine, UCSF, San Francisco, CA; 2) Golden Gate Pediatrics, San Francisco, CA; 3) Clinical Cytogenetics Lab, UCSF Medical Center, San Francisco, CA.

A mosaic karyotype with gain of a rearranged X: 47,XX,add(X)(q?22)[11]/46,XX[9], was detected by cytogenetic analysis of chorionic villous sampling due to advanced maternal age. Microarray and additional studies indicated that the rearranged X carried a 39.5 Mb inverted duplication between Xq11.1 and Xq22.1, a 53.8 Mb deletion distal to Xq22.1 and a satellited Xqter. Both parental karyotypes were normal. Gain of this rearranged X was confirmed in all analyzed cells from a follow-up amniocentesis and in approximately 92.5% of the cells from a postnatal cord blood sample. Nonetheless, prenatal ultrasonography imaging including fetal echocardiography appeared to be unremarkable and a full-term baby girl was born at 39 4/7 week via a spontaneous vaginal delivery. The infant appeared to have normal development at the 2-month follow-up examination. Late replication studies showed that the rearranged X was inactivated in all abnormal cells examined. This skewed X-inactivation might suppress potentially detrimental phenotypic effects of genomic imbalance. Although, the possibility of tissue-specific mosaicism that result in mild phenotypes cannot be excluded. Gain of X chromosomes, particularly rearranged X chromosomes, often presents challenges for prenatal genetic counseling. It is well-known that gain of X chromosomes is associated with abnormal phenotypes, and the increase of additional Xs is often associated with more severe phenotypes. However, the regions responsible for the abnormal phenotypes on X chromosome are poorly understood. Long-term follow-up of this case may provide better understanding of phenotypic effects of X inactivation and X chromosome abnormalities.

964T

A de novo 0.24-Mb microdeletion of 16q24.3 encompassing ANKRD11 in a patient with KBG syndrome. *HJ. Cho[1], EJ. Seo[1,2], JH. Lim[1,2], JO. Lee[4], HW. Yoo[1,3], CK. Cheon[5]*. 1) Medical Genetics Center, Asan Medical Center, Seoul, South Korea; 2) Department of Laboratory Medicine, Asan Medical Center, Seoul, Korea; 3) Asan Medical Center Children's Hospital, University of Ulsan College of Medicine, Seoul, Korea; 4) Asan Institute for Life Sciences, Asan Medical Center, Seoul; 5) Department of Pediatrics, Pusan National University, Pusan, Korea.

KBG syndrome is a rare condition characterized by macrodontia of the upper central incisors, distinctive craniofacial findings, skeletal anomalies, short stature, and neurologic involvement that includes seizure, intellectual disability, and global developmental delay. The diagnosis of KBG syndrome is mostly based on clinical features. Recently, it has been reported that the mutations in ANKRD11 cause KBG syndrome, Now we present a boy with the phenotype of KBG syndrome with the microdeletion of ANKRD11. The patient at the age of 6 years was referred to our hospital because of short stature and developmental delay. He showed short stature, macrodontia of permanent dentition, and craniofacial dysmorphism including triangular face, prominent forehead, epicanthal fold, anteverted nostrils, broad mouth, and prominent ears. He had also neurologic involvement such as delay of speech and motor function, intellectual disability (IQ 70) and partial seizure on EEG. Brain MRI showed a thickened corpus callosum. Bone X-ray showed narrowing and elongation of both iliac wing and body, small femoral heads, small distal tibial epiphyses and tibiotalar slanting, and delayed bone age. Conventional chromosome analysis was normal, but a de novo heterozygous 0.24-Mb deletion at 16q24.3 including ANKRD11 and a part of the SPG7 genes was detected by the 244k oligonucleotide based array-CGH (Agilent). Recent study reported that patients with heterozygous microdeletion of ANKRD11 showed the phenotype similar to that of KBG syndrome. ANKDR11 encoding ankyrin repeat domain 11 seems to function in transcriptional repression and promoting transcription. However, it is uncertain how the pathogenic mechanism of ANKRD11 mutations cause KBG syndrome. The phenotype shown in this patient was compatible with clinical features described in KGB syndrome with ANKRD11 mutations. We suggest that the heterozygous microdeletion of ANKRD11 could also cause the KBG syndrome.

965W

The Evaluation of Subtelomeric Fluoresent in situ Hybridization Analysis of Idiopathic Mental Retardation / Multiple Congenital Anomaly Cases. *B. Durmaz[1], E. Karaca[1], A. Durmaz[1], T. Atik[2], H. Akin[1], O. Cogulu[1,2], F. Ozkinay[1,2]*. 1) Department of Medical Genetics, Ege University, Izmir, Turkey; 2) Department of Pediatrics, Ege University, Izmir, Turkey.

Mental retardation is a life-long disability that occurs in approximately 2–3% of the population. The etiology of intellectual and developmental anomalies remain unknown in about 50% of cases which limits the effectiveness of genetic counseling on recurrance risks, prognosis and treatment. Conventional cytogenetic analysis is the standard method for investigating chromosomal abnormalities however, small (<5 Mb) DNA changes could not be detected. Fluoresent in situ Hybridization (FISH) analysis is used to show those cryptic changes. The subtelomeric regions of chromosomes which can be visualized by FISH, are considered to be gene-rich and rearrangements involving these regions have been shown to cause mental retardation in many studies varying from 5 to 10% of the affected cases. In this study, subtelomeric FISH analysis of 161 cases with idiopathic mental retardation / multiple congenital anomalies (MR/MCA) which do not have any abnormalities in classical cytogenetic analysis were retrospectively evaluated between January 2005 - April 2012. In 9 patients (5.59%), subtelomeric deletions were detected . These were deletions of 14q in 2 cases, 18q in 2 cases, 18p in 2 cases, 15q in 1 case, 7p in 1 case and 8p in 1 case. Deletions of 18p, 18q and 8p were familial, the others were de novo. The prevalence of subtelomeric deletions in our cohort of moderate to severe mental retardation patients is consistent with the literature. In conclusion, subtelomeric FISH was shown to be feasible and cost-effective in determining the etiology of MR/MCA patients.

966T

Mechanism-based Analysis of Human Cell Lines of Prader-Willi Syndrome. *Z. TANG, D. Berlin, M. Wineburg, A. MacMillan, D. Altamuro, L. Toji, C. Beiswanger, S. Madore, N. Gerry*. Coriell Institute for Medical Research, Camden, New Jersey 08103.

Prader-Willi Syndrome (PWS) is one of the most common conditions seen in genetics clinic, and it is the most common genetic cause of obesity seen in both sexes and all races. The National Institute of General Medical Sciences Human Genetic Cell Repository (NIGMS repository) housed at Coriell Institute has collected a group of cell lines originating from PWS patients and their family members. A total of 32 lines consisting of 24 proband and 8 family members are currently available to the research community. We have characterized each of these cell lines using conventional G-banded karyotyping and molecular cytogenetic methods such as fluorescence in situ hybridization (FISH) and chromosomal microarray analysis (CMA). Based on these findings and potential genetic mechanisms causing PWS, the cell lines can be categorized into 4 different groups. Group 1 consists of 18 cell lines (16 proband and 2 family members) containing a deletion ranging from 45 kb (partial or whole SNRPN gene) to >6 Mb (the whole region) affecting 15q11.2q13. Group 2 includes 2 cell lines that have no deletion in 15q11.2q13 but do contain a maternal uniparental disomy (matUPD) in this region. One line is homozygous, indicating two identical alleles, while the other is heterozygous. Group 3 includes 2 cell lines with a microdeletion affecting the imprinting center (IC). Group 4 contains 4 cell lines with an as yet uncharacterized PWS genotype. Two of the cell lines in Group 4 have a chromosomal translocation involving the 15q11.2q13 region as detected by G-banded karyotyping and FISH studies, but no copy number change in this region was seen by CMA. Further studies such as DNA sequencing for potential imprinting mutations are needed to clarify the underlying mechanism(s) in these 4 cell lines. These groups of cell lines comprise a useful PWS-specific reference panel for research and clinical diagnostic applications.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

967W

A child with Pitt-Hopkins syndrome and maternal somatic mosaicism at 18q21.2. *K. Doudney[1], V. Bickley[1], M. Robertson[2], J. Watt[2], J. Taylor[2], A. Kidd[3], P. George[4].* 1) Molecular Pathology, Canterbury Health Laboratories (CHL); 2) Cytogenetics, CHL; 3) Clinical Geneticist, CHL; 4) Clinical Director, CHL, Canterbury District Health Board, Christchurch, New Zealand.

Pitt-Hopkins syndrome (PTHS) is a neurodevelopmental disorder often caused by heterozygous mutations in TCF4, a transcription factor important during embryonic development. Typical PTHS characteristics are severe intellectual disability, breathing abnormalities and abnormal facial morphology including a wide mouth. Following a diagnosis of global developmental delay at 21 months of age, patient 1 was referred to our laboratory for arrayCGH. We detected an approximately 600kb heterozygous deletion at 18q12.2, encompassing the 5' half of TCF4 and 300kb of upstream genomic sequence. Analysis of the proband's mother revealed a mosaic pattern of deletion at this locus, identified by FISH in 11% of interphase nuclei and confirmed in metaphase cells. Mosaicism at 18q21.2 (TCF4) is rare, with only a few examples reported (1–3). This finding provides further evidence that TCF4 underlies Pitt-Hopkins syndrome and establishes haploinsufficiency at this locus as underlying the disorder's dominant mode of inheritance. 1 Steinbusch et al. (2012) Clin Genet doi: 10.1111/j.1399-0004.2012.01857.x. [Epub ahead of print] 2. De Pontual et al. (2009) Hum Mutat 30(4):669–676 3. Stavropoulos et al. (2010) Eur J Med Genet 2010: 53(6): 396–399.

968T

Microarray benefits diagnostic pregnancy loss studies. *D.L. Pickering, B.J. Dave, D.M. Golden, A.A. Haggerty, D.L. Bishay, R.L. Smith, S.A. Fisher, W.G. Sanger*. Human Genetics Laboratories, Munroe Meyer Institute, University of Nebraska Medical Center, Omaha, NE.

Chromosome studies of spontaneous abortion tissue provide important diagnostic information; however, these tests are hampered by culture failure and maternal cell contamination. FISH analysis circumvents the need for cell culture yet aneuploidy screening for selected chromosomes does not provide comprehensive genomic data. Microarray studies in products of conception (POC) are useful because they eliminate the need for cell culture, interrogate the entire genome, and can distinguish maternal contamination upon comparative analysis with maternal blood specimen. From August 2011-May 2012, our laboratory performed 54 pregnancy loss microarray tests using a combination 180K array with approximately 140,000 oligonucleotide probes for copy number analysis and 40,000 SNPs for genotype assessment. DNA extraction was successful and microarray results were obtained in all 54 cases, 35 females and 19 males. Nine out of 54 (16.6%) cases were abnormal by microarray analysis. There were six cases with whole chromosome aneuploidies including trisomy 2 (one case), trisomy 21 (three cases), and monosomy X (two cases). Structural abnormalities of 22q were observed in three cases. One was a common 22q deletion observed in a male POC specimen. The other two were female POC studies that showed duplications of 22q11.2 in maternal blood and POC tissue. Microarray studies thus helped determine subtle maternal abnormalities during POC genomic investigations. Maternal contamination was determined in nine out of 35 female POCs (7/29 normal; 2/6 abnormal) which could not have been detected by FISH. Maternal contamination was also detected in a male POC specimen and confirmed by interphase FISH in 70% of cells from a tissue touch preparation, emphasizing the ability of microarray to detect mosaicism. Furthermore, the use of microarray reduced the turnaround time of pregnancy loss studies to 8.2 days. Incorporation of microarray testing for miscarriage tissues has thus improved patient care by increasing success and abnormality rates, reducing the time to report, and providing more accurate results in cases with maternal contamination. Additionally, detection of maternal abnormalities during pregnancy loss investigations underscores the need for genetic counseling and peripheral blood microarray studies among couples with recurrent miscarriage.

969W

Method Comparison Study of the Illumina InfiniumDx CytoSNP-12 Assay for the Detection of Chromosomal Abnormalities in DNA Extracted from Peripheral Blood. *M. Porter, R. Golshani, J. Fabian, M. Harris.* Illumina Inc, San Diego, CA.

Introduction / Background: The InfiniumDx™ CytoSNP-12 Assay is intended for use as an aid in post-natal diagnosis of autism spectrum disorders, developmental delay and/or intellectual disability, congenital anomalies, and dysmorphic features in conjunction with other currently used information. To demonstrate the utility of this assay, 402 samples with reported chromosomal abnormalities were analyzed using the InfiniumDx CytoSNP-12 Assay and at least one other method, including karyotype, FISH, PCR, Multiplex Ligation-dependent Probe Amplification (MLPA), qPCR, or an in-house confirmatory method for loss of heterozygosity (LOH). Materials and Methods: The InfiniumDx CytoSNP-12 Assay, based on Illumina BeadArray™ technology and run on the iScanDx System, consists of CytoSNP-12 assay kits and KaryoStudioDx™ analysis/reporting software. The 402 samples were run in singlet. Karyotype, MLPA, FISH, and PCR were performed externally with commercially available assays. qPCR and LOH were performed using methods validated in-house. Sample success rate (percent of samples passing defined QC metrics/total samples run), positive percent agreement (percent of regions reported by a comparator method and the assay/total regions expected) and false positive rate (percent of regions detected by the assay but not by a comparator method/total regions detected by the assay) were calculated to measure the performance of the InfiniumDx assay against the comparator methods. Results: Sample success rate upon first attempt was 98% (393/402). The 95% lower bound for positive percent agreement for all chromosomal abnormality types was 93% (1034/1098). For gains, losses, and copy-neutral LOH (CNLOH), the 95% lower bound for positive percent agreement for each type was 87% (22/224), 91% (25/365), and 95%(17/445), respectively. False positive rate across all regions detected by the assay was 12% (144/1170). For gains, losses, and CNLOH, the false positive rate for each type was 10% (25/248), 9% (37/392), and 15% (82/530), respectively. Conclusions: Our study demonstrates comparable performance of the InfiniumDx™ CytoSNP-12 Assay to many of the standard clinical methods used to determine copy number variation. This assay reports all chromosomal abnormalities within the entire genome of a sample, making it more comprehensive than FISH, PCR, MLPA, and qPCR. In addition, this assay reports CNLOH regions, unlike karyotyping or FISH. This test is For Investigational Use Only.

970T

A CML case with normal karyotype and positive *BCR/ABL1* fusion. *C.A. Tirado[1], JW. Boles[2], J. Ochoa[2], R. Collins[3].* 1) Pathology & Laboratory Medicine, UCLA, Los Angeles, CA90024, CA; 2) Department of Ecology & Evolutionary Biology UCLA; 3) UT Southwestern Medical Center- Department ofHematology & Oncology.

Chronic myeloid leukemia (CML) is cytogenetically characterized by the presence of the Philadelphia chromosome (Ph) in 90–95% of cases. Ph variants are found in approximately 5% of CML patients. In these cases, FISH is always useful to determine the location of the BCR/ABL1 fusion gene. Herein, we present a 47 year old patient who presented with splenomegaly, dizziness, vertigo and high WBC (> 700×10^9). Morphologic examination of the bone marrow (BM) biopsy was consistent with CML in chronic phase with no increase in blasts. Chromosome studies revealed a 46,XY,t(9;22) in all cells examined. He started with hydroxyurea and Gleevec. After six months his WBC was $53x\ 10^9$ with mild thrombocytopenia and recurrent chronic phase with no increase of blasts but 95% cellularity. Chromosome analysis on BM at this time showed a normal chromosome complement but FISH on interphase nuclei showed only one fusion *BCR/ABL1* signal in 195/200 of the cells examined. Seventeen of those cells showed an extra *BCR* signal. FISH on previously G-banded metaphases showed a fusion signal on the Ph which confirmed the cryptic translocation. This finding was characterized as ins(22;9)(q11.2;q34q34). The patient then had increase in the dose of Gleevec (800 mg) but due to some pancytopenia this was switched to 600 mg. Pancytopenia did not improve so he was switched to sprycel 70 mg. After treatment his WBC are 4.2, Hb 9.6, platelets 67×10^9 although mild anemia and thrombocytopenia was noted. To date, the patient has remained stable in a chronic phase with partial hematological remission. Our study confirms the advantage of using FISH to determine the location of the *BCR/ABL1* fusion gene in cases where conventional cytogenetics is normal and FISH on interphase nuclei show evidence of *BCR/ABL1* fusion.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

971W

Computational methods for detecting whole-genome triploid samples with Agilent CGH+SNP arrays. *A. Vadapalli[1], B.J. Peter[2], A. Ashutosh[1], J. Ghosh[1], B. Curry[2], A. Rajkovic[3], U. Surti[3], S.A. Yatsenko[3]*. 1) Genomics R&D, Agilent Technologies, Santa Clara, CA; 2) Agilent Labs, Agilent Technologies, Santa Clara, CA; 3) Department of Obstetrics, Gynecology and Reproductive Sciences, University of Pittsburgh, PA.

Array-based comparative genomic hybridization (array CGH) is a tool to provide high-resolution copy number measurements of the genome. Array CGH has become important for research in identifying constitutional diseases, cancer biology and in conducting pre or postnatal studies. The addition of single nucleotide polymorphism (SNP) measurements to CGH microarrays enables the detection of copy-neutral loss of heterozygosity (cnLOH) events and allelic imbalances.

Fetal loss, or miscarriage, is one of the most common complications of pregnancy, occurring in 15 to 20% of clinically recognized pregnancies. Cytogenetic abnormalities are present in approximately 50% of miscarriages, and the majority of these are due to whole chromosome aneuploidy or triploidy. Karyotyping has limited resolution and depends on successful culturing of cells from the product of conception. High resolution array CGH analysis can be performed on genomic DNA, but copy number analysis alone cannot detect cases of triploidy when a whole haploid set of chromosomes is present in three copies.

A novel algorithm has been developed in Agilent Genomic Workbench 7.0 software that allows for the detection of triploidy when CGH+SNP arrays are used. In an earlier version of the algorithm, samples with whole-genome triploidy were normalized and previously assigned to be non-aberrant. For constitutional triploid samples, the algorithm identifies a pattern of SNP log ratio distribution of the majority of genome consistent with triploidy. The parameter values of the fitted peaks in the SNP log ratio distribution are used to automatically detect and adjust log ratios to represent the non-diploid state across a majority of the chromosomes in such samples. Upon determining the correct ploidy, the total copy number and the SNP genotype of the genomic regions is adjusted accordingly.

A new algorithm has been applied to study samples with 69,XXX, 69,XXY, and 70,XXY,+2 chromosome complements. We have shown that the new algorithm allows researchers to determine the genotypes, the total copy numbers and SNP allele specific copy numbers in triploid samples. Since 1–2 % of all pregnancies are triploid, this novel algorithm provides an improved research tool for the detection of chromosomal abnormalities.

972T

Genome-wide resolution study of CNV detection by the CytoScan® HD Cytogenetics Array through a large-scale experiment. *Y.S. Wang, A.H. Roter*. Informatics, Affymetrix, Santa Clara, CA.

Accurate estimation of functional performance of genomics technologies is challenging due to limited access to clinical samples representing the wide range of aberration possibilities across the genome. The nature of their non-probability sampling and low sample availability increases the risk of bias in estimates of population parameters. We hypothesize that random sampling with available samples can overcome these drawbacks. We tested our hypothesis using this method to estimate array functional resolution for Affymetrix' CytoScan® HD Array. CytoScan HD Array is a 2.6 million marker high-density microarray designed to detect whole-genome copy number variation (CNV) for constitutional and hematological malignancy cytogenetic research applications. To assess functional performance, it is critical to determine of the number of markers required to make an accurate and reproducible CNV call across the genome. A non-existent ideal sample set is a collection of biological DNA samples with CNV regions of different sizes at all genome locations. Developers could randomly sample and run microarrays to estimate performance. Alternatively, developers could use synthetic or sub-cloned DNAs and targeted deletions spiked into normal genomic DNA. Arrays could then be run to estimate performance. Conducting such experiments is expensive and impractical. As a highly effective alternative, we designed an in silico experiment that simulates the above studies using the actual behavior of markers on real clinical samples. We obtained a set of samples from the Coriell Institute with abnormal copy number regions collectively covering a large proportion of the genome and ran these on CytoScan HD Arrays. We randomly sampled DNA pieces of different sizes from the karyotype-confirmed large CNV regions. Each sample was flanked with copy number markers from normal regions and subsequently processed for CNV detection. This experiment used real clinical samples and enabled an estimation of sensitivity. The derived statistics summarized millions of specific cases across the genome. Based on these data, we have established the relationship between number of markers to make a high-confidence CNV call and the associated sensitivity of CNV detection. In addition, we have characterized the difference in sensitivity between copy number losses and gains. This study demonstrates the feasibility of accurate estimation of population parameters using limited non-probabilistic samples.

973W

A new case of inverted duplication with terminal deletion of chromosome 12p caracterized by FISH and SNP array analysis. *E. Pipiras[1], B. Jeandidier[2], S. Chantot-Bastaraud[3], B. Benzacken[1], A. Delahaye[1]*. 1) Hopital Jean Verdier, Service de Cytogénétique, APHP, Bondy, France; 2) Hopital Jean Verdier, Service de Pédiatrie, APHP, Bondy, France; 3) Hopital Trousseau, Service de Génétique, APHP, Paris, France.

Objective: So far, the inverted duplication with concomitant terminal deletion (invdupdel) has been identified for several chromosomes but was not studied for the short arm of chromosome 12. We describe clinical features and molecular characterization of an invdupdel 12p in a child with a developmental delay. **Methods:** Clinical work up included clinical histories, physical, neurological, and ophthalmologic examinations. A standard karyotyping was performed. The structural abnormality was characterised by molecular cytogenetic methods as FISH (Fluorescence in situ Hybridisation) using whole chromosomes painting 12, 12pter subtelomeric probe, bacterial artificial chromosome (BAC) probes and chromosomal microarray analysis (CMA) using whole genome HumanCyto-12 SNP (Illumina) array. **Results:** The patient is a one year-old girl (non consanguineous parent) who was referred to clinic because of hypotonia, hyperlaxity, psychomotor retardation, developmental delay, mild dysmorphic features including upslanting palpebral fissures and a hockey-stick palmar crease. A 46,XX,add(12)(p13) was suspected in the routine karyotype analysis of the patient while both parents revealed an apparently normal karyotype. The subtelomeric probe 12p showed terminal deletion. Painting of chromosome 12 excluded translocation and confirmed the interstitial duplication. The inverted orientation of the duplication was determined by either the banding pattern and dual coloured FISH. Molecular characterization of the deleted region by SNP array confirmed 12p terminal deletion, a 0,8Mb region that comprises 9 genes. The inverted duplication region spans 27,2Mb and encompasses 272 genes (hg18-build 36.3). SNP array failed to identify single copy region, between the duplication regions. **Conclusions:** We are reporting the first case of inverted duplication of 12p13.33p11.22 with a deletion of 12p13.33-pter. Our cytogenetic molecular results and their correlation with the phenotype of the patient are discussed, mechanisms (U-type exchange or NAHR) of invdupdel also are reviewed. In this study we have described a case in which, molecular cytogenetics (FISH and CMA analysis) can determine breakpoints and the exact size of the chromosome abnormality detected by karyotype.

974T

GISH Analysis of Genome Discrimination in Two Interspecific Hybrid of Sweet Potato. *Z. Li.* Jiangsu Normal University, Xuzhou, China.

To understand the genomic structure and chromosomal composition of two interspecific hybrids 59-1 and 67-1 from I. batatas cv. Xushu18 and I. grandifolia and I. hederacea. Jaq., comparative genomic in situ hybridization was used to study the sweet potato, wild species and interspecific hybrids. The results significantly indicated the vast majority of genome of 59-1 and 67-1 were from that of Xushu18. The chromosome number of the two hybrids was 75 and 90, respectively, different from the theoretical number. GISH detected two or more chromosomes and or some chromosomal segments from the male parent wide species. Sequences homologous with the genome of male parent wide species were more distributed in the regions of centromere, subcentromere, telomere, subtelomer. GISH signals of large size and strong intensity were distributed along entire lengths of most chromosomes. Presumably, in the process of nurturing, We speculate that in the process of hybrid cultivation, a or of the parents may have occurred in genome of the two hybrids. the parental genome in hybrids 59-1 and 67-1 occurred chromosome combination or genome reorganization, amplification, and structural rearrangement.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

975W

Latest innovations in oligo FISH enable high resolution detection of chromosomal aberrations. *M. Ruvolo[1], V. Kulkarni[1], B. Mullinax[1], A. Bergstrom Lucas[1], P. Tsang[2], N. Faravashi[1], A. De Witte[1], M. Srinivasan[1], E. LeProust[1], S. Fulmer-Smentek[1]*. 1) Agilent Technologies Genomics, Santa Clara, CA; 2) Agilent Laboratories, Santa Clara, CA.

The discovery of complex structural variations that exist within individual genomes has prompted a need to visualize chromosomes at a higher resolution than previously possible. In response to this need, we have developed a new generation of fluorescent in situ hybridization probes for the detection of chromosome aberrations, including translocations and copy number changes. These new probes, Agilent SureFISH probes, are designed using an in silico design strategy that specifically avoids placing oligonucleotides in repetitive regions of the genome, including repeat-masked regions and segmental duplications, allowing for highly specific detection of the region of interest. Each SureFISH probe is designed to a specific region of the genome with different design strategies employed for different applications. The probes are generated from complex libraries containing hundreds to thousands of unique high quality long oligonucleotides. The resulting probes provide high specificity and enable users to detect aberrations in targeted regions of the genome as well as aberrations near highly repetitive elements. These results can be obtained using a workflow compatible with existing technologies, with the improvement that suppressive hybridization agents are not required and the hybridization time can be shortened. Because of their unique design methodology, Agilent SureFISH probes provide superior resolution and specificity as compared to other available technologies.

976T

Deletion 2q24.1q24.2 in a boy with generalized hypotonia and developmental delay: further delineation. *V. del Castillo[1], E. Yokoyama[1], C. Villarroel[1], J.L. Castrillo[2], S. Avila[2], S. Sánchez[3], B. Molina[3], S. Frías[3]*. 1) Departamento de Genética Humana, Instituto Nacional de Pediatria, Mexico, DF, Mexico; 2) Laboratorio Genetadi, Bilbao, España; 3) Laboratorio de Citogenética. Departamento de Genética Humana, Instituto Nacional de Pediatria, Mexico, DF, Mexico.

BACKGROUND: Many patients with hypotonia and developmental delay/mental retardation have a chromosomal etiology, and often the submicroscopic deletion or duplication is only evidenced by molecular cytogenetic methods, as array-CGH. To date, there are only three previous cases with an interstitial deletion involving 2q24.2 so we present the first Mexican patient with this alteration. CASE REPORT: The patient is the first child of an 18-years-old mother and 20-years-old father. The parents were healthy and non-consanguineous. He was born after an uneventful pregnancy. There was no family history of children born with congenital disease, intellectual disability, autism, seizures, neurologic disorders, metabolic disorder, recurrent pregnancy loss or infertility. Birth weight was 2600 g (centile 5), height was 47 cm (centile <5). At 6 months of life the patient showed generalized hypotonia, developmental delay and cryptorchidism. Karyotype was normal. At this time he is almost 7 years-old boy who have no started to speech and has not bladder and bowel control. Facial appearance is normal. His weight is 12,5 Kg (centile <5), height is 102 cm (centile <5), and cranial circumference 48 cm (centile <5). Tubular renal acidosis and gastroesophageal reflux disease has been documented. Cerebral computed axial tomography imaging (TAC) did not reveal any alteration. RESULTS: Oligonucleotide array-CGH was performed using the Agilent Human Genome CGH microarray 44K (Agilent Technologies, Santa Clara, CA, USA) and revealed a 7.3Mb interstitial deletion of chromosome region 2q24.1q24.2 encompassing 30 genes, which was absent in parents. DISCUSSION: Only one of the three patients has almost the same interstitial deletion, and our patient share clinically with the other cases the hypotonia and the developmental delay/mental retardation. Palumbo O et al, suggested 5 candidate genes, however only 3 of them (SLC4A10, TBR1, KCNH7) were present in all patients including our case. We considered that these genes have an important role in the brain development and the haploinsufficiency of them could be related to the developmental delay in all patients. Finally, in conclusion we agree with Palumbo O et al that deletions of chromosome region 2q24.1q24.2 could be associated to a specific phenotype. REFERENCES: 1) Palumbo O, et al. Mol Cytogenet 2012;5:1. 2) Takatsuki S, et al. Am J Med Genet 2010;152A:1020-5. 3) Magri C, et al. Eur J Med Genet 2011;54:361-4. CONACYT-SALUD 87792-08.

977W

SNP arrays provide incremental, clinically significant information for patients undergoing chromosome microarray testing. *V. Aggarwal[1], O. Nahum[1], B. Marmol[1], K. Anayane-Yeboa[2], B. Levy[1]*. 1) Pathology and Cell Biology, Columbia University, New York, NY; 2) Pediatrics, Columbia University, New York, NY.

Chromosome microarray analysis is currently recommended as first tier testing in evaluation of children with developmental disability, autism or multiple congenital abnormalities. Most of the literature supporting this is experience with array comparative genomic hybridization (aCGH) platforms. We describe our experience using the Affymetrix SNP arrays for 150 patients referred for the same reasons mentioned above. An abnormal result was identified in 20% of the patients. Strikingly, most of the abnormalities detected are clinically significant copy neutral changes. These changes can be identified with the help of the allele tracks plotted using the SNP probes on the array. Copy neutral changes identified include isodisomy for entire chromosome 7 in a patient with clinical history and findings suggestive of Russell- Silver Syndrome as well as mosaic segmental isodisomy for the short arm of chromosome 11 in a pediatric patient with hepatoblastoma. Another copy neutral change detected is identity by descent (IBD). Regions of long contiguous stretches of homozygosity (LCSH) identified in such patients can be used to delineate genes for identification of causative mutations. As an example, to identify causative mutations in a patient with deafness, we are currently sequencing CDH23 gene which was narrowed down by LCSH regions present as a result of consanguinity. We thus demonstrate the use of SNP arrays in providing incremental clinically significant information in our patients. Based on our results, we emphasize the importance of a SNP microarray analysis in pediatric patients and underscore its utility in the clinical setting.

978T

Non-mosaic duplication of chromosome 6p in a child, inherited from his mother with a mosaic unbalanced der(14)t(6;14)(6pter→6p22.3::14p13→14qter). *S.M. Bain[1], J.K. Nicholl[1], W. Waters[1], A.H. Attwood[1], J. Samuel[1], T.A. Hocking[1], S.G. Brown[1], E.A. Haan[2], S. Yu[1]*. 1) Cytogenetics, SA Pathology, Women's & Children's Hospital, North Adelaide, South Australia, Australia; 2) South Australian Clinical Genetics Service, SA Pathology, Women's & Children's Hospital, North Adelaide, South Australia, Australia.

A female patient was referred at 10 weeks of pregnancy. Her clinical features included mild intellectual disability, short stature, abnormalities of chest and finger shape and past history of ventricular septal defect. Chromosome analysis of venous blood showed a mosaic karyotype, 46,XX, add(14)(p12)[9]/46,XX[111]. Subtelomere MLPA and locus specific FISH failed to identify the origin of the additional euchromatin due to the low level of mosaicism. Prenatal diagnosis was recommended but was declined. The baby had normal chromosomes at birth. The mother's next child had cytogenetic testing soon after birth. He was born at 33 weeks gestation with weight and length on the 10th centile, dysmorphic facial features and minor hypospadias and had necrotising enterocolitis in the neonatal period. Both array CGH and routine cytogenetics were initiated. Cytogenetic studies showed a non-mosaic karyotype 46,XY, add(14)(p12)mat and array CGH documented a terminal 15.9 Mb gain of 6p22.3-p25.3. The add(14) could now be described as der(14)t(6;14)(6pter→6p22.3::14p13→14qter) mat. Once the origin of the additional euchromatin on 14p had been identified, two different microarray platforms were used to assess the mosaicism that had been detected in the mother by cytogenetic methods (7.5% mosaicism). A BlueGnome oligo ISCA (4080-5) 8×60K array using BlueFuse software and an Illumina HumanCytoSNP-12 BeadChip using KaryoStudio software were used to analyse the mother's original DNA sample. Both platforms failed to call the low level mosaicism identified by chromosome analysis. Array CGH has become the standard approach in our laboratory for the investigation of intellectual disability, autism, dysmorphic features and malformations and would have detected the chromosome abnormality in the affected boy. It would not have detected the mother's mosaicism and her clinical features would have gone unexplained. Only empiric risk counselling could have been offered prior to her first pregnancy and an opportunity to prevent the birth of her more severely affected son would have been missed. Nonetheless, array CGH will remain the first line investigation in spite of its well known limits for detection of mosaicism. As in other situations, the clinician needs to consider other lines of investigation when faced with a patient with clinical features suggestive of a chromosome abnormality.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

979W

Diagnostic array analysis and the necessity of high quality SNP information. *R. Pfundt, J. Hehir-Kwa, N. Leijsten, S. van Gessel, M. Wunderink, T. Machielsen, R. van Beek, M. Banning, M. Rosario, N. de Leeuw.* Human Genetics, Radboud University Nijmegen Medical Centre, Nijmegen, Netherlands.

Genetic centres worldwide, have adopted the microarray technology as the replacement of routine chromosome analysis by karyotyping, for the detection of genomic copy number variation (CNV). In our genome diagnostics laboratory, array analysis was implemented in 2003, initially using "homemade" BAC arrays, that were replaced by commercially available SNP arrays in 2007 (Affymetrix 250K NspI SNP arrays). To date over 9000 samples have been analysed in our diagnostic array facility. SNP based microarrays not only provide reliable data for CNV analysis, but also generate a genotyping profile for every individual that is hybridized. Here we present our diagnostic workflow in which we have optimized the use of this genotype information to both enhance the quality assurance of this diagnostic test and to increase the diagnostic yield of the array analysis. For the quality assurance of the array test, genotyping data of patient and parental samples is used to check sample consistency in patient-parent duo and/or trio combinations. By doing so, we can check for sample mix-ups, potential non-paternity and determine the parental origin of de novo CNVs. In addition to this we have increased our diagnostic yield with >1 % by analysing the genotyping data for homozygous stretches that has led to the identification of >25 cases of uniparental disomy (UPD), and has pinpointed the location of pathogenic mutations in disease genes that were located in stretches of homozygosity identified in the array test. In 2011 we switched to the CytoscanHD array of Affymetrix that harbours over 2.6 million array probes of which ~750,000 generate high quality SNP calls that can be used to calculate allele ratios. This high quality SNP data has led to a further increase in diagnostic yield by revealing not only stretches of homozygosity but also more subtle allelic imbalances caused by e.g. mitotic recombination leading to stretches of mosaic UPD and (low level) mosaic CNVs. We conclude that high quality and high resolution genotyping information from SNP based array platforms is essential for a practical diagnostic array workflow that maximises the diagnostic yield of a single array test.

980T

Copy-number variation and UPD(16) mosaicism in a patient with apple peel intestinal atresia, ocular anomalies, microcephaly and bone-marrow failure. *F.A.T. de Vries[1], Y. van Bever[1], I.M. Appel[2], H. Kart[1], L.J.C.M. van Zutven[1], H.B. Beverloo[1].* 1) Clinical Genetics, Erasmus Medical Centre, Rotterdam, Netherlands; 2) Pediatrics Div. Oncology/Hematology, Erasmus Medical Centre- Sophia Children's Hospital, Rotterdam, Netherlands.

We report a 6-year-old female patient born from healthy, non-consanguineous parents. Amniocentesis and cytogenetic studies following ultrasound detection of a double bubble showed a 46,XX karyotype. The proband was born at 34+3 weeks of gestation. Postpartum she had a large hematoma at the left occipito-temporal region, probably due to forceps assisted extraction. The double bubble was caused by apple peel jejunal atresia, for which she had surgery. There was a slight neonatal thrombopenia explained by the hematoma and surgery. The patient has mild facial dysmorphisms, microcephaly, brittle hair, microphthalmia and major visual impairment due to corneal clouding. These ocular anomalies are associated with Axenfeld-Rieger malformation/Peters anomaly. Motor and mental development at 2 years were estimated to be within normal limits. Analysis of *PITX2, FOXC1, PAX6* and *MYCN* genes, implicated in Axenfeld-Rieger/Peters anomaly showed no mutations. Additional MLPA (MRC-Holland) analysis of these genes and all subtelomeric regions showed no deletion/duplication. (van Bever et al., 2008) At the age of ~6 years, the child developed pancytopenia and mild aplastic anemia. Hemato-cytologic or immunophenotypic bone-marrow examinations exposed no signs of hematologic malignancies or a myelodysplastic syndrome. Cytogenetic analysis of the bone-marrow showed a normal female karyotype. Fanconi anemia was excluded in a chromosome breakage test on both blood and fibroblasts. Because of the presence of congenital abnormalities, SNP-array analysis (Illumina HumanCytoSNP-12) was performed and identified 3 maternally inherited copy-number variations: a 2,5 Mb loss in chromosome 2q13 (published as a risk factor for developmental delay and dysmorphism), an 0.2 Mb loss in 5q21.1 (unknown clinical significance) and a 1.6 Mb gain in Xp22.31 (published as causative of intellectual disability or a benign variant). Moreover an aberrant B-allele frequency (BAF) pattern was observed for chromosome 16q. LogR results of this region were normal and a (mosaic) deletion/duplication was excluded by FISH. The aberrant BAF suggested mosaic uniparental disomy for 16q. Further investigations showed that this partial upd(16) resulted from mosaic loss of the maternal allele. Since the upd(16q)pat was not observed in DNA sampled in the first year of life, this acquired upd may be linked to the bone-marrow failure. This is currently under investigation.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

981W

Tumor suppression in *Apc*$^{Min/+}$ mice by the prostaglandin D_2 receptor PTGDR. *H. Lin[1], A. Kwong[1], O. Lee[1], V. Buslon[2], S. French[2], E. Salido[3].* 1) Division of Medical Genetics, Harbor-UCLA Medical Center, Torrance, CA; 2) Department of Pathology, Harbor-UCLA Medical Center, Torrance, CA; 3) CIBERER, Hospital Universitario Canarias, La Laguna, Tenerife, Spain.

Prostaglandin D_2 may prevent intestinal tumors, as shown by more adenomas in *Apc*$^{Min/+}$ mice that lack hematopoietic prostaglandin D synthase (HPGDS) and fewer tumors in *Apc*$^{Min/+}$ mice that over-express this enzyme (Park et al. Cancer Res 2007; 67: 881-9). Knockout of PGD_2 receptor PTGDR also led to more tumors, in our earlier work. Here, we expand data on *Ptgdr*, with *Apc*$^{Min/+}$ mice examined at 6 weeks (50 mice, including 13 controls) and 14 weeks (64 mice, including 22 controls). Intestines were removed, opened, fixed, coiled into Swiss rolls, embedded, sectioned, and mounted on slides. To score adenomas, we used 10 Swiss roll sections spaced 250 Δm apart. Small tumors were those seen on only 1 section, whereas large tumors were those seen on multiple sections. Heterozygous *Ptgdr* knockouts did not increase the median number of total tumors in *Apc*$^{Min/+}$ mice at 14 weeks (84 tumors versus 71; P = 0.16). However, the median number of large tumors was higher with heterozygous *Ptgdr* knockouts (33 versus 24; P = 0.023). Homozygous *Ptgdr* knockouts raised median numbers of total tumors (100 versus 71; P = 0.0060) and large tumors (38 versus 24; P = 0.0040) at 14 weeks. At 6 weeks, median numbers of total and small tumors were higher for both heterozygous and homozygous *Ptgdr* knockout mice (e.g., for homozygous knockouts versus controls — 65 versus 43 for total tumors, P = 0.0031; 58 versus 37 for small tumors, P = 0.0025). Numbers of colon tumors also increased (0–6 for homozygous knockouts versus 0–1 for controls; P = 0.015). We examined younger *Apc*$^{Min/+}$ mice (one mouse at 1 week; 6 mice at 3 weeks), for data on when tumors appear. We saw tumors by 3 weeks. Roughly 60% of tumors that will occur over 14 weeks arise during the first 6 weeks — when tumors form 1.5-fold faster in *Ptgdr* knockout mice, compared to controls. After 6 weeks, tumors occur at similar rates in knockout and control mice, perhaps allowing controls to "catch up."

Mice are born without intestinal crypts, which form by rapid expansion of the crypt stem cell pool, followed by a switch to production of non-stem cells (Itzkovitz et al. Cell 2012; 148: 608-19). More tumors in *Ptgdr* knockout mice in the first 6 weeks of life support an interpretation that PTGDR slows proliferation of crypt stem cells. PTGDR agonists may be an approach for preventing tumors in familial adenomatous polyposis.

982T

Effects of Western-style diet and genetic predisposition on the protein expression and cancer susceptibility in normal colonic mucosa. *S. Lowndes[1], D. Dermadi Bebek[1], M. Pussila[1], N. Reyhani[2], L. Sarantaus[1], M. Nyström[1].* 1) Department of Biosciences, University of Helsinki, Finland; 2) Aalto University, School of Science, Helsinki, Finland.

Lifestyle and diet have a major effect on the development of colorectal cancer (CRC). Dietary habits of Western populations in particular are recognized as a risk factor for CRC. Most probably unhealthy features in Western diet (WD) induce expression changes in key regulatory pathways that affect normal metabolic processes in colonic mucosa. Since cancer development needs accumulation of several genomic changes, the individuals, who already have the inherited susceptibility, may be extremely sensitive for diet effects. Indeed, disease phenotypes such as age of onset and tumor spectrum vary considerably between mutation carriers even in a same CRC family suggesting implication of lifestyle factors on risk accumulation. Here, by using a mouse model for intestinal cancer we studied whether and how WD and a genetic predisposition may separately and in interaction change the protein expression and thus increase a cancer risk in normal colonic mucosa.

Mlh1$^{+/-}$ mice were used as a model for Lynch syndrome (LS), the most common CRC syndrome, and *Mlh1*$^{+/+}$ wild type littermates were used as healthy controls. After a 12 months feeding experiment the effects of WD compared to effects of a control diet (AIN-93G) in different mouse genotypes were studied using 2D-DIGE proteomics and differently expressed proteins were identified by using mass spectrometry. Our results suggest that individuals with genetic predisposition are more susceptible to diet effects and that together they may increase the risk for CRC. Altogether 394 significant protein expression changes were detected in *Mlh1*$^{+/-}$ WD group when compared to control diet group. In *Mlh1*$^{+/-}$ mice 19 out of the 394 changes were related to the effect of diet or combination of diet and genotype. Identification of these proteins revealed substantial changes in energy metabolism linked to fatty acid catabolism, glycolysis, pentose phosphate pathway, TCA cycle and oxidative phosphorylation. Furthermore, an increased number of neoplasia/hyperplasia together with signs of inflammation, deregulation of apoptosis and higher rate of proliferation were detected in the normal colonic mucosa of *Mlh1*$^{+/-}$ mice fed with WD. Moreover, the results indicate that colonic mucosa cells exposed to WD are under higher proliferative state than cells exposed to healthy control diet, suggesting the diet itself to be a considerable risk factor.

983F

The Jackson Laboratory Rare and Orphan Disease Center. *S. Rockwood, C. Lutz, K. Fancher, A. Picard, M. Sasner, L. Reinholdt, S. Murray, D. Bergstrom, J. Morgan, L. Donahue, The GRS Team.* Genetic Resource Science, The Jackson Laboratory, Bar Harbor, ME.

In the United States, diseases are categorized as rare or orphan if the number of afflicted patients numbers 200,000 or less.About 7000 such diseases have been described, with another ~250 described each year.Historically, this subset of diseases has been associated with significant unmet medical need.Government incentives (direct grants, tax credits, market exclusivity) put in place over the last several decades have contributed to a growing interest in the orphan disease drug market. Importantly, there is growing recognition by pharmaceutical and biotech companies that orphan drug development is an economically viable activity.To further facilitate these positive developments, The Jackson Laboratory Rare and Orphan Disease Center was recently launched.The Center focuses on partnering with scientists, foundations, and other experts around the world to enable the development, standardization, optimization, and rapid distribution of preclinical mouse models for drug discovery. Being able to offer the resources and expertise to enable the design, construction and management of preclinical mouse models of disease, in combination with a global delivery system, expertise in technical transfer issues and genetic quality control, uniquely positions the Center to put new tools into the hands of scientists and thereby accelerate drug discovery.Complementing the Center's model development pipeline is one of the most comprehensive mouse repositories available, consisting of a growing collection of over 6000 distinct mouse lines.The wide array of easily accessible, well-characterized tool and mutant allele strains serves as an unmatched companion resource for the building of novel disease models with applications in translational research. Contributing your disease model enables researchers across the globe to have greater access to tools for drug efficacy testing and discovery.If you would like to donate your mouse strain to the Jackson Laboratory Mouse Repository, please see: **www.jax.org/donate-a-mouse**. Ongoing rare disease research projects involve partnerships with foundations, companies and principle investigators.Our emphasis is on the development, standardization and distribution of validated preclinical mouse models for drug discovery.We are interested in partnering with additional groups across all therapeutic areas of research. To learn more about The Jackson Laboratory Rare and Orphan Disease Center, please visit our website at **www.jax.org/rare.**.

984W

Genetic differences in transcript responses to low-dose ionizing radiation identify tissue functions associated with breast cancer susceptibility. *A.J. Wyrobek, F. Marchetti, S. Bhatnagar, N. Duru, J. Han, Z. Hu, J-H. Mao, J.W. Gray, A.M. Snijders.* Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA.

High dose ionizing radiation (IR) is a well-known risk factor for breast cancer but the health effects after low-dose (LD, <10cGy) exposures remain highly uncertain. We explored a systems approach that compared LD-induced chromosome damage and transcriptional responses in strains of mice with genetic differences in their sensitivity to radiation-induced mammary cancer (BALB/c and C57BL/6) for the purpose of identifying mechanisms of mammary cancer susceptibility. Unirradiated mammary and blood tissues of these strains differed significantly in baseline expressions of DNA repair, tumor suppressor, and stress response genes. LD exposures of 7.5cGy (weekly for 4 weeks) did not induce detectable genomic instability in either strain. However, the mammary glands of the sensitive strain showed early transcriptional responses involving: (a) diminished immune response, (b) increased cellular stress, (c) altered TGFβ-signaling, and (d) inappropriate expression of developmental genes. One month after LD exposure, the two strains showed opposing responses in transcriptional signatures linked to proliferation, senescence, and microenvironment functions. We also discovered a pre-exposure expression signature that is predictive for poor survival among human cancer patients (p=0.0001), and a post-LD-exposure signature that is predictive for patient survival (p<0.0001) with concordant expression in the LD-exposed sensitive strain, in human DCIS and in human breast tumors. Our findings support the hypothesis that mechanisms that control susceptibility to low-dose radiation induced mammary cancer in mice are similar to those that determine poor survival in breast cancer patients. We observed non-linearity of the LD responses providing molecular evidence against the LNT risk model and obtained new evidence that LD responses are strongly influenced by genotype. Our findings suggest that the biological assumptions concerning the mechanisms by which LD radiation is translated into breast cancer risk should be reexamined and suggest a new strategy to identify genetic features that predispose or protect individuals from LD-induced breast cancer.[Supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02–05CH11231 and by LBNL-LDRD].

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

985T

Association of the *HLA-G* gene polymorphisms with cervical cancer risk. *T. Chang[1], Y. Yang[2,4], Y. Lee[1,3,5], T. Chen[2], S. Chang[1], W. Lin[1].* 1) Med Res Dept, Mackay Memorial Hosp, New Taipei, Taiwan; 2) Gynecology and Obstetrics, Mackay Memorial Hosp, Taipei, Taiwan; 3) Pediatrics, Mackay Memorial Hosp, Taipei, Taiwan; 4) Gynecology and Obstetrics, Taipei Medical University, Taipei, Taiwan; 5) Pediatrics, Taipei Medical University, Taipei, Taiwan.

Cervical cancer is a major cause of cancer-related death in women worldwide, in spite of extensive efforts dedicated to its preventive or treatment managements. While cervical infection with high-risk types of human papillomaviruses (HPVs) is critical to cervical carcinogenesis, infection alone is not sufficient to cause cervical cancer. It is now recognized that host immunogenetic background plays an important role in the control of HPV infection and the development of cervical cancer. HLA-G is a nonclassical HLA class I molecule with multiple immunoregulatory properties. It has been shown to inhibit the cytotoxic activity of both natural killer and cytotoxic T cells. The aim of this study is to investigate the association between *HLA-G* gene single nucleotide polymorphisms (SNPs) and the risk of cervical cancer in the Taiwanese women. We genotyped 2 putatively functional SNPs, an insertion/deletion polymorphism of 14bp (rs1704) and a missense polymorphism (rs12722477 *C/A*), in 342 cervical squamous cell carcinoma (CSCC) patients and 344 age/sex matched healthy controls by using PCR and the Pre-Developed TaqMan Allelic Discrimination Assay. The presence and genotypes of HPV in CSCC patients were determined by PCR. Our findings show significant increases of rs12722477 *C/C* genotype and *C* allele frequencies among patients with CSCC as compared with controls (OR = 1.65, 95% CI 1.13–2.41, $P = 0.02$, $P_c = 0.04$; OR = 1.54, 95% CI 1.14–2.06, $P = 0.004$, $P_c = 0.008$, respectively). In addition, significant associations of *C/C* genotype and *C* allele frequencies were observed between HPV-16 positive CSCC patients and controls (OR = 2.29, 95% CI 1.14–3.66, $P = 4.0 \times 10^{-4}$, $P_c = 8.0 \times 10^{-4}$; OR = 2.18, 95% CI 1.48–3.22, $P = 6.62 \times 10^{-5}$, $P_c = 1.32 \times 10^{-4}$, respectively). Haplotype analysis revealed the distribution of haplotype Del-C was significantly increased in women with CSCC and HPV-16 positive CSCC (OR = 1.51, 95% CI 1.15–1.98, $P = 0.003$, $P_c = 0.012$; OR = 1.60, 95% CI 1.16–2.21, $P = 0.004$, $P_c = 0.016$, respectively). In conclusion, our results suggest that rs12722477 *C/A* polymorphism of the *HLA-G* gene conferred susceptibility to CSCC in the Taiwanese population.

986F

Association of the HOXB13 G84E Variant with Familial, Sporadic, and Aggressive Prostate Cancer: the Mayo Clinic Experience. *M.S. DeRycke[1], L.S. Tillmans[2], S.K. McDonnell[3], S.M. Riska[3], J. Xu[4], Y.W. Asmann[3], D.J. Schaid[3], S.N. Thibodeau[2], International Consortium for Prostate Cancer Genetics.* 1) Department of Health Sciences Research, Mayo Clinic, Rochester, MN; 2) Department of Laboratory Medicine & Pathology, Mayo Clinic, Rochester, MN; 3) Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, MN, USA; 4) Center for Cancer Genomics, Wake Forest School of Medicine, Winston-Salem, NC.

Recently a rare, germline mutation in HOXB13 (G84E) was reported to be associated with increased risk of familial prostate cancer (PC). Increased risk was highest among men diagnosed at a young age and with a strong family history of disease. In this study, we genotyped men with familial (n=467), sporadic (n=439) and aggressive (n=196) PC, and in population controls (n=494) for the G84E variant to determine if this variant is also associated with increased risk of PC in sporadic and aggressive PC, as well as familial PC. Genotyping was performed using the MassARRAY iPLEX and genotype frequencies were compared between sporadic or aggressive PC cases versus controls using the Armitage test for trend in proportions. Familial cases were compared to a combined set of 63 unaffected relatives and the population controls (n=557), adjusting for the correlation among related subjects. The Odds Ratio (OR) for all analyses ranges from 2.27–5.17, consistent with the original report for the G84E/HOXB13 association. Although the MAF is small, the G84E variant was present statistically significantly more often in individuals with aggressive disease compared to controls, resulting in an OR of 5.17 (95% CI, 1.28–20.88). A similar magnitude of risk was found for familial PC cases versus control (OR=4.04; 95% CI, 1.39–11.72), yet not as large a risk for sporadic cases versus controls (OR=2.27; 95% CI, 0.56–9.12). These results suggest that the G84E/HOXB13 does have implication for the general population; additional, larger studies will be required to confirm this finding.

987W

The miR-29 family of microRNA genes and let-7i associate with gastric cancer in European populations: Results from the EPIC-EURGAST study. *Y. Espinosa-Parrilla[1], X. Muñoz[2], C. Bonet[2], N. García[2], A. Venceslá[2], E. Riboli[3], C.A. González[2], N. Sala[2] on behalf of the EPIC gastric cancer working group.* 1) Ciencies Experimentals i de la Salut (CEXS-UPF), Institut de Biologia Evolutiva (UPF-CSIC), Barcelona. Spain; 2) Unit of Nutrition, Environment and Cancer, Catalan Institute of Oncology-IDIBELL, Hospitalet de LLobregat, Barcelona. Spain; 3) Department of Epidemiology and Public Health, Imperial College London, London. United Kingdom.

MicroRNAs (miRNAs) are recognized post-transcriptional gene repressors involved in the control of almost every biological process including tumorigenesis. Even though deregulation of miRNA expression has been associated with various human cancers, the contribution of these regulators to the genetic susceptibility to cancer is still unclear. We performed association studies of miRNA genes with gastric cancer (GC) risk in samples from the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort. From 719 human miRNAs (release 13.0, miRBase) we selected 90 miRNAs located in 39 miRNA regions (25 miRNA clusters and 14 isolated miRNAs) that were either involved in various types of cancer or predicted to regulate candidate genes for GC (TargeScan and miRanda predictions). A panel of 156 tagSNPs was designed and genotyped in 365 incident gastric adenocarcinoma cases and 1284 matched controls from a nested case-control study in the EPIC cohort. Regarding the histological type of tumours, 35% were intestinal and 35% diffuse. After quality control analysis and data filtering, association was analyzed by unconditional logistic regression, adjusting for age, sex and country. Six SNPs in a miR-29 cluster region were associated with the risk of diffuse gastric cancer (Best hit: OR=1.72, 95%CI: 1.30–2.28; pvalue= 0.0002) and three SNPs tagging let-7i were associated with increase risk of general gastric cancer (Best hit: OR=1.69, 95%CI: 1.25–2.29; pvalue= 0.0008). Evaluation of target genes for let-7i and miR-29 using the miRanda and TargetScan algorithms revealed that at least 17 and 13 target genes for miR-29 and let-7i, respectively, are involved in cancer: Among these there are *VEGFA* and the DNA methyltransferases *DNMT3A* and *DNMT3B* target genes for miR-29 and *TP53* for let-7i. Interestingly, SNPs in *VEGFA* and in *DNMT3B* showed a significant interaction with miR-29 SNPs. Furthermore, miR-29 has been shown to up-regulate p53 and induce apoptosis in a p53-dependent manner. Sequencing of miR-29 and of let-7i genes did not reveal any causal variant for the association. These results indicate that miR-29 and let-7i are associated with the genetic susceptibility to GC and that this association could be at least partially related to the regulation of p53 and its function as tumour suppressor. Supported by “LaCaixa” (BM06-130-0); Spanish Ministry of Health (PI070130 and PI081420) and Science (BFU2010-18477); European Commission FP7 (PIOF-GA-2009-236836).

988T

3' UTR Variation in *PTEN* and *PTENP1*: implications for microRNA binding. *B.G. Hernandez, C. LaViolette, T. Begay, J.A. Wilder.* Biological Sciences, Northern Arizona University, Flagstaff, AZ.

PTEN is a tumor suppressor gene that is associated with prostate cancer when downregulated. Dysregulation of the gene can be caused by several mechanisms, including microRNA interactions with the mature transcript. *PTENP1* is a processed pseudogene of *PTEN* that shares a similar 3' untranslated region (UTR). Due to this similarity *PTENP1* acts as a sink for microRNAS that may otherwise bind to and downregulate *PTEN*. In this project, we aim to characterize single nucleotide polymorphisms (SNPs) in the 3' UTRs of both *PTEN* and *PTENP1* and to assess *in silico* the effect of these SNPs on microRNA binding. We surveyed the 3' UTRs of *PTEN* and *PTENP1* in 41 Native American genomic DNA samples and publicly available polymorphism data. Several SNP loci are predicted to affect microRNA binding at both *PTEN* and *PTENP1*. Variability at these sites has the potential to affect *PTEN* expression, both directly and indirectly, and therefore may affect cancer progression.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

989F

Replacement therapy for lung adenocarcinoma risk in never-smoking women. *C. Hsiao[1], K. Chen[2], G. Chang[3], Y. Tsai[4], W. Su[5], Y. Chen[6], M. Huang[7], C. Hsiung[1], C. Chen[8], P. Yang[2], GELAC Study Group.* 1) Biostatistics & Bioinformatics, Natl Health Research Inst, Miaoli County, Taiwan; 2) Division of Pulmonary Medicine, Department of Internal Medicine, National Taiwan University Hospital and College of Medicine, Taipei, Taiwan; 3) Division of Chest Medicine, Department of Internal Medicine, Taichung Veterans General Hospital, Taichung; Department of Medicine, School of Medicine, National Yang-Ming University, Taipei, Taiwan; 4) Division of Pulmonary and Critical Care Medicine, Department of Internal Medicine, Chang Gung Memorial Hospital, Taipei, Taiwan; 5) Division of Hematology/Oncology, Department of Internal Medicine, National Cheng Kung University, Tainan, Taiwan; 6) Chest Department, Taipei Veterans General Hospital; School of Medicine, National Yang-Ming University, Taipei, Taiwan; 7) Department of Internal Medicine, Kaohsiung Medical University Hospital; School of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan; 8) Genomics Research Center, Academia Sinica; and Graduate Institute of Epidemiology, College of Public Health, National Taiwan University, Taipei, Taiwan.

EGFR single nucleotide polymorphisms(SNPs) and hormone replacement therapy (HRT) have been demonstrated to be risk factors for lung carcinogenesis. This study is to investigate the interactive effect of EGFR SNPs and HRT on the lung adenocarcinoma risk in never-smoking women. This case-control study included 532 never-smoking female lung adenocarcinoma patients and 532 controls. All the subjects are Taiwanese. The EGFR SNPs were retrieved from a genome-wide association study. The associations of EGFR SNPs with the lung adenocarcinoma risk were estimated by multivariate-adjusted logistic regression. The Haploview program was used to select tagged EGFR SNPs interacted with HRT and to construct haplotype blocks. The Benjamini and Hochberg method was used to reduce the multiple testing effect. Among 84 EGFR SNPs retrieved, 11 tagging SNPs showed a potentially interaction with HRT use and lung adenocarcinoma risk. Most SNPs were located near the EGFR tyrosine kinase domain. Eight of the tagged SNPs were in two haplotype blocks, block 1 and 2. The interactions for HRT use and numbers of high-risk EGFR SNPs on lung adenocarcinoma risk in block 1 and 2 are significant (p for interaction =0.0002, and 0.0032 respectively). A trend of increased lung adenocarcinoma risk was found in subjects with HRT harboring an increasing number of at-risk EGFR SNP genotypes in block 1 and 2 (p=0.0008 and 0.0023 respectively). HRT use may modify the association of EGFR SNPs with lung adenocarcinoma risk. The EGFR SNPs have a cumulative effect on increasing the lung adenocarcinoma risk in never-smoking women with HRT use.

990W

Application of molecular diagnostic techniques on the study done on phosphatase and tensin homologue deleted on chromosome 10 (PTEN). *R. Issa.* imolecular diagnostics, institute of genetic engineering, cairo, Egypt.

Background:PTEN is one of the tumor suppressor genes involved in secondary cancer progression prevention.In egypt,prostate adenocarcinoma,recorded the highest incidence and mortality among cancers of the genito-urinary system. Aim:The aim of this stuyis to detect the presence of mutations in the tumor suppressor PTEN gene in begnin prostate hyperplasia patients,and prostate cancer patients,and corrolate between PTEN and the pathogenesis of both patient groups. Patients and methods:Twenty begnin and malignant patients were recruited,as well as seven apparently normal subjects taken as control.Whole blood samples were collected from diagnosed patients using TRUS-biopsies.Control subjects were free of chronic disorders and tumors,yet some had family history of begin and malignant tumors,yet some had family history of begnin and malignant tumors.Studied group had age range between 54 and 77 years,and average PSA level of 2,3,19.9 and 35 ng/ml for control,BPH and PCa groups respectively.PCR were performed using intronic primers,specific for exons 5 and 7,to map possible mutations(for positive PCR reactions sequensing done using forward and reverse primers) Results:Analyzing mutations mapped for the selected studied groups proved the likeliness of correlating between both BPH and PCa patients.PTEN association in PCa patients was corrolated with their pathology finfings and clinical picture.One patient out of ten showed complete deletion and the highest Gleason score(9)and worst clinical picture,two out of ten showed partial deletion, (one losing the phosphatase function and the other losing the C2 domain of the direct protein binding).The remaining results showed the high occuring genotypes.All 5 potentially significant mutations (g.74482dupA,g,74818dupG,g.74845dupC and g.99506C G,g.99582dupA) shared the presence of the highly occurring genotypes except g.99506cG of exon 7 begnin patient(7b6) Conclusion:Mutations mapped have revealed some potentially significant mutant variants of PTEN pointing towards its role in both BPH and PCa pathology and the possibility of a bridging correlation between both begnin and malignant proatate tumors,.

991T

Validation, fine-mapping, and characterization of six genes related to testicular germ cell tumor in a large Swedish-Norwegian case-parent, case-control sample. *R. Karlsson[1], W. Kristiansen[2], K.E. Andreassen[3], E.L. Aschim[2], R.M. Bremnes[4], O. Dahl[5], S.D. Fossa[6,7], O. Klepp[8], C.W. Langberg[9], A. Solberg[10], S. Tretli[3], P.K.E. Magnusson[1], H.-O. Adami[1,11], T.B. Haugen[2], T. Grotmol[3], F. Wiklund[1].* 1) Department of Medical Epidemiology and Biostatistics, Karolinska institutet, Stockholm, Sweden; 2) Faculty of Health Sciences, Oslo and Akershus University College of Applied Sciences, Oslo, Norway; 3) Department of Etiologic Research, Cancer Registry of Norway, Oslo, Norway; 4) Translational Cancer Research Group, Institute of Clinical Medicine, University of Tromsø and Department of Oncology, University Hospital North Norway, Tromsø, Norway; 5) Section of Oncology, Institute of Medicine, University of Bergen and Department of Oncology, Haukeland University Hospital, Bergen, Norway; 6) Department of Clinical Cancer Research, The Norwegian Radium Hospital, Oslo University Hospital, Oslo, Norway; 7) Faculty of Medicine, The Norwegian Radium Hospital, University of Oslo, Oslo, Norway; 8) Department of Oncology, Ålesund Hospital, Helse Sunnmøre HF, Ålesund, Norway; 9) Cancer Centre, Ullevål University Hospital, Oslo, Norway; 10) Department of Oncology, St Olavs University Hospital, Trondheim, Norway; 11) Department of Epidemiology, Harvard School of Public Health, Boston, MA, USA.

Purpose: Testicular germ cell tumors (TGCT) are rare, but the most common malignancy in young men. Genome-wide association studies (GWAS) have identified single nucleotide polymorphisms (SNPs) associated with TGCT risk in the genes *ATF7IP*, *BAK1*, *DMRT1*, *KITLG*, *SPRY4*, and *TERT*. The present study aimed to validate and characterize these associations, and to assess effect modification by parental sex and histological subtype.

Methods: We invited TGCT cases and parents identified from Swedish and Norwegian national registries to the study. 2162 cases consented, distributed as 1004 case-parent triads, 440 dyads, and 718 singletons. 1135 TGCTs were seminomas, and 935 non-seminomas or mixed tumors. 118 SNPs tagging common variation in the genes were genotyped in saliva-derived DNA on Sequenom MassARRAY iPLEX Gold chemistry. 3922 unrelated controls were included from TWINGENE, a population-based sample of Swedish twins. TWINGENE DNA was blood-derived, and genotyped on the Illumina OmniExpress platform. Additional SNPs were imputed using BEAGLE software and reference haplotypes from the 1000 genomes project. SNP-TGCT association was tested in the combined sample using UNPHASED software. Forward stepwise regression within each gene was applied to find independent association signals. Effect modification was tested for the strongest associations in each gene.

Results: 831 triads, 474 dyads, 712 singletons, 3922 controls, and 98 SNPs passed sample and genotyping quality control. After imputation 852 SNPs were analyzed. All genes had significant associations with TGCT (lowest unadjusted P: *ATF7IP* 6.2e-06; *BAK1* 2.1e-10; *DMRT1* 6.7e-25; *KITLG* 2.1e-48; *SPRY4* 1.4e-29; *TERT* 1.8e-18). Most genes had more than one independent association, and one signal in *TERT* has not been reported previously to our knowledge. A significant parent-of-origin effect was observed for rs10463352 in *SPRY4* (interaction P=0.0008, OR_{mat}=1.56, OR_{pat}=1.03). No significant effect modification by histological subtype was found.

Conclusions: We validated previous associations, and may have found a new, independent association in *TERT*. Assessment of parent-of-origin effects indicated a *SPRY4* SNP as only associated with TGCT when inherited maternally.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

992F

Genome-wide germline DNA autosomal copy number variation analysis in Finnish *BRCA1/2*-founder mutation-negative hereditary breast and/or ovarian cancer individuals. *KM. Kuusisto[1], O. Akinrinade[1], M. Vihinen[1,2], S-L. Sallinen[3], J. Schleutker[1,4]*. 1) Institute of Biomedical Technology / BioMediTech, University of Tampere and Fimlab Laboratories, Tampere, Finland; 2) Department of Experimental Medical Science, University of Lund, Lund, Sweden; 3) Department of Pediatrics, Genetic Outpatient Clinic, Tampere University Hospital, Tampere, Finland; 4) Department of Medical Biochemistry and Genetics, University of Turku, Turku, Finland.

Introduction: Genetic factors account for 5–10% of all breast cancer (BrCa) cases. Several BrCa susceptibility genes have been identified and the two most important ones are *BRCA1* and *BRCA2* being responsible for approximately 20% of hereditary breast and/or ovarian cancer (HBOC) cases in Finland. However, genetic factors underlying majority of the HBOC cases remain unexplained and this makes genetic counseling challenging. Recently, copy number variations (CNVs) have been discovered to play an important role in diseases. Germline CNVs have been implicated to predispose many types of cancers including neuroblastoma, colorectal, and prostate. We aimed to investigate germline DNA autosomal CNVs in HBOC individuals in the Finnish population and to get new clues to genetic counceling. **Methods:** Whole genome SNP genotyping analysis by Illumina HumanCytoSNP-12 BeadChip was performed for 81 well-characterized, hereditary breast and/or ovarian cancer *BRCA1/2*-founder mutation-negative individuals and 35 healthy controls in a Finnish population to identify breast/ovarian cancer-associated CNV regions. Four algorithms, PennCNV, cnvPartition, QuantiSNP, and CNstream, were applied for the analysis. CNV carrier frequencies between 81 HBOC individuals and 35 controls were compared by using the Fisher's exact/chi-square test. Gene Ontology (GO) analysis, the Database of Genetic Variants search and the Online Mendelian Inheritance in Man Database search were performed to analyze the importance of the identified CNV affected genes. Four CNV regions were further validated by quantitative RT-PCR. **Results:** A total number of identified CNV regions in HBOC individuals and controls (n=116) ranged from 211 to 404 by different algorithms and the majority of the CNV regions sized from 10 to 50 kb. The most frequent CNV identified by four algorithms was a 14.6 kb loss at 3p11.1 observed in 9.9% (8/81) of the HBOC individuals but not in controls. Other frequent CNVs observed in HBOC individual included losses at 1q21.1, 2q35, and 5q15, and gain at 19q13.41. Several CNV affected genes have been previously associated with human disorders. **Conclusions:** This is the first study to analyze germline DNA autosomal CNVs in HBOC individuals in Finland. Our study provides new information of potential CNV regions that could be associated with the increased risk of HBOC susceptibility in Finland but the current results warrant validation in larger study group.

993W

***Mediator complex subunit 12* (*MED12*) mutations in uterine leiomyosarcomas.** *M.M. McGuire[1], M. Jones[2], G. Trucco[2], U. Surti[1], R.P. Edwards[1], A. Rajkovic[1]*. 1) Department of Obstetrics, Gynecology and Reproductive Sciences, Magee-Womens Research Institute, Pittsburgh, PA; 2) Department of Pathology, Magee-Womens Hospital of UPMC, Pittsburgh, PA.

Uterine leiomyosarcomas (ULMS) arise from the smooth muscle cells of the uterine wall, the myometrium. These rare malignant tumors are often solitary and large (>10cm, diameter), and their behavior is highly variable. The annual incidence rate of ULMS is less than 1% worldwide, and they primarily occur in women 40 to 60 years of age around the time of menopause. The etiology of ULMS is largely unclear, and their rare nature has limited research on identifying possible risk factors. It is common for benign uterine leiomyomas (fibroids), which also arise from uterine smooth muscle tissue, and ULMS to be present in the same uterus, though it is debated whether uterine leiomyomas can transition into ULMS. Recent histopathologic data suggests that a subset of ULMS may derive from uterine leiomyomas. *Mediator complex subunit 12* (*MED12*), an X-linked gene which encodes a protein involved in transcriptional regulation of RNA polymerase II-dependent genes, was identified as being strongly associated with uterine leiomyoma development. Exon 2 of *MED12* harbors damaging somatic variants in 60–70% of uterine leiomyomas. Therefore, we examined exon 2 of *MED12* in genomic DNAs from 60 FFPE ULMS samples. A gynecologic pathologist confirmed the absence of contamination of ULMS samples with benign leiomyoma tissue. In this sample set, 6.7% (4/60) of ULMS samples harbored variants in exon 2 of *MED12*, and all variants were single heterozygous missense mutations confined to codon 44. These data support that a subset of ULMS harbor *MED12* mutations which are also found in uterine leiomyomas. Our results therefore support the hypothesis that a subset of ULMS may arise from uterine leiomyomas.

994T

Association of breast cancer risk factors and HLA-G gene InDel 14bp polimorphyms in a Brazilian population. *Y.CN. Muniz[1], M. Damiani[1], L.D. Hausmann[1], B.S. Almeida[1], B. Fernandes[2], R.T. Simões[3], A.R. Marrero[1], I.R. Souza[1]*. 1) LAPOGE, Federal University of Santa Catarina, Florianopolis, Santa Catarina, Brazil; 2) HU, Federal University of Santa Catarina, Florianopolis, Santa Catarina, Brazil; 3) Instituto de Ensino e Pesquisa da Santa Casa de Belo Horizonte - IEP/SCBH. Laboratório de Biologia Molecular e Biomarcadores.

Breast cancer is the second most common type of cancer in the world. The HLA-G protein is related to the escape of immuno response and was found at high levels in tumor cells of breast cancer patients. The insertion of the 3'UTR 14bp of HLA-G is related to an increased expression by promoting the development of tumor cells. The aim was to investigate the influence of polymorphism in the manifestation of breast cancer related to risk factors. The DNA was extracted from whole blood, genotyping by PCR carried out, viewed PAGE and stained with silver nitrate. We genotyped 198 patients and 200 controls (matched for age). In the comparison of risk factors were more patients with familial history of disease (22%) and smokers (16%) and lower proportion of patients who became pregnant (8%) when compared with control (13%, 7%, 15%, respectively). The presence of precoce menarche, late menopause and hormonal contraceptive use were similar between the two groups. The insertion frequency was 0.492 and 0.453 in the patient group in the control group, respectively. There was no statistical difference in the comparisons between the polymorphism and disease, as well as considering the risk factors. To clarify the relationship of the 3'UTR of the HLA-G with breast cancer and its risk factors, other markers related to mRNA stability will be further analyzed by the group (Financial support: PNPD CAPES, FAPESC, CNPq).

995F

Fine-mapping of prostate cancer susceptibility locus 11q13.5. *R. Nurminen[1], R. Lehtonen[2], T.L. Tammela[3], T. Wahlfors[1], J. Schleutker[1,4]*. 1) Institute of Biomedical Technology/BioMediTech, University of Tampere and Fimlab Laboratories, Tampere, Finland; 2) Department of Biosciences, University of Helsinki, Helsinki, Finland; 3) Department of Urology, University of Tampere and Tampere University Hospital, Tampere, Finland; 4) Department of Medical Biochemistry and Genetics, University of Turku, Turku, Finland.

Prostate cancer is a complex disease with estimated heritability of 19–40% but so far only few susceptibility genes have been identified. It would be especially important to identify genetic variants that predispose to aggressive disease in order to recognize those men at risk early enough and to develop treatment methods for life-threatening disease outcome. Finnish population is well suited for finding disease associated genetic variants because it is genetically homogeneous.

Chromosomal region 11q13 has been connected with prostate cancer in linkage and genome-wide association studies. Previously we screened *EMSY* at 11q13.5 in a Finnish case-control study in over 3000 samples. A rare, intronic variant IVS6–43A>G was observed to associate with familial disease and aggressive disease outcome in unselected prostate cancer patients. The functionality of the variant is unclear.

The aim of this study was to determine if IVS6–43A>G is in linkage disequilibrium (LD) with another variant(s) at 11q13.5 and to characterize the region further in relation with prostate cancer predisposition. The fine-mapping study included 2128 unselected Finnish prostate cancer patients and 906 controls. A total of 31 tagging single nucleotide polymorphisms (tagSNPs) covering 0.4 Mb at 11q13.5 were genotyped using TaqMan assay, KASPar assay, High Resolution Melt analysis and sequencing. Imputation was conducted with 1000 Genomes and HapMap3 reference panels using IMPUTE v2.2. The association of the imputed SNPs with prostate cancer was tested using SNPTEST v.2 and the association of the tagSNPs using PLINK v1.07. TagSNPs that deviated from Hardy-Weinberg equilibrium in controls were excluded from the analysis prior imputation. Haplotype blocks and LD structure were visualized and the haplotype association was calculated using Haploview v4.2.

The previously identified variant IVS6–43A>G was not observed to be in high LD with any of the tagSNPs or the imputed variants. However, the strongest association with prostate cancer was observed for a tagSNP in close approximation to IVS6–43A>G ($p = 0.0008$). In addition, a haplotype including the tagSNP allele was associated with prostate cancer ($p = 0.005$). The results further support the importance of 11q13.5 in prostate cancer predisposition.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

996W

Role of Copy Number Variants in the Nucleotide Excision Repair Pathway Genes and Breast Cancer Risk in Puerto Rican Women. *A. Pacheco Torres[1], J. Matta[2], J. Dutil[1]*. 1) Biochemistry, Ponce School of Medicine and Health Sciences, Ponce, PR; 2) Toxicology and Pharmacology, Ponce School of Medicine and Health Sciences, Ponce, PR.

Common polymorphisms in DNA repair genes can alter an individuals' DNA repair capacity and modify the effect of environmental exposures on cancer risk. Previous studies suggest that genetic variations in nucleotide excision repair (NER) genes may increase the risk to develop certain types of cancer, including breast cancer. However, most of these studies have been limited to single nucleotide polymorphisms (SNPs). Copy number variations (CNVs) are segments of DNA that are present at a variable number of copies in comparison to a reference genome. Together with SNPs, CNVs and may underlie heritability of complex diseases. Our objective is to determine whether CNVs in the genomic regions of the NER genes are involved in modulating DNA repair capacity (DRC) and breast cancer risk in Puerto Rican women. TaqMan CNV assays for the genes XPD/ERCC2, XPC, XPE/DDB1, LIG1, RAD23A and CHEK2 were used to genotype breast cancer cases and controls. DRC levels were measured using a modified host-cell reactivation assay. For XPD/ERCC2, CHEK2, LIG1, XPC and RAD23A genes there was no significant difference for the distribution of the CNV genotypes in cases versus controls (p=0.091, p=0.067, p=0.478, p=0.103, p=0.835 respectively). For DDB1/XPE, cases (n=49) were more likely to have more than two copies while controls (n=60) showed to be more likely to have less than two copies at this locus (p<0.001). We next compared the DRC values in study participants grouped by CNV genotype. Controls carrying more than two copies of the XPE/DDB1 genotype had a significant increase in DRC (p=0.003). This association was not observed in breast cancer cases (p=0.532). The Database of Genomic Variants (DGV) reports a total of five other genes in the CNV encompassing XPE/DDB1, including VWCE/URG11. This gene has been associated with gastric cancer and epithelial to mesenchymal transition. Therefore, the association of the CNV in the DDB1/XPE region with breast cancer risk may be due to the VWCE rather than DDB1/XPE. Our results show that CNV genotypes in NER gene regions may be associated with the modulation of DRC and breast cancer risk. To our knowledge, this is the first study assessing the role of CNV using a comprehensive pathway-oriented approach. It provides evidence supporting a role for CNVs in breast cancer susceptibility.

997T

Screening of Finnish RAD51C founder mutations in prostate and colorectal cancer patients. *L.M. Pelttari[1], R. Nurminen[2], A. Gylfe[3], L. Aaltonen[3], J. Schleutker[2,4], H. Nevanlinna[1]*. 1) Department of Obstetrics and Gynecology, University of Helsinki and Helsinki University Central Hospital, Helsinki, Finland; 2) Institute of Biomedical Technology/BioMediTech, University of Tampere and Fimlab Laboratories, Tampere, Finland; 3) Department of Medical Genetics, Genome-Scale Biology Research Program, University of Helsinki, Helsinki, Finland; 4) Department of Medical Biochemistry and Genetics, University of Turku, Turku, Finland.

RAD51C is involved in DNA damage repair by homologous recombination. Rare, heterozygous germline mutations in the *RAD51C* gene have been identified in breast and ovarian cancer families and a homozygous missense mutation was found in Fanconi anemia like disorder. We have previously identified two founder mutations c.837+1G>A and c.93delG in *RAD51C* in the Finnish population increasing the risk of ovarian cancer. The mutations were not associated with an increased risk of breast cancer in the absence of ovarian cancer. Many of the mutation carriers had family history of diverse cancer types, for instance colon and prostate cancer. Although several studies have identified moderate penetrance mutations in the *RAD51C* gene in breast and/or ovarian cancer families they seem to be very rare with unique mutations in outbred populations. This poses considerable challenges for investigating these in large series of patients with different cancer types.
To assess the role of *RAD51C* mutations in other common cancers we took advantage of the founder mutations in Finland and genotyped the c.837+1G>A and c.93delG mutations in genomic DNA samples from 1083 prostate and 802 colorectal cancer patients using TaqMan real-time PCR.None of the prostate or colorectal cancer patients harboured the *RAD51C* founder mutations c.837+1G>A or c.93delG. The absence of the mutations among the patients suggests that *RAD51C* does not play a significant role in the development of prostate or colorectal cancer but primarily is a susceptibility gene for ovarian cancer. These results have clinical value in assessing the cancer risks associated with *RAD51C* mutations. However, to fully assess the role of *RAD51C* in susceptibility to other cancer types, screening of the whole coding region of the gene in large sample sets would be needed.

998F

Exome Sequencing in a Familial Testicular Germ Cell Tumor Kindred. *K.A. Schrader*[1], S. Shah*[1], V. Joseph[1], R. Murali[2], R. Rau-Murthy[1], K. Sarrel[1], I. Dolgalev[2], S. Cheguri[1], C. Manschreck[1], N. Socci[3], A. Viale[4], A. Heguy[2], G. Bosl[5], D. Feldman[5], M. Robson[1], Z. Stadler[1], K. Offit[1], * equal contributions.* 1) Clinical Genetics Service, Department of Medicine, Memorial Sloan-Kettering Cancer Center, New York, NY; 2) Department of Pathology, and Human Oncology and Pathogenesis Program, Memorial Sloan-Kettering Cancer Center, New York, NY; 3) Computational Biology Program, Memorial Sloan-Kettering Cancer Center, New York, NY; 4) Genomics Core Laboratory, Memorial Sloan-Kettering Cancer Center, New York, NY; 5) Genitourinary Oncology Service, Division of Solid Tumor Oncology, Department of Medicine, Memorial Sloan-Kettering Cancer Center, New York, NY.

Introduction: Genome-wide association studies of susceptibility to testicular germ cell tumors (TGCT) have demonstrated moderately increased risks conferred by variants in or near *BAK1*, *DMRT1*, *TERT-CLPTM1L* and *KITLG*. To date, no high-penetrance susceptibility genes have been identified. To identify candidate high-penetrant TGCT susceptibility genes, we undertook whole exome sequencing in a family with three first cousins who developed TGCT at ages 26, 40 and 19 years, respectively, to identify shared rare variants seen at a very low or zero frequency in public databases. All affected individuals were related through their mothers. The family provided informed consent and the research was IRB-approved.

Methods: DNA was extracted from saliva and subjected to exome capture using the Agilent SureSelect Human All Exon 50Mb kit followed by 75bp paired-end sequencing on an Illumina HiSeq-2000 system. Alignment and variant calling was performed by the MSKCC Bioinformatics core using BWA and Best Practice Variant Detection with the GATK v2. Downstream analysis consisted of filtering out low quality variant calls and those already reported in public databases. Under the assumption of autosomal dominant or X-linked inheritance, we identified, novel, functional (truncating, indel, splice site, nonsynonymous) variants shared by all three of the affected individuals in keeping with the patterns of inheritance. To filter out rare familial variants less likely to be associated with the TGCT phenotype, we also performed exome sequencing on an unaffected 68 year-old uncle who has two unaffected sons in their 40's, based on the assumption that he does not carry the highly penetrant TCGT susceptibility allele.

Results: The three affected first cousins each carried, respectively, 370, 395 and 436 functional single nucleotide variants (SNV) and 75, 83 and 84 indels absent in dbSNP132. Forty-eight SNVs and 38 indels were shared. After filtering out the variants seen in the uncle, 26 SNVs and 18 indels remained. Following further subtraction of reported variants and orthogonal Sanger sequence validation, the number of candidates was reduced to 9, restricted to chromosomes 1 and X.

Interpretation: The candidate variants of greatest interest occur in a Rho GTPase, a chromatin modifier gene and a gene related to the spindle assembly checkpoint. Tumor analysis and replication in other families for mutations in the candidate genes and pathways is underway.

999W

SLX4 mutation in hereditary breast cancer. *S. Shah[1], Y. Kim[2], I. Ostrovnaya[1], F. Lach[2], R. Murali[1], K. Schrader[1], V. Joseph[1], K. Sarrel[1], R. Rau-Murthy[1], N. Hansen[1], S. Cheguri[1], J. Littman[1], L. Zhang[1], K. Offit[1], A. Smogorzewska[2]*. 1) Memorial Sloan Kettering Cancer Center, New York, NY; 2) The Rockefeller University, New York, NY.

SLX4 is a DNA repair protein that serves as a docking platform for three structure-specific endonucleases. Recent studies have reported mutations in SLX4 in a new subtype of Fanconi Anemia (FA-P). Monoallelic defects in several Fanconi Anemia (FA) genes are known to confer susceptibility to breast and ovarian cancer. Therefore, we resequenced the entire SLX4 coding region in 739 (288 Ashkenazi Jewish and 480 Non-Ashkenazi Jewish) breast cancer patients with 3 or more family members affected by breast cancer and no known BRCA1/BRCA2 mutations, to determine if SLX4 is involved in breast cancer susceptibility. We found a novel nonsense (p.W823*) mutation in one patient. In addition we also found 13 novel, 23 rare (MAF<0.1%), and 15 common (MAF1%) missense variants, of which 22 (5 novel and 17 rare) are predicted as damaging by Polyphen2 (score= 0.65–1). We performed functional complementation using p.W823* and 4 others (3 novel and 1 rare) mutant SLX4 cDNA in a SLX4 null cell line. While, wildtype SLX4 and all the other variants fully rescued the sensitivity to Mitomycin C (MMC), campthothecin(CMT), and PARP inhibitor, p.W823* SLX4 mutant failed to do so.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1000T

Contribution of ARLTS1 to prostate cancer susceptibility. *S. Siltanen[1], H. Rauhala[1], TLJ. Tammela[2], P. Kujala[3], T. Visakorpi[1], T. Wahlfors[1], J. Schleutker[1,4]*. 1) Institute of Biomedical Technology/BioMediTech, University of Tampere and Fimlab Laboratories, Tampere, Finland; 2) Department of Urology, University of Tampere and Tampere University Hospital, Tampere, Finland; 3) Department of Pathology, Tampere University Hospital, Tampere, Finland; 4) Department of Medical Biochemistry and Genetics, University of Turku, Turku, Finland.

ARLTS1 (ADP-ribosylation factor-like tumor suppressor gene 1, also known as ARL11) is a tumor suppressor gene located at 13q14.3, a region frequently deleted in both sporadic and familial prostate cancer (PCa). Previously, we have shown that ARLTS1 variants, especially Cys148Arg (T442C), revealed a statistically significant association with an increased risk for homozygous CC carriers both in unselected and familial PCa cases. Cys148Arg (T442C) was also the only variant observed with a greater frequency among malignant tissue samples, PCa cell lines and xenografts, supporting its role in prostate carcinogenesis. Further, down regulation of ARLTS1 in clinical samples and its co-expression signature showed that ARLTS1 expression was strongly associated with immune processes. In this study, we investigated the mechanisms behind the observed associations. Quantitative-RT-PCR was used to examine the ARLTS1 expression status in eight previously 5-aza/TSA treated and demethylated PCa cell lines. Bisulphite sequencing was used to examine whether promoter methylation is behind the downregulation of ARLTS1 also in PCa. By staining prostatectomy FFPE block tissue slides with an antibody we examined the differences in lymphocyte infiltration, between cancers with CC genotype and TT genotype. We studied also the ARLTS1 tumor suppressor properties, apoptosis, cell growth and migration in PCa cell lines 22Rv1 and LAPC4, by blocking ARLTS1 expression with ARL11 siRNA. TSA alone was able to restore the ARLTS1 expression in PCa cell line LnCaP and in normal prostate epithelial cell line PrEC. Both TSA and 5-aza were needed to restore the expression in PC3 and LAPC4 PCa cell lines revealing that both DNA hypermethylation and histone modifications may regulate the expression. However, bisulphite sequencing did not work as expected on the CpG islands of promoter region. Lack of official CpG islands in the CpG databases supported this result which is contrary to previous reports on ARLTS1 structure in lung and ovarian cancers. Studies on gene silencing and immunostaining are still on-going.

1001F

Meta-analysis of 25 prostate cancer associated SNPs in high-risk prostate cancer families: new evidence from the International Consortium for Prostate Cancer Genetics (ICPCG). *C. Teerlink[1], S. Thibodeau[2], D. Schaid[3], K. Cooney[4], E. Lange[5], C. Maier[6], J. Stanford[7], E.A. Ostrander[8], J. Schleutker[9], G. Cancel-Tassin[10], O. Cussenot[10], R. Eeles[11], D. Easton[12], W. Isaacs[13], J. Xu[14], J. Carpten[15], J. Bailey-Wilson[16], F. Wiklund[17], A. Whittemore[18], W. Catalona[19], W. Foulkes[20], N. Camp[1], L. Cannon-Albright[1,21], International Consortium for Prostate Cancer Genetics.* 1) Dept Internal Medicine, Univ Utah Sch Med, Salt Lake City, UT; 2) Department of Lab Medicine and Pathology, Mayo Clinic, Rochester, MN 55905; 3) Department of Health Sciences Research, Mayo Clinic, Rochester, MN 55905; 4) Department of Internal Medicine, University of Michigan Medical School, Ann Arbor, MI 48109; 5) Departments of Genetics and Biostatistics, University of North Carolina, Chapel Hill, NC 27599; 6) Institute for Human Genetics, University of Ulm, Germany; 7) Fred Hutchinson Cancer Research Center (FHCRC), Division of Public Health Sciences, Seattle, WA 98195; 8) Cancer Genetics Branch, National Human Genome Research Institute (NHGRI), National Institutes of Health (NIH), Bethesda, MD 20892; 9) Department of Medical Biochemistry and Genetics, University of Turku, Turku 20520, Finland; 10) CeRePP ICPCG Group, Hopital Tenon, Assistance Publique-Hopitaux de Paris, Paris75020, France; 11) The Institute of Cancer Research, Sutton, Surrey SM2 5NG, UK; 12) Strangeways Laboratory, Worts Causeway, Cambridge CB1 8RN, UK; 13) Johns Hopkins University ICPCG Group, Department of Urology, Johns Hopkins Medical Institutions, Baltimore, MD 21287; 14) Center for Human Genomics, Wake Forest University School of Medicine, Winston-Salem, NC 27157; 15) Translational Genomics Research Institute, Phoenix, AZ 85004; 16) Johns Hopkins University, Baltimore, MD 21224; 17) Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden; 18) Department of Health Research and Policy, Stanford School of Medicine, Stanford, CA 94305; 19) Northwestern University ICPCG Group, Northwestern University Feinberg School of Medicine, Chicago, IL 60611; 20) Program in Cancer Genetics, McGill University, Montreal, Quebec H3T 1E2, Canada; 21) George E. Wahlen Department of Veterans Affairs Medical Center, Salt Lake City, UT 84148.

Previous GWAS studies have reported associations between various SNPs and prostate cancer using cases unselected for family history. We report the results of a validation study of 25 SNPs reported to be associated with prostate cancer using a novel analysis of related familial prostate cancer cases. Fourteen study sites contributed a total of 11,783 genotyped samples, including 3,786 unaffected controls, 5,628 cases with non- or undetermined-aggressive disease, and 2,368 cases with aggressive disease. The majority of genotyped samples (n = 8,573) originated from 2,276 pedigrees, the remainder were individually sampled cases. Each site contributed its own controls. We used PedGenie software to account for known relationships. We conducted separate analyses for all prostate cancer and for aggressive prostate cancer only. For the non-specific prostate cancer phenotype, our analysis showed that 16 of the 25 SNPs were statistically significant ($p < 2E\text{-}3$ (= 0.05/25)), including SNPs on chromosomal bands 6q25, 7p15, 8q24, 10q11, 11q13, 17q12, 17q24, and Xp11. For the aggressive prostate cancer phenotype, our analysis showed that 8 of the 25 SNPs, were statistically significant ($p < 2E\text{-}3$), including most of the same chromosomal bands except 6q25 and 10q11 but also included a SNP at 2p15. The results of this analysis validate strong association to familial prostate cancer for the majority of SNPs considered and demonstrate the power of a GWAS approach that integrates data for related cases.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1002W

CDH1 mutations in patients with lobular carcinoma of the breast. *A.L. Valente[1], S. Rummel[1], C.D. Shriver[2], R.E. Ellsworth[3].* 1) Clinical Breast Care Project, Windber Research Institute, Windber, PA; 2) Clinical Breast Care Project, Walter Reed National Military Medical Center, Bethesda, MD, USA; 3) Clinical Breast Care Project, Henry M. Jackson Foundation, Windber, PA, USA.

Background: E-cadherin (CDH1) is involved in cell adhesion and maintenance of tissue architecture. Loss of CDH1 has been detected in a number of cancers characterized by loose cell-cell adhesion such as hereditary diffuse gastric, endometrial, ovarian and invasive lobular carcinoma (ILCA) of the breast, and germline mutations in CDH1 have been associated with increased risk of developing ILCA but not other breast tumor types. To improve understanding of the role of CDH1 mutations in breast cancer, mutation status was determined in a breast cancer cohort of patients with ILCA with and without family history of other tumor types. Methods: The Clinical Breast Care Project database was queried to identify all patients with ILCA with or without a family history of gastric/stomach, endometrial or ovarian cancer. Genomic DNA was isolated from peripheral blood samples. DNA variants were detected for each exon of CDH1 using high-resolution melting technology and the underlying change identified by direct sequencing. Variant A617T was assayed using TaqMan SNP assay rs33935154. Results: Of the 72 patients with ILCA, five had a family history of stomach cancer. Two private mutations were detected: P825Q in a patient with a family history of stomach cancer, and G879S in a patient with no family history of cancer. In addition, the A592T variant, which has been reported as non-pathogenic, was detected in a Hispanic woman while the causative mutations 1137G-A, identified in a Caucasian woman, and A617T identified in two African American women were detected. The A617T mutation was then evaluated in an additional 166 African American patients: 12 patients carried the mutation, nine of which had IDCA, two with microinvasive disease and one with mixed ductal and lobular features. Conclusions: The frequency of CDH1 mutations in this patient population was low (<10%), with only two mutations with known clinical consequences identified. In addition, the detection of the A617T mutation previously associated with endometrial cancer in African American women, in patients with predominantly ductal carcinomas questions whether this mutation is associated with increased risk of ILCA. The low frequency of germline mutations in CDH1 suggests that inherited mutations are not the predominant mechanism but rather alternate methods, such as epigenetic modification or miRNA silencing, may be driving the suppression of E-cadherin expression and the development of ILCA.

1003T

Fine-mapping of Breast Cancer Genome-Wide Association Studies Loci in Women of African Ancestry. *Y. Zheng[1], T.O. Ogundiran[2], K.L. Nathanson[3], E.M. John[4,5], A. Hennis[6], S. Ambs[7], C. Adebamowo[8], S.M. Domchek[3], T.R. Rebbeck[9], M.S. Simon[10], B. Nemesure[11], S.Y. Wu[11], M.C. Leske[11], Q. Niu[1], J. Zhang[11], E.R. Gamazon[11], N.J. Cox[11], O.I. Olopade[11], D. Huo[12].* 1) Department of Medicine, The University of Chicago, Chicago, IL, USA, IL; 2) Department of Surgery, College of Medicine, University of Ibadan, Ibadan, Nigeria; 3) Department of Medicine, University of Pennsylvania, Philadelphia, PA, USA; 4) Cancer Prevention Institute of California, Fremont, CA, USA; 5) Department of Health Research & Policy, Stanford University School of Medicine and Stanford Cancer Institute, Stanford, CA, USA; 6) Chronic Disease Research Centre, Tropical Medicine Research Institute, University of the West Indies, Bridgetown, Barbados; 7) Laboratory of Human Carcinogenesis, National Cancer Institute, Bethesda, MD, USA; 8) Department of Epidemiology & Preventive Medicine, University of Maryland, Baltimore, MD, USA; 9) Department of Statistics & Epidemiology, University of Pennsylvania, Philadelphia, PA, USA; 10) Department of Oncology, Karmanos Cancer Institute at Wayne State University, Detroit, MI, USA; 11) Department of Preventive Medicine, State University of New York at Stony Brook, Stony Brook, NY, USA; 12) Department of Health Studies, The University of Chicago, Chicago, IL, USA.

Globally, breast cancer affects more than 1.1 million women each year and is the most common cancer diagnosis among women in the United States. Breast cancer is a genetically heterogeneous disease characterized by different penetrance, complex phenotypes, and a polygenic pattern of inheritance. Dozens of novel single nucleotide polymorphisms (SNPs) for breast cancer susceptibility have been identified by genome-wide association studies (GWAS). However, those susceptibility loci were primarily discovered and validated in women of European ancestry and Asians. Given that linkage disequilibrium (LD) is quite different among ethnic populations, we sought to evaluate common genetic variants at 23 breast cancer GWAS susceptibility regions using 1,509 breast cancer cases and 1,383 controls of African ancestry. The average age (mean ± standard deviation) was 48.0 ± 12.0 years in cases and 47.2 ± 17.2 years in controls. Ascertainment of cases and controls occurred in Ibadan, Nigeria (681 cases and 282 controls), Barbados (93 cases and 244 controls), and multiple sites in the US (735 cases and 857 controls). A total of 965 tagging SNPs (selected to capture either the targeted LD blocks or the GWAS index SNPs) and 30 ancestry informative markers were genotyped using the Illumina GoldenGate Genotyping platform. Imputation for non-genotyped SNPs across the regions was conducted using MACH with phased YRI and CEU data from HapMap Phased II (release 22) and 1000 Genome Project Nov 2010 release as the reference panels. Unconditional logistic regression models adjusting for study site and estimated African ancestry proportion were then fit by SAS and PLINK. Both allele dosage effects and genotypic effects were examined. Our fine-mapping analysis found several novel breast cancer susceptibility markers in loci such as 5p12 (P = 0.0040), 5q11 (P = 0.00053), and 10p15 (P = 0.0015), which are unique for women of African ancestry. Our study underscores the importance of fine mapping of reported breast cancer GWAS loci in diverse and understudied populations. Future studies and new GWAS in women of African ancestry hold promise to uncover additional variants for breast cancer susceptibility with implications throughout the Diaspora.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1004F

L1 encoded ORF1p facilitates tumorigenesis and suppresses nucleus accumulation of Smad-4 in HepG2 Cells. *Y. Zhu[1,2], F. Feng[3], X. Gao[4], J. Yu[5], M. Hu[6], F. Zhang[1], Y. Wang[5], Q. Zhang[7], Y. Guo[1].* 1) Tumor Center in Beijing 301 Hospital, Beijing 100853, China; 2) The institute of life sciences and bio-engineering in Beijing Jiaotong University, Beijing 100044,China; 3) 3 Beijing Institute of Pharmacology and Toxicology, Beijing 100850, China; 4) The center for liver treatment in Beijing 302 Hospital, Beijing 100039, China; 5) Beijing Institute of Basic Medicine Science, Beijing 100850, China; 6) Department of tumor molecular in Beijing 307 Hospital, Beijing 100071, China; 7) Institute of Liver Transplantation, General Hospital of Chinese People's Armed Police Force, Beijing, 100039.

L1 encoded ORF1p facilitates tumorigenesis and suppresses nucleus accumulation of Smad-4 in HepG2 Cells L1 (LINE-1, Long interspersed nucleotide element-1), an autonomous retrotransposon, comprises about 17% of the human genome, most of which appear as 5' truncated forms. The full-length sequence contains 5' untranslated region (UTR), which encompass an internal promoter for RNA polymerase II, two open reading frames (ORF1 and ORF2), and a 3' terminal polyadenylation site [2]. There are thousands of full-length L1 in human genome and many of which encode ORF1p without making ORF2p. The competent L1 is considered to have 80–100 copies. L1 is activated in early stage of germ cell and embryo development and, however, unscheduled activation in various cancer tissues. Suppressing endogenous L1 by L1-ORF2 targeted RNAi and RT inhibitors in vitro and in vivo promotes the cell transition from highly proliferating, transformed phenotypes to low proliferating, differentiated phenotypes. But however, ORF1p remains to be elucidated, even that it has an excessive expression comparing to ORF2p. L1-encoded ORF1p (L1-ORF1p), a RNA binding protein, facilitating the RNA: DNA duplex in target site suggests that ORF1p encompasses some important functions beyond retrotransposition. To elucidate the function of L1-ORF1p, the transfected HepG2 cells with over-expression and down-regulation of ORF1p were analyzed and the proteins associated with L1-ORF1p were identified by co-immunoprecipitation. As the results, ORF1p over expression facilitates cells proliferation significantly in vitro and in vivo and its down-regulation results in opposed phenomenon. By co-immunoprecipitation analysis, ORF1p is demonstrated to form a complex with Smad4 and, in turn, suppress the nucleus accumulation of Smad4 elicited by TGF-β. MH2 domain in Smad4 was identified to be responsible for its interaction with L1-ORF1p. So it suggests that (1) upon the scenario of cell regulation L1-ORF1p beyond the function of retrotransposition might involve in some important regulation in tumorigenesis or at least acts as a key molecule in this process and alteration of its expression can switch cell statues; (2) ORF1p plays some important role for tumor growth and Smad4 may act as a mediator involved in this process. Key words: LINE-1; retrotransposon; ORF1p; Smad4; interaction.

1005W

Copy number variation analysis in 222 mutation negative polyposis patients reveals potential new candidate genes. *S. Aretz[1], S. Horpaopan[1], S. Vogt[1], I. Spier[1], A.M. Zink[1], S. Herms[1,2], A. Laner[3,4], K. Wöllner[1], S. Pasternack[1], M. Draaken[1,2], D. Stienen[1], S. Uhlhaas[1], E. Holinski-Feder[3,4], M.M. Nöthen[1,2], P. Hoffmann[1,2].* 1) Institute of Human Genetics, University Hospital Bonn, Germany; 2) Deptartment of Genomics, Life & Brain Center, University of Bonn, Germany; 3) University Hospital of the Ludwig-Maximilians-University, Campus Innenstadt, Munich, Germany; 4) MGZ - Center of Medical Genetics, Munich, Germany.

Background: Adenomatous polyposis syndromes are characterized by multiple colorectal adenomas and a high lifetime risk of colorectal cancer. Germline mutations of the APC and MUTYH genes cause the autosomal dominant Familial Adenomatous Polyposis (FAP) and autosomal recessive MUTYH-associated Polyposis (MAP), respectively. However, in up to 50% of families no germline mutation could be identified. Copy number variants (CNVs), in particular heterozygous microdeletions, contribute significantly to the mutation spectrum of hereditary tumor syndromes and thus it can be hypothesized that those heterozygous deletion CNVs might also be the underlying cause in yet unidentified genes responsible for adenomatous polyposis syndromes. Methods: Genomic DNA from 222 unrelated mutation negative polyposis patients was used for genome-wide SNP genotyping with the HumanOmni1-Quad BeadArray (Illumina). Putative CNVs were identified by the QuantiSNP v2.2 algorithm, filtered according to various criteria by use of the Cartagenia Bench™software, by in-silico-analysis, and by comparison with 531 healthy controls, and validated by qPCR. The gastrointestinal expression of genes covered by the identified CNVs was examined by cDNA analysis of normal mucosa. Results: 69 unique heterozygous deletion CNVs (size 10–664 kb) containing 88 protein coding genes could be identified in 60 patients (27%) but not in controls. Of them, 62 genes are partly or completely deleted, in 26 more the deletion affects intronic regions only. All but ten patients harbor just one CNV. Almost all CNVs are present only once in the whole cohort. After further prioritisation and expression analysis, 55 deleted genes remain. These candidate adenoma genes include protein kinases, transcription factors, and potential tumor suppressors. A few of them have been reported to be associated with CRC phenotypes. Conclusions: By applying stringent filter criteria, we identified a group of rare deletion CNVs which might contain predisposing genes for adenoma formation. After prioritization of the included genes according to function, pathway, and literature, present work includes sequencing the coding regions of the most interesting candidates in all patients to look for pathogenic point mutations. Additional steps are loss-of-heterozygosity analyses in tumor DNA. The study was supported by the German Cancer Aid.

1006T

Evaluation of the RHINO gene for breast cancer predisposition in Finnish breast cancer families. *T. Heikkinen[1], E. Huovari[1], S. Vilske[1], C. Blomqvist[2], J. Schleutker[3,4], A. Kallioniemi[4], K. Aittomäki[5], H. Nevanlinna[1].* 1) Department of obstetrics and gynecology, Helsinki university central hospital, Helsinki, Finland; 2) Department of Oncology, University of Helsinki and Helsinki University Central Hospital, Helsinki, Finland; 3) Department of Medical Biochemistry and Genetics, University of Turku, FI-20014 Turku, Finland; 4) Institute of Biomedical Technology/BioMediTech, University of Tampere and Fimlab Laboratories, FI-33520 Tampere, Finland; 5) Department of Clinical Genetics, University of Helsinki and Helsinki University Central Hospital, Helsinki, Finland.

Genetic predisposition to breast cancer is caused by mutations in multiple loci. Among the high and moderate penetrance risk genes are many genes involved in DNA damage response, including BRCA1, BRCA2, and PALB2. A recently identified gene RHINO (C12orf32), interacting with Rad9-Rad1-Hus1 complex and with ATR activator TopBP1, is a novel candidate for DNA damage response and thus to breast cancer susceptibility. We sequenced the coding regions of the RHINO gene in blood DNA samples of 466 index patients of breast cancer families and in 507 population controls. We identified nine variants including four missense changes and one single nucleotide substitution creating a premature stop codon c.250C>T, R84X (rs140887418). The R84X nonsense variant truncates 65% of the 238 amino acid RHINO protein and therefore seems the most likely functional variant among the RHINO changes. There was no significant difference between the frequencies of any of the missense variants in the sequenced cases and controls. To investigate the role of the R84X nonsense variant in breast cancer predisposition we determined the genotype of the R84X variant in altogether 2214 cases and 1169 controls from the Helsinki region and in 730 cases and 521 controls from the Tampere region. The variant was present in cases and controls in similar frequencies of 0.4% and 0.8% respectively (P=0.106). Involvement of RHINO in breast cancer predisposition is unlikely.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1007F

A gene-based association approach prioritizes candidate susceptible genes for chronic myeloid leukemia. *J.-H. Park[1], H.-H. Won[2], J.-W. Kim[3], D.H. Kim[4], S.-T. Lee[3], S. Kim[2], S.-H. Kim[3], C.W. Jung[5].* 1) Dept. of Health Sciences and Technology, Graduate School, Samsung Advanced Institute for Health Science and Technology, Sungkyunkwan University, Korea; 2) Samsung Biomedical Research Institute, Samsung Medical Center, Korea; 3) Dept. of Laboratory Medicine and Genetics, Sungkyunkwan University School of Medicine, Samsung Medical Center, Korea; 4) Chronic Myelogenous Leukemia Group, Department of Hematology/Medical Oncology, Princess Margaret Hospital, University Health Network, University of Toronto, Toronto, Canada; 5) Department of Hematology/Oncology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea.

Gene-based genome-wide association methods have been developed to overcome the limited power of a single SNP-based approach especially for variants with mild effects and to investigate relationships between genes and variants. It is also important to prioritize many candidate genes from the associated loci because functional studies for all located genes are not feasible. Therefore, the gene-based approach holds promise for selecting the most plausible candidate variants that are functionally relevant to genes and etiology of diseases. In this study, we further investigated candidate genes using the gene-based genome-wide association approach for the previously identified novel locus 6q25.1 associated with susceptibility to chronic myeloid leukemia (CML) (Kim *et al*., Blood, 2011, 117(25):6906-11). We utilized the versatile gene-based association study (VEGAS) method (Liu *et al*., AJHG, 2010, 87(1):139-45) that is an association test at gene level by considering the association results from the single SNP-based approach and linkage disequilibrium between SNPs in each gene. We carried out imputation and used approximately 1.6 million autosomal SNPs after quality control. We tested the assigned 17,542 genes with genotyped and imputed SNPs using the VEGAS and CML association analysis results. We also used other gene-based methods that combine P values of assigned and correlated SNPs for a corresponding gene. As a result of the gene-based approach, we confirmed three genes (*RMND1*, P=0.000011; *C6orf211*, P= 0.000016; and *ZBTB2*, P=0.00029) in the candidate locus (6q25.1) of our previous study. The results of other gene-based methods (ProxyGeneLD (Hum Genet, 2009, 126(2):289–301) and GSA-SNP (Nucleic Acids Research, 2010, 38(2):749-54)) also strongly supported our outcomes including *RMND1*. We also examined whether the identified variants could affect gene expression functionally using accessible expression quantitative trait loci (eQTL) databases. The most significant gene,*RMND1* included several SNPs (rs963193,rs13201167,rs6931104 and *et al*.) with eQTL information (http://eqtl.uchicago.edu/), which might affect regulation or expression of the corresponding gene. *RMND1* was also shown to be expressed in whole blood in public databases. Taken together with the result that the SNPs were replicated in independent Korean and European populations in our previous study, these results suggest that *RMND1* is a highly potential, susceptible gene for the development of CML.

1008W

Proteomic analysis of gastric cancer from individuals of Northern Brazil. *M.F. Leal[1], J. Chung[2], D.Q. Calcagno[1], P.P. Assumpçao[3], S. Demachki[4], I.D.C.G Silva[5], R. Chammas[6], R.R. Burbano[7], M.A.C. Smith[1].* 1) Department of Morphology and Genetics, Federal University of Sao Paulo, Sao Paulo, SP, Brazil; 2) Department of Microbiology, Immunology and Parasitology, Federal University of Sao Paulo, Sao Paulo, SP, Brazil; 3) Surgery Service, João de Barros Barreto University Hospital, Federal University of Para;, Belem, PA, Brazil; 4) Pathology Service, Joao de Barros Barreto University Hospital, Federal University of Para, Belem, PA, Brazil; 5) Department of Gynecology, Federal University of Sao Paulo, Sao Paulo, SP, Brazil; 6) School of Medicine, University of Sao Paulo, Sao Paulo, SP, Brazil; 7) Institute of Biological Sciences, Federal University of Para, Belem, PA, Brazil.

Gastric cancer is the second leading cause of cancer-related death worldwide. The identification of new cancer biomarkers is necessary to reduce the mortality rates through the development of new screening assays and early diagnosis, as well as new target therapies. In this study, we performed a differential proteomic analysis of noncardia gastric neoplasias of individuals from Northern Brazil. All samples presented H. pylori infection. The proteins were analyzed by two-dimensional electrophoresis and mass spectrometry. For the identification of differentially expressed proteins, we used statistical tests with bootstrapping resampling to control the type I error in the multiple comparison analyses. We identified 111 proteins involved in gastric carcinogenesis. The computational analysis revealed several proteins involved in the energy production processes and reinforced the Warburg effect in gastric cancer. ENO1, HSPB1 and ANXA2 expression were further evaluated. ENO1 was selected due to its role in aerobic glycolysis that may contribute to the Warburg effect. Although we observed two up-regulated spots of ENO1 in the proteomic analysis, the mean expression of ENO1 was reduced in gastric tumors by western blot. However, mean ENO1 expression seems to increase in more invasive tumors. This lack of correlation between proteomic and western blot analyses may be due to the presence of other ENO1 spots that present a slightly reduced expression, but with a high impact in the mean protein expression. In neoplasias, HSPB1 is induced by cellular stress to protect cells against apoptosis. In the present study, HSPB1 presented an elevated protein and mRNA expression in a subset of gastric cancer samples. However, no association was observed between HSPB1 expression and clinicopathological characteristics. ANXA2 was selected due to its suggested role in cellular transformation. *ANXA2* mRNA expression was up-regulated in more than half of gastric tumors compared to control tissues and it was associated with the presence of lymph node metastasis. However, we did not observe a correlation between gene and protein expression, underling the complex control of mRNA and protein expression. The present study enhances efforts to expand knowledge about gastric carcinogenesis, especially for neoplasms of patients from Northern Brazil. These biomarkers may be also useful for the assessment of prognosis and stratification for therapy if validated in larger clinical study sets.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1009T

Fine-mapping of IL-16 gene and prostate cancer risk in African Americans. *K. Batai[1], E. Shah[1], M. Ruden[1], J. Newsome[1], A.B. Murphy[2], C. Ahaghotu[3], R.A. Kittles[1,4,5]*. 1) Institute of Human Genetics, University of Illinois at Chicago, Chicago, IL; 2) Department of Urology, Feinberg School of Medicine, Northwestern University, Chicago, IL; 3) Division of Urology, Department of Surgery, Howard University Hospital, Washington, D.C; 4) Section of Hematology/Oncology, Department of Medicine, University of Illinois at Chicago, Chicago, IL; 5) Division of Epidemiology and Biostatistics, School of Public Health, University of Illinois at Chicago, Chicago, IL.

Background: Prostate cancer (Pca) is the most common type of cancer among men in the United States, and its incidence and mortality rates are disproportionate among ethnic groups. While genome-wide association studies (GWAS) of European descents have identified candidate loci associated with Pca risk, including a variant in the IL-16 gene, replication studies in African Americans (AAs) have been inconsistent. Here we explore SNP variation in IL-16 in AAs and test for association with Pca.

Methods: We tested association of 2,257 genotyped and imputed SNPs spanning IL-16 in 605 AAs from Washington, DC. Washington, D. C. samples were genotyped as a part of a large GWAS consortium for Pca in AAs using the Illumina Infinium 1M-Duo bead array. Imputation was performed using IMUTE2 and 1000 Genomes Project data as a reference panel. Eleven of the SNPs were also genotyped in a replication population of AAs from Chicago (n=506). Allelic associations were tested adjusting for age, Global and Local west African ancestry.

Results: Analyses of genotyped and imputed SNPs revealed that a cluster of IL-16 SNPs were significantly associated with Pca risk. The strongest association was found for rs7175701 (P=9.78 × 10^{-8}). In the replication population, another SNP (rs11556218), which was not associated with Pca in the Washington, D.C. AAs, was associated with Pca risk (P=0.01). In the pooled analysis, we identified three independent loci within IL-16 that were associated with Pca risk. SNP eQTL analyses revealed that rs7175701 is predicted to influence the expression of IL-16 and other cancer related genes.

Conclusion: We were successful in the `local' replication revealing that IL-16 variants associated with Pca risk in two independent samples of AAs. Our study provides evidence that IL-16 polymorphisms likely play a role in Pca susceptibility among AAs. The significantly associated SNPs are likely to affect gene expression. Our findings are significant given that there has been limited focus on the role of IL-16 genetic polymorphisms on Pca risk in AAs.

1010F

Fas signaling pathway-based analysis of genome-wide data and risk of gastric cancer. *P.L. Hyland[1,2], S-W. Lin[1,2], N. Hu[1], Z-Z. Tang[3], L. Wang[1], C. Wang[1], H. Su[1], T. Ding[3], J-H. Fan[4], Y-L. Qiao[4], X. Xiong[5], W. Wheeler[5], C. Griffen[5], K. Yu[1], Z. Wang[6], L. Burdett[6], S.J. Chanock[6], S.M. Dawsey[1], M.A. Tucker[7], N.D. Freedman[1], A.M. Goldstein[1], C.C. Abnet[1], P.R. Taylor[1].* 1) DCEG, National Cancer Institute (NCI), NIH, Bethesda, MD; 2) CPFP, Division of Cancer Prevention, NCI, NIH, Bethesda, MD; 3) Shanxi Cancer Hospital, Taiyuan, People's Republic (PR) China; 4) Dept of Epidemiology, Cancer Institute (Hospital), Chinese Academy of Medical Sciences, Beijing, People's Republic (PR) China; 5) IMS, Inc, Silver Spring, MD; 6) Core Genotyping Facility, NCI-Frederick, SAIC-Frederick Inc, and DCEG, Bethesda, MD; 7) Human Genetics Program, DCEG, NCI, NIH, Bethesda, MD.

The populations of the Shanxi and Linxian regions in north central China are at high risk for gastric cancer (GC) and altered FAS-mediated cell signaling and/or apoptosis may contribute to this risk. In this study, we used a pathway-based approach to incorporate a priori knowledge into the analysis of existing genome wide association study (GWAS) data. We examined the association of 554 single nucleotide polymorphisms (SNPs) in 53 genes related to FAS signaling using a pathway-based (gene-set) approach in 1758 GC cases (including 1126 gastric cardia adenocarcinomas [GCA], and 632 gastric noncardia adenocarcinomas [GNCA]) and 2111 controls from the GWAS of gastric cancer in ethnic Chinese. SNP-, gene- and pathway-based associations with risk of GC overall as well as GCA and GNCA were evaluated using the adaptive rank-truncated product (ARTP) method and the additive model within unconditional logistic regressions adjusted for age and sex. Statistical significance was evaluated empirically by permutation. Significant associations were observed for genes involved in the Fas signaling pathway with risk of overall GC (P = 5.5E-04) and GCA (P = 6.3E-03), but not GNCA (P = 8.1E-02). Ten genes including *MAP2K4, FAF1, MAPK8, CASP10, CASP8, CFLAR, MAP2K1, CAP8AP2, PAK2* and *IKBKB* were significantly associated with risk of GC (nominal P < 0.05). *FAF1* and *MAPK8* were significantly associated with risk of both GCA and GNCA (nominal P < 0.05) in our high risk population. In contrast, *MAP2K4* and *IKBKB* were significantly associated with GNCA risk, but not GCA; and *CASP8, CASP10, CFLAR* and *MAP2K1* were significantly associated with risk of GCA, but not GNCA. Our examination of genetic variation in the FAS signaling pathway is consistent with an association of altered FAS signaling and/or apoptosis with risk of GC. Results suggest that this association may be driven by genetic variation in specific genes with important roles in FAS-signaling in cells. As one of the first attempts to investigate a pathway-level association, our results suggest that these genes and the FAS signaling pathway warrant further evaluation in relation to GC risk in other populations.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1011W

Fine-mapping identifies multiple prostate cancer risk loci at 5p15, one of which associates with TERT expression. *Z. Kote-Jarai[1], E. Saunders[1], D. Leongamornlert[1], M. Tymrakiewicz[1], T. Dadaev[1], S. Jugurnauth-Little[1], H. Ross-Adams[2], A. Al-Olama[3], S. Bennloch[3], S. Halim[2], R. Russel[2], A. Dunning[3], D. Neal[4], F. Hamdy[5], J. Donovan[6], D. Easton[3], R. Eeles[1], The UK Genetics Prostate Cancer Study Collaborators, The PRACTICAL Consortium.* 1) Oncogenetics, The Institute of Cancer Research, 15 Cotswold Rd, Sutton, Surrey, United Kingdom; 2) Cancer Research UK Cambridge Research Institute, Li Ka Shing Centre, Robinson Way, Cambridge, UK; 3) Centre for Cancer Epidemiology, Strangeways Laboratory, Worts Causeway,CAmbridge, UK; 4) Surgical Oncology, University of Cambridge, Hills Rd, CAmbridge, UK; 5) Nuffield Department of Surgical Sciences, Univeristy of Oxford, Oxford, UK; 6) School of Social and Communiy Medicine, University of Bristol, Whatley Rd, Bristol, UK.

We have previously reported an association between prostate cancer (PrCa) risk and rs2242652 on 5p15. rs2242652 lies in intron 4 of TERT, which encodes telomerase reverse transcriptase, the catalytic subunit of the telomerase ribonucleoprotein complex. Telomerase catalyzes the de novo addition of telomere repeat sequences on to chromosome ends and thereby counterbalances telomere-dependent replicative senescence. Associations between SNPs in the TERT region and multiple cancer types have been reported; however no correlation has been observed thus far between the cancer-associated SNPs in TERT and gene expression or telomere length. To comprehensively evaluate the association between genetic variation across this region and PrCa we performed a fine-mapping analysis by direct genotyping using either a custom Illumina iSelect array (114 SNPs on iCOGS) or Sequenom MassArray iPlex (25 SNPs) followed by imputation of 1,094 SNPs in 22,301 PrCa cases and 22,320 controls in the PRACTICAL consortium. To determine independently associated variants in this region, we performed forward and backward stepwise logistic regression; SNPs were included in the model if they were significant at $P<10-4$ after adjustment for other SNPs. Both regression models identified multiple independent associations, reflecting the complexity of this region. These SNPs fall into four regions consisting of clusters of highly or moderately correlated variants, and there is only weak LD between SNPs in different regions. To investigate whether SNPs in any of these regions were associated with TERT gene expression, we performed qPCR assays on RNA from benign prostate tissue samples using the Fluidigm Biomark™ HD system. We found evidence that the risk reducing alleles of several variants in region one associated with elevated TERT expression, providing a plausible mechanism for the differential effect of SNPs on PrCa risk. Deep re-sequencing of these loci may help to further refine this region and facilitate selection of prospective causal variants for functional validation studies.

1012T

Fine mapping of 12q24 shows a promising association locus for renal cell carcinoma. *J.R. Toro[1], S. Han[1], M. Yeager[1], L. Moore[1], M. Purdue[1], M. Johansson[2], G. Scelo[2], V. Gaborieau[2], C. Berg[3], R. Grubb[4], V. Stevens[5], M. Thun[5], W. Diver[5], D. Albanes[1], S. Weinstein[1], J. Virtamo[6], L. Burdett[1], A. Brisuda[7], J. McKay[2], J. Fraumeni[1], N. Chartterjee[1], P. Rosenberg[1], N. Rothman[1], P. Brennan[2], W.H. Chow[1], M. Tucker[1], S. Chanock[1].* 1) Division of Cancer Epidemiology and Genetics, National Institutes of Health, Bethesda, MD, USA; 2) International Agency for Research on Cancer (IARC), Lyon, France; 3) Division of Cancer Prevention, Department of Health and Human Services, National Cancer Institute, Bethesda, MD, USA; 4) Division of Urologic Surgery, Washington University School of Medicine, St. Louis, MO, USA; 5) Epidemiology Research Program, American Cancer Society, Atlanta, GA, USA; 6) Department of Chronic Disease Prevention, National Institute for Health and Welfare, Helsinki, Finland; 7) Department of Urology, University Hospital Motol, Prague, Czech Republic.

A recent genome-wide association study (GWAS) of renal cell carcinoma (RCC) observed a promising association locus on 12q24, which maps to SCARB1, scavenger receptor class B member 1 gene. SCARB1 codes for a cell surface receptor that binds to high-density lipoprotein cholesterol (HDL-C) and mediates HDL-C uptake. In follow-up, we performed fine mapping of the 12q24 locus, by genotyping 53 single nucleotide polymorphisms (SNPs) in 2,278 cases and 3,719 controls from four independent studies in United States, Finland, and Central and Eastern Europe. We used reference panels from 1000 Genomes Project and Phase III HapMap to impute genotypes of common SNPs across the region. Genome-wide level of significance was found with two SNPs that were in high linkage disequilibrium (LD). These variants could either tag or potentially cause the principal risk of the locus, but the high LD makes it difficult to assign specific causality. While this SNPs marks a promising association, further confirmatory work is required to establish its association with RCC risk. SCARB1 is a novel gene associated with RCC risk that may lead to new etiological insights.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1013F

Array CGH Analyses in a Patient with Multifocal Recurrent Meningioma. *Z. Yilmaz Celik[1], Y.K. Terzi[1], N. Altinors[2], F.I. Sahin[1].* 1) Department of Medical Genetics, Baskent University Faculty of Medicine, Ankara, Turkey; 2) Department of Neurosurgery, Baskent University Faculty of Medicine, Ankara, Turkey.

Multiple meningiomas occur in 1–9% of all meningiomas and refer to a condition in which at least two tumors are present at different sites in one patient without neurofibromatosis. Although meningioma is mostly a benign tumor, it is a very unusual tumor because of its many cytogenetic, biological, and clinical aspects. Several studies have described genetic alterations associated with meningioma initiation. Here we report a case of a sporadic recurrent multiple meningioma. A 39 year old male has been operated three times for multiple intracranial meningiomas at a different center during a time frame of nearly 20 years. He was brought to our attention by his relatives due to inability to walk for the last two weeks and inability to speak for the last six months. The attending hospital notified the family as no further treatment is indicated for this patient. The cranial MR showed multiple intracranial meningiomas including four different localizations. The patient was operated upon with the total removal of the two biggest tumors causing edema and shift of the midline structures. The histoptahologic evaluation showed atypical meningioma with extensive brain invasion of the tumor in the left frontoparietal lobe. Cytogenetic study revealed a complex karyotype from solid tumor tissue cultures taken at the time of surgery, and was reported as; 38~39,XY,-1,der(1),der(4),-5,-6,-7,der(7),-8,-9,der(9),del(11q),-12,der(12),-14,-17,del(18q),-21,der(21),-22,inc[cp8]/46,XY[1]. Roche NimbleGen Human CGH 3x720K Whole-Genome Tiling v3.0 Array was used for aCGH analysis on genomic DNA isolated from the sample. The data were analyzed by using web based Genoglyphix program (Signature Genomics). A large number of chromosome copy gains and losses were observed. Copy gains were identified on chromosomes 6, 9, 10, 15 and 22. Copy losses were identified on chromosomes 1, 6, 7, 9, 10, 14, 17, 18 and 22. According to our results, 30 chromosome regions were identified as clinically significant. 12 were of unclear clinical significance and 6 were regarded as benign. Our aCGH results mostly confirmed our conventional cytogenetic results. A conventional karyotype needs live cells and is a laborious process during both cultures and analyses, aCGH promises a more definitive and explanatory result in similar patients.

1014W

Haplotype-resolved sequencing and structural analysis of the HeLa cancer genome. *A.C. Adey, J.O. Kitzman, J. Burton, J. Shendure.* Genome Sciences Dept, University of Washington, Seattle, WA.

We present the first haplotype-resolved sequence of a cancer genome. A PubMed search of “HeLa” returns over 68,000 hits. However, a comprehensive genomic analysis of the oldest and arguably most prevalent human cancer cell line has yet to be performed. We sequenced the HeLa genome to ~88X shotgun coverage in order to accurately identify variants in the context of extreme aneuploidy. Next, we constructed and sequenced nearly 300 fosmid clone pools, each containing ~2 percent representation of the genome, thereby providing islands of haplotype-specific coverage. In addition, we performed single chromosome flow-sorting and pooled sequencing in order to boost our ability to determine haplotype relationships amongst germline variants and somatic mutations across entire chromosomes, as well as to accurately catalog chromosomal rearrangements. Shotgun, fosmid clone pool, chromosome sorting, and 4–40 Kb mate-pair data were integrated to generate the first haplotype-resolved sequence of a cancer genome. The genome of HeLa may yield insights into the biological basis for its persistence and robust growth properties. Furthermore, we are using the haplotype information to estimate the timing and order of somatic events over the 60 year history of HeLa. Finally, it has been previously shown that a number of chromosomal fragments are present in populations of HeLa cells at varying frequencies. To further investigate this, we performed flow sorting and shotgun sequencing of single nuclei to characterize the content and frequencies of these abnormal chromosomes. The methods presented here may prove broadly useful for improving the comprehensiveness and resolution at which we characterize cancer genomes.

1015T

Role of somatic mutations in the etiopathogenesis of Maffucci Syndrome. *M. Amyere[1], M. Limaye[1], N. Mulliken[2], J.B. Dompmartin[3], A. Enjorlas[4], O. Kaitila[5], I. Docquier[6], P-L. Godfraind[7], L-M. Boon[1,8], M. Vikkula[1,9].* 1) Human Molecular Genetics, de Duve Institute, Brussels, Belgium; 2) Department of Plastic Surgery, Children's Hospital and Harvard Medical School, Boston, MA, USA; 3) Department of Dermatology, Université de Caen Basse Normandie, CHU Caen; 4) Consultation des Angiomes, Hôpital Lariboisière, Paris, France; 5) Department of Clinical Genetics, Helsinki University Central Hospital, Helsinki, Finland; 6) Division of Orthopedic Surgery, Cliniques Universitaires St-Luc, Université catholique de Louvain, Brussels, Belgium; 7) Division of Pathology, Cliniques Universitaires St-Luc, Université catholique de Louvain, Brussels, Belgium; 8) Center for Vascular Anomalies, Division of Plastic Surgery, Cliniques Universitaires St-Luc, Université catholique de Louvain, Brussels, Belgium; 9) Walloon Excellence in Lifesciences and Biotechnology (WELBIO), Université catholique de Louvain, Brussels, B-1200, Belgium.

Maffucci Syndrome is a non-hereditary enchondromatosis syndrome, characterized by multiple benign central cartilaginous tumors, accompanied by vascular tumors in the form of spindle cell hemangiomas. Enchondromas can undergo malignant transformation, and patients show an increased incidence of cancers, such as glioma. The etiology of the Maffucci syndrome is unknown, and differs from the Ollier disease. Mutations in isocitrate dehydrogenases IDH1 and IDH2 have been described in gliomas, as well as solitary enchondromas, making them attractive candidates to play a role in the pathogenesis of the Maffucci syndrome. Targeted sequencing of the IDH1 and IDH2 genes identified mutations in 3\4 of tissues resected from our series of patients with the Maffucci syndrome. The mutations were somatic and mosaic, affecting particular protein residues: R132 (C) in IDH1, and R172 (G) in IDH2. Yet, patients with inherited D-2-hydroxyglutaric aciduria, which is caused by germline IDH2 mutations, do not show an increased incidence of tumors, suggesting that IDH mutations are not necessary and/or sufficient to cause tumors. Therefore, to identify the origin of Maffucci syndrome, we have performed Affymetrix SNP-array analyses to look for chromosomal alterations in a series of tissues vs. blood, from 10 Maffucci patients and two patients with Ollier disease. While chromosomal anomalies were absent in constitutional DNA, several CNVs were identified in tissues, including losses and gains. These identified chromosomal regions are shared in subsets of Maffucci patients, and are thus candidates to harbor causative mutations. Exome sequencing of these samples will allow us to further analyze these loci. (miikka.vikkula@uclouvain.be).

1016F

Association of polymorphisms in NNK metabolizing genes and lung cancer risk. *J.L. Engle[1], A.K. Hull[2], G. Liu[1], A.S. Berg[1], C.J. Gallagher[1,3], P. Lazarus[1,3], J.E. Muscat[1].* 1) Department of Public Health Sciences, Penn State College of Medicine, Hershey, PA, United States; 2) Biology Department, Lincoln University, PA, United States; 3) Department of Pharmacology, Penn State College of Medicine, Hershey, PA, United States.

Lung cancer is the leading cause of cancer related deaths among both men and women in the US. While smoking is estimated to account for over 80% of all lung cancer cases, only about 15% of lifetime smokers develop this malignancy. This discrepancy among smokers suggests that individual susceptibility to tobacco carcinogens may be modulated by genetic factors. One of the major carcinogens in tobacco smoke is the nicotine derived nitrosamine 4- methylnitrosamino-1-(3-pyridyl)-1-butanone (NNK). In order to be detoxified, NNK must first be reduced to another potent lung carcinogen 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol (NNAL), which can be glucuronidated by the uridine glucuronosyltransferase (UGT) enzyme family and excreted. Metabolism of NNK to NNAL is catalyzed by four enzymes, aldo-keto reductase 1C1 and 1C2 (AKR1C1 and AKR1C2), carbonyl reductase 1 (CBR1), and hydroxysteroid dehydrogenase 11β1 (HSD11β1). The current study focuses on polymorphisms that provide high coverage in the NNK reduction pathway; *AKR1C1*, *AKR1C2*, *CBR1*, and *HSD11β1* genes. We hypothesized that the polymorphisms in the genes that metabolize NNK to NNAL may be associated with lung cancer risk in individuals exposed to tobacco smoke. We genotyped 448 whites with histologically-confirmed lung cancer (no past history of other tobacco-related cancers) and 786 white controls using a custom 48-SNP Illumina GoldenGate genotyping panel and Taqman real-time PCR. Tag SNPs were chosen in the four NNK metabolizing genes using HapMap genotypes in haploview software (tagger option) with r^2>0.8. SNP association analysis was performed adjusting for known lung cancer confounders (age, sex, and pack-years of smoking) and Hardy-Weinberg Equilibrium analysis was conducted for each SNP. Overall significant associations were observed for three SNPs from the *AKR1C1*, *HSD11β1*, and *CBR1* genes and risk for lung cancer. Smoking and sex stratification and histological subgroup analysis also showed significant findings associated with selected SNPs from all four NNK metabolizing genes. Comprehensive haplotype analysis on all NNK metabolizing genes is ongoing. The identification of genetic variants that regulate the activation and detoxification of tobacco smoke carcinogens is necessary to identify high-risk smokers and develop targeted interventions.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1017W

Myeloproliferative neoplasms and somatic mosaicism in the 23andMe participant community. *D.A. Hinds[1], K.E. Barnholt[1], J.L. Zehnder[2], A.K. Kiefer[1], C.B. Do[1], N. Eriksson[1], J.L. Mountain[1], U. Francke[1], J.Y. Tung[1], R.L. Levine[3], R.A. Mesa[4], J.R. Gotlib[5].* 1) 23andMe, Inc., Mountain View, CA; 2) Department of Pathology and Department of Medicine/Hematology, Stanford University School of Medicine, Stanford, CA; 3) Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, NY; 4) Division of Hematology & Medical Oncology, Mayo Clinic, Scottsdale, AZ; 5) Department of Medicine/Hematology, Stanford University School of Medicine, Stanford, CA.

Background: Myeloproliferative neoplasms (MPNs) are disorders that result in unregulated overproduction of one or more myeloid blood cell types by the bone marrow. Polycythemia vera (PV), essential thrombocythemia (ET), and primary myelofibrosis (PMF) comprise the three classic MPNs. A somatic *JAK2* mutation, V617F, is present in 95% of PV and 50–60% of ET and PMF patients. Past work has identified germline variation in *JAK2* associated with risk of developing a V617F-positive MPN. **Methods:** We have recruited a web-based participatory cohort of patients with MPNs to better understand the genetic basis of these conditions. We have enrolled and collected saliva samples from more than 600 participants, including the following self-reported diagnoses: systemic mastocytosis (n=130), ET (n= 115), PV (n=107), PMF (n=50), chronic myeloid leukemia (n=70), and 79 with overlapping diagnoses. Participants have been genotyped using a derivative of the Illumina OmniExpress with additional custom content, including probes for V617F. **Results:** We can detect V617F with high specificity and sensitivity, indicating that a portion of the DNA we extract is of myeloid origin. We also observe large-scale chromosomal alterations previously seen in MPNs, including uniparental disomy of chromosome 9p. We detect V617F and larger chromosomal alterations at low frequency in the broader 23andMe community. These variants have been associated with increased cancer risk and we are developing strategies for communicating this information and conducting follow-up studies with our participants. In genome-wide association analyses, we replicate the known association between V617F and germline variation in *JAK2* (rs12340895: odds ratio=3.2, P=3.2e-41). We also detect a novel association between MPNs—ET, PV, and PMF—and rs2853677 (odds ratio=1.5, P=8.7e-08) in telomerase reverse transcriptase (*TERT*). Variation in *TERT* has previously been associated with multiple non-hematological cancers, and our lead SNP is in linkage disequilibrium with the previously associated *TERT* variants. The high-risk allele of rs2853677 is also predictive of V617F status among 23andMe participants who were not recruited into the MPN community (odds ratio=1.6, P=0.0014). **Conclusion:** We have identified a germline variant in *TERT* which is a novel predisposition allele for MPNs. These results demonstrate the potential for web-based recruitment methods to contribute to genetic research for uncommon diseases.

1018T

Acute myeloid leukemia associated with chromosome 6: Clinical and phenotypical implications in two pediatric patients with chromosome trisomy and translocation. *L. Hurtado-Hernandez[1], J. M. Aparicio[2,6], F. L. Cuellar[3], M. A. L. Cubillo[4], M. A. H. Garrido[5], S. M. Chatelain[7].* 1) Cytogenetics; 2) Genetics; 3) Urology; 4) Physical therapy; 5) Oncohematology, Hospital para el Nino Poblano, Pue, puebla; 6) Estomatology, Benemerita Universidad Autonoma de Puebla; 7) Biotechnology, Universidad Autónoma Metropolitana, Mexico.

Chromosome aberrations are considered alterations in the chromosome number or structure. They are mainly considered due to gametogenesis inborn error or during the zygote first cellular divisions. Among 4617 chromosomal studies performed during 19 years (from 1992 to 2011), at Hospital Para El Niño Poblano in México, 34.6% (1596 patients) had chromosomal alterations. Among these studies population, a male and female pediatric patients are described, with 6;9 translocation and trisomy of chromosome 6, were chromosome changes are classified as structural or numeric alterations respectively, where a number of leukemias has been associated with specific chromosomal translocations. Both cases were described in this study analyzing their hematological, clinical features, medical treatments and prognosis.

1019F

Correlating Genetic Alterations From Array Comparative Genome Hybridization And Next-Generation Sequencing With Histology And Clinical Outcome In Melanocytic Neoplasms. *Y. Liu, Y. Zhou, E. George.* Dept Pathology, Univ Washington, Seattle, WA.

The incidence of melanoma in the United States is increasing and is about one in seventy currently. Due to limited systemic therapies for treating advanced stages of melanoma, early detection is the key to improve survival. Although particular melanoma subtypes are clinically and histopathologically distinct, they do not provide independent prognostic value. Molecular studies revealed some genetic defects of several genes and pathways in melanoma, however, the etiology of melanoma development is still largely unknown. The goal of this study is to use of high-resolution, genome-wide approaches to identify genomic aberrations that contribute to melanoma transformation, development and progression and correlate these genetic findings with the histology and clinical outcome in hope of generating a molecular classification scheme for melanocytic neoplasms. The FFPE melanoma samples used in this study include malignant melanoma, metastasized melanoma, atypical nevi and benign nevi. We employed both whole genome scan using CCMC 180K CGH+SNP array and panel of actionable gene sequencing using Next-Generation Sequencing in this study. The analyses and results of these samples will be presented and its clinical utility will be discussed.

1020W

Molecular profiling of the therapy-resistant clones in the recurrent Embryonal Carcinomas of the testes based upon FISH and NGS of the isolated chromosomes. *R. Malecki[1,2], M. Malecki[1,3,4].* 1) PBMEF, San Francisco, CA; 2) SFSU, San Francisco, CA; 3) WUHS, Pomona, CA; 4) UW, Madison, WI.

Molecular profiling of the therapy-resistant clones in the recurrent Embryonal Carcinomas of the testes based upon FISH and NGS of the isolated chromosomes **Introduction.** Orchidectomy - the radical surgery of the therapy-resistant, clinically advanced cancer of the testes, leads unequivocally to infertility. Fertility sparing surgery, followed by systemic radiation or chemotherapy, may lead to genetic mutations manifested in the off-spring. Therefore, identification of the unique biomarkers is essential for specific diagnosis [1], including distinguishing between different types of testicular cancers, as the first step towards forging targeted therapy, capable of overcoming the resistance to therapeutics [2]. The specific aim of this project was to seek clones with the unique biomarkers of the cancer initiating stem cells in the recurrent, therapy-resistant embryonal carcinoma of the testes. **Methods.** Based upon the CDS library, we genetically engineered fluorescent and superparamagnetic scFvs targeting cancer stem cells: SSEA-3, SSEA-4, TRA-1-60, TRA- 1-81. Using these scFvs, we isolated single, living cancer stem cells for clonal expansion. We also designed the chromosome-specific probes and used them to isolate single chromosomes from the selected cells clones. We collected the cells from the patients suffering advanced stages of the therapy-resistant, recurrent embryonal carcinomas of the testes. **Results.** The cells were isolated using FACS and MACS as they expressed biomarkers of the embryonic stem cells SSEA-4 and TRA-1-60, confirmed with the flow cytometry, confocal microscopy, and nuclear magnetic resonance. The isolated cells retained high viability and were suitable for clonal expansion. From those clones, the isolated chromosomes helped to identify genetic mutations with FISH and NGS. The clonally expanded cells and amplified transcriptomes facilitated revealing expression of the genes specific for pluripotency Oct3/4, Sox2, Nanog. **Conclusion.** The novel biotechnology facilitated identifying, within the therapy resistant embryonal carcinomas, the clones of cancer stem cells carrying sets of very specific biomarkers. They may serve not only for the fine diagnosis, but also for designing targeted therapy. **Bibliography.** [1] Malecki M, Szybalski W. Gene. 2012 Feb 1;493(1):132-9. Epub 2011 Dec 1. [2] Mueller T et al. Tumour Biol. 2006 27(2):71–83. Epub 2006 Mar 24.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1021T

Detection of Large Rearrangements in *BRCA1* and *BRCA2* by Microarray-CGH. *D. Mancini-DiNardo, T. Judkins, N. Woolstenhulme, C. Burton, S. Chen, J. Schoenberger, M. Ryder, A. Murray, M. Balzotti, N. Gutin, J. Holladay, J. Craft, C. Colvin, J. Trost, L.A. Burbidge, C. Arnell, E. Rosenthal, B. Roa.* Myriad Genetic Laboratories, Inc. Salt Lake City, UT.

Conventional Sanger sequencing technology is capable of detecting the vast majority of mutation types in *BRCA1* and *BRCA2*, but it cannot detect large rearrangements (e.g. deletions and duplications). The proportion of clinically significant defects in *BRCA1/2* attributable to genomic rearrangements is estimated to be 6–10%. The Microarray BRAC*Analysis* Rearrangement Test (microarray BART) is a newly validated testing platform based on oligonucleotide array Comparative Genomic Hybridization (oligo aCGH) technology, which is able to detect large rearrangement (LR) mutations. We validated microarray BART by blinded analysis of 313 blood and buccal DNA samples that were previously examined for LR mutations in *BRCA1/2* using our clinically validated quantitative multiplex PCR assay. Results of our method comparison showed that microarray-CGH was 100% concordant with the original quantitative multiplex PCR results, including the precise identification of 37 mutation positive samples. Furthermore, these validation studies correctly identified two instances of an Alu insertion in *BRCA2* specific to the Portuguese population (156_157insAlu), among the 313 under investigation. Microarray-CGH offers better resolution of LR mutations due to the greater probe number and more comprehensive gene coverage. This methodology is not as sensitive to technical artifacts caused by single nucleotide polymorphisms at probe binding sites. Furthermore, this platform is amenable to high throughput processing, which can enhance testing capacity for *BRCA1/2* LR mutations. This feature is particularly significant given the recently published National Comprehensive Cancer Network (NCCN) guidelines that recommend comprehensive testing, which encompasses full *BRCA1/2* sequencing and detection of large gene rearrangements in patients who meet HBOC testing criteria. In all, microarray-CGH is a clinically validated test that can reliably detect the presence of deletions and duplications in *BRCA1/2*. Since clinical microarray BART testing was offered in May 2012, limited data have been generated mostly on high-risk patients for whom sequencing and LR testing were run concurrently. Our collective experience has shown that 6–10% of mutation positive samples had a LR, with ~88.4% in *BRCA1* and ~11.6% in *BRCA2*. We also detected two Portuguese insertions. Results will be updated as additional patients undergo clnical microarray-CGH testing for *BRCA1/2* large rearrangements in the coming months.

1022F

Effect of insecticide exposure in the generation of alterations in MLL, a gene associated with acute lymphoblastic leukemia. *M.P. Navarrete Meneses[1,2], M. Betancourt[3], E. Bonilla[3], M. Altamirano[4], A. Reyes[1], P. Pérez-Vera[1].* 1) Laboratorio de Cultivo de Tejidos, Instituto Nacional de Pediatría, Mexico City, Mexico; 2) Posgrado en Ciencias Biológicas, UNAM, Mexico City, Mexico; 3) Departamento de Ciencias de la Salud, Universidad Autónoma Metropolitana-Iztapalapa, Mexico City, Mexico; 4) FES-Zaragoza, UNAM. México, D.F, Mexico City, Mexico.

Acute lymphoblastic leukemia (ALL) is the most common childhood cancer. It is characterized by the presence of numerical and structural chromosomal abnormalities. Translocations involving MLL gene are detected in 80% of patients under 1 year old; it is proposed they arise in utero. Development of ALL and the uterine exposure to pesticides have been reported in numerous epidemiological studies; however, there is limited biological evidence supporting this association. The aim of this study was to detect alterations in MLL gene induced by two extensively used insecticides, permethrin and malathion, in human lymphocytes in vitro. Method: Lymphocytes from a healthy volunteer were cultured for 72h. Chemical grade permethrin/malathion were added (100–300μM) for the last 24h. Etoposide and insecticides solvents were used as positive and negative controls, respectively. Alterations in MLL gene were analyzed by FISH. Results: Both pesticides showed increased percentage of cells with numerical and structural alterations, compared to negative controls; the highest percentage of damage was detected with permethrin treatment. Even though damage was found in all samples, more complex abnormalities were observed in the cells treated with pesticides. Discussion: Numerical alterations, not previously reported, were detected using these insecticides. Some of the structural alterations found in this study have been observed in cells exposed to etoposide. Direct damage found in MLL gene using permethrin/malathion, resembles the effect described for etoposide, a well-known leukemogenic agent. Acknowledgement to CONACyt MPNM-385279.

1023W

Genomic analyses of paired diagnostic/relapse childhood acute lymphoblastic leukemias identify relapse-associated genetic features of pathogenetic and clinical importance. *L. Olsson[1], A. Castor[2], M. Behrendtz[3], A. Biloglav[1], E. Forestier[4], K. Paulsson[1], B. Johansson[1].* 1) Clinical genetics, University and Regional Laboratories, Skåne University Hospital, Lund University, Lund, Sweden; 2) Department of Pediatrics, Skåne University Hospital, Lund, Sweden; 3) Department of Pediatrics, Linköping University Hospital, Linköping, Sweden; 4) Department of Clinical Sciences, Pediatrics, University of Umeå, Umeå, Sweden.

Despite high cure rates of childhood acute lymphoblastic leukemia (ALL), a subset of the patients relapse and treatment failure remains a significant problem. The mechanisms underlying such unfavorable events are unclear. The fact that relapses occur not only in high risk groups but also in the standard risk group indicates that the present prognostic factors are insufficient. Therefore, it is important to identify specific relapse-associated genetic features already at the time of diagnosis in order to ensure proper risk stratification. We performed single nucleotide polymorphism (SNP) genomic array analysis, using Illumina 1M SNP arrays, on 77 ALL cases with events (60 cases with relapse, 9 with resistant disease, and 8 with induction failure) to ascertain molecular lesions associated with relapsed and resistant disease. The patients are part of a large series of patients (n = 307) diagnosed and treated according to the Nordic Society of Pediatric Hematology and Oncology ALL protocols between 1992 and 2011 at the Departments of Pediatric Oncology and Hematology, Lund and Linköping University Hospitals, Sweden. Among the 77 cases, 17 paired diagnostic/relapse samples, 41 diagnostic samples from patients that subsequently relapsed, and 4 relapse samples (without corresponding diagnostic samples) were informative. Fifteen cases were excluded due to either too low DNA concentration or because of too many normal cells in these samples. The paired samples were genetically related, identical, or harbored additional alterations in the relapse sample. However, 29% lacked some of the copy number aberrations present at diagnosis, suggesting the presence of a preleukemic clone from which the relapse clone evolved. No single aberration was linked to only relapse, but deletions involving IKZF1 (24.2% vs. 8.2%; P = 0.003) and NR3C1 (6.5% vs. 0.5%; P = 0.019) were more common compared with patients remaining in complete remission. SNP array analysis is not only an excellent tool for genome-wide profiling of cryptic structural alterations in cancer cells but also provides insights into the origin of relapse clones. Many of the changes identified at relapse are probably present in minor clones already at the time of diagnosis, and detection of these clones at diagnosis should result in improved risk stratification and hopefully in better outcome.

1024T

Loss of chromosomes is the primary event in near-haploid and low hypodiploid acute lymphoblastic leukemia. *S. Safavi[1], E. Forestier[2], I. Golovleva[2], G. Barbany[3], K. H.Nord[1], A. V. Moorman[4], C. J. Harrison[4], B. Johansson[1], K. Paulsson[1].* 1) Clinical Genetics, Laboratory medicine, Lund, Sweden; 2) Department of Medical Biosciences, Medical and Clinical Genetics,University of Umeå,Sweden; 3) 3Department of Molecular Medicine and Surgery and Center for Molecular Medicine, Karolinska Institutet, Stockholm, Sweden; 4) Leukaemia Research Cytogenetics Group, Northern Institute for Cancer Research, University of Newcastle, Newcastle Upon Tyne, UK.

Near-haploid (23–29 chromosomes) and low hypodiploid (HoL; 33–39 chromosomes) acute lymphoblastic leukemia (ALL) is associated with a poor prognosis. Both subtypes are characterized by chromosomal loss and duplicated stemlines, the latter which may lead to misclassification as high hyperdiploid ALL. Nothing is known about the pathogenetic consequences and mechanisms underlying near-haploidy and HoL. In the present study, we used SNP array analysis to investigate 12 cases of near-haploid (n=8) and HoL (n=4) ALL. A characteristic pattern of monosomies/uniparental isodisomies and retained heterodisomies/tetrasomies was identified, with the latter involving chromosomes X/Y, 14, 18, and 21 in near-haploid cases and chromosomes 1, 5, 6, 8, 10, 11, 14, 18, 19, 21, and 22 in HoL. All cases displayed widespread loss of heterozygosity (LOH). Two cases were initially diagnosed as high hyperdiploid childhood ALL and both relapsed; notably, these were unambiguously identifiable as near-haploid with the SNP array analysis, showing the potential of this technique for correct risk stratification. Cryptic genetic events were found in five cases, including four CDKN2A deletions. The pattern of deletions and LOH indicated that microdeletions occurred in a subsequent step in one case, strongly suggesting that the chromosome loss is the primary event in near-haploid and HoL ALL.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1025F

Detection of the t(8;14) chromosomal rearrangements in paraffin embedded tissue sections of Burkitt Lymphoma patients with quantitative polymerase chain reaction (PCR) and fluorescence in situ hybridization (FISH) methods. *N. Selvi[1], M. Hekimgil[2], B. Tezcanli Kaymaz[1], S. Aktas[2], N. Ozhan[2], C. Gunduz[1], B. Kosova[1], Y. Dodurga[3], N. Topcuoglu[1].* 1) Ege University, Medical Biology Department, Medicine school, izmir Turkey; 2) Ege Univiersity, Medical school, Department of pathology, izmir, Turkey; 3) Pamukkale University, Medical Biology and genetics Department, Medicine school, Denizli, Turkey.

Burkitt lymphoma (BL) is an aggressive B-cell lymphoma, with a characteristic clinical presentation, morphology, and immunophenotype. In hemato-oncologic malignancies, determination of the specific tumor chromosomal translocations and the gene amplifications has an ascending importance in view of diagnosis and prognosis of these lesions. In the current study, 17 paraffin-embedded tissue samples, which indicate the lymphoid hyperplasia feature, comprised control group. Total 26 samples were evaluated, the study group of which were composed of 18 paraffin sections, four fresh tissue, two peripheral bloods and two imprint samples diagnosed with Burkitt' lymphoma. The t(8;14) chromosomal rearrangement in the samples was investigated in terms of retrospectives using “expand long template” PCR (LD-PCR) and fluorescent in situ hybridization (FISH) molecular techniques and the results obtained using the two methods were compared with each other. t(8;14) translocation was investigated by expand long PCR and FISH methods in cases diagnosed as Burkitt lymphoma. Fresh tissue samples of three cases, imprint samples of two cases, and peripheral blood samples of two cases were determined positive for MYC/lgH rearrangement. MYC/lgH rearrangement was determined positive in paraffin sections of two samples. The paraffin tissue samples of 16 patients could not be evaluated due to DNA degradation. MYC/lgH rearrangement was determined as positive by FISH analysis in three fresh tissue samples, two imprint and two peripheral blood samples of our cases. In paraffin tissue samples, 9 cases were found positive for t(8;14) translocation and partial trisomy of 8th and 14th chromosomes was determined in one case. As a result, we think that using both different PCR methods and FISH method will facilitate the identification of tumor's genetic characterization. In conclusion, we think that basic molecular methods will bring new dimensions to current therapeutic approaches, also developments in molecular diagnosis will accelerate the step of individual specific treatment methods.

1026W

Molecular changes in Lynch syndrome associated and sporadic ovarian carcinomas. *A. Tieva[1], S. Kaur[1], L. Renkonen-Sinisalo[2], H.J. Järvinen[2], J-P. Mecklin[3], R. Bützow[4], P. Peltomäki[1].* 1) Department of Medical Genetics, University of Helsinki, Helsinki, Finland; 2) Second Department of Surgery, Helsinki University Central Hospital, Helsinki, Finland; 3) Department of Surgery, Jyväskylä Central Hospital, Jyväskylä, Finland; 4) Departments of Obstetrics, Gynecology & Pathology, Helsinki University Central Hospital, Helsinki, Finland.

STATEMENT OF PURPOSE Epithelial ovarian cancer has the highest fatality rate of all gynecological malignancies among women worldwide. Increased understanding of the pathogenesis is required for improvement of the diagnosis, prognosis and treatment of ovarian carcinoma. Ovarian carcinoma is overrepresented in Lynch syndrome (LS). Our aim was to classify LS associated and sporadic ovarian carcinomas by studying epigenetic, genetic and protein expression alterations. METHODS USED Normal and tumor samples of all available LS associated ovarian carcinomas were collected (21 paired samples) through a nation-wide registry. Sporadic ovarian tumors (95 cases) were chosen to match with tumors from LS families relative to histological type (serous, endometrioid and clear cell) and mismatch repair (MMR) status of the tumor. We used microsatellite instability (MSI) and MMR protein expression to study MMR defects. Epigenetic changes were studied by the methylation-specific multiplex ligation-dependent probe amplification test (MS-MLPA, ME001-C1 tumor suppressor-1 test) which detects methylation of 24 tumor suppressor gene (TSG) promoters. In addition, the samples will be examined for global hypomethylation, p53 expression, and K-RAS mutations. SUMMARY OF RESULTS Among LS associated ovarian carcinomas, 89% (16/18) showed loss of MMR protein expression corresponding to the germline mutation, and high-degree MSI was present in all the cases. In sporadic cases all four MMR proteins were expressed in the majority of cases (87%), and MSI was detected only in 5% of the cases. Ovarian carcinomas showed inactivation of certain TSGs by promoter methylation. The most frequently methylated TSGs in the entire set of ovarian tumors were RASSF1 (62%), CDH13 (60%), CDKN2B (23%), APC (22%), TP73 (16%) and GSTP1 (14%). Methylation frequencies of these six genes differed remarkably between LS associated and sporadic ovarian carcinomas as well as between different histological subtypes of ovarian cancer. Methylation of CDKN2B was specific to LS associated ovarian carcinoma (81%) in contrast to sporadic cases (10%). The methylation frequency was very low (4 to 12% depending on the TSG) in serous ovarian carcinomas compared to clear cell and endometrioid types (sporadic and LS combined). CONCLUSION MMR defects are significantly more common in LS vs sporadic ovarian carcinoma. TSG promoter methylation is frequent with patterns specific to LS vs sporadic and histological type.

1027T

A Comprehensive Genetic and Epigenetic Study of Hepatocellular Carcinoma with Microarrays. *G. Zhang[1,7], C. Wu[2,7], D. Yu[2], J. Chang[2], W. Sun[3,1], M. Li[4], L. Liang[5], Y. Li[1,3,6], D. Lin[2].* 1) Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC; 2) State Key Laboratory of Molecular Oncology and Department of Etiology and Carcinogenesis, Cancer Institute and Hospital, Chinese Academy of Medical Science and Peking Union Medical College, Beijing, China; 3) Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC; 4) Department of Biostatistics and Epidemiology, University of Pennsylvania School of Medicine, Philadelphia, PA; 5) Department of Epidemiology and Department of Biostatistics, Harvard School of Public Health, Boston, MA; 6) Department of Computer Science, University of North Carolina at Chapel Hill, Chapel Hill, NC; 7) These authors contributed equally to this work.

Hepatocellular carcinoma (HCC) caused by viral hepatitis is one of the most common cancers in countries where hepatitis is endemic, such as China. Recently, the emergence of various high-throughput technologies have enabled integrated genetic and epigenetic analysis of multiple omics datasets, which holds the promise to identify genetic factors of complex diseases such as HCC. Here we used multiple types of microarrays to obtain SNPs and copy number variations/aberrations (CNV/CNAs) genotypes, DNA methylation, and expressions of protein coding genes and miRNAs from 30 paired primary tumor and normal liver tissues from patients with HCC in China. Our comprehensive genetic and epigenetic analysis has revealed a remarkable number of alterations in HCC samples. In total, we identified 2528 ($p<1.8\times10^{-6}$) differentially expressed transcripts out of 28536 examined and 440 ($p<7.1\times10^{-8}$) differentially methylated probes out of 707812 in tumor versus normal liver tissue using conservative Bonferroni correction. Pathway enrichment analysis detected several significant pathways enriched of genes highly expressed/methylated in tumor tissues including cell cycle, DNA repair, inflammatory and several signaling pathways. SNP array data also identified frequent CNAs, such as gains at 1q, 5, 7, 17p, and losses at 1p, 4q, 6q, 16, 17q, 21 in tumors. Combining the expression, methylation and genetic data, we were able to illustrate the genetic and epigenetic effects on gene expression patterns that are specific in tumor or normal tissues. Our integrated analysis lays the ground for further molecular studies of HCC by identifying novel genetically and epigenetically regulated genes potentially involved in HCC development or progression.

1028F

Congenital Leukemia with Insertion of MLL into 19p13.3 and a Deletion of Half a Megabase of 19p13.3: Case Report. *H. Al-Kateb, M. Evenson, S. Kulkarni.* Pathology and Immunology, Washington University School of Medicine, St. Louis, MO.

Congenital and neonatal leukemia represent less than 1% of all childhood leukemia. The incidence of leukemia in infants under 1 month of age is 4.7 per million live births per year. The vast majority of cases of congenital and neonatal leukemia are acute myeloid leukemia (AML). Mixed-lineage leukemia gene (MLL) rearrangements occur in approximately 14.4% of cases of which t(11;19) translocation was found in half of the cases. Two translocations were described that fuse MLL to different genes on 19p, the first one is t(11;19)(q23.3;p13.3) and fuses 5' MLL to 3' of mixed-lineage leukemia translocated to 1(MLLT1); and the second one is t(11;19)(q23.3;p13.1) and fuses 5'MLL to 3' elongation factor RNA polymerase II (ELL) gene. We report about a 36-week gestation female infant who was born with blueberry muffin rash, hepatosplenomegaly and coagulopathy. Pathology examination revealed AML with monoctyic differentiation. Initial cytogenetic analysis revealed a t(11;19)(q23.3;p13.1) translocation in all cells examined. However, fluorescence in situ analysis (FISH) using MLL probe revealed that the 3' part of MLL went to chromosome 19. However, FISH studies with E2A at 19p13.3 and subtelomeric probes showed that those loci were intact on the derivative chromosome 19 indicating that 3'MLL was inserted into 19p13.3 and not translocated. Subsequent microarray analysis revealed a half megabase deletion on chromosome 19 within 19p13.1 band and no loss or gain on chromosome 11. The proximal deletion breakpoints were within ELL gene where the first 6 exons were deleted. This corresponds to deleting 289 amino acids of ELL N-terminal. Accordingly the karyotype was modified to 46,XX,ins(19;11)(p13.1;q23). arr 19p13.1(18428039–18961490)x1. In all the previously reported MLL-ELL transcripts, which have been associated with AML, MLL has been fused to ELL exon 2 or exon 3 in one case. Although the deletion of ELL in the present case affected exons 1–6, it preserved the minimal ELL region required for immortalization. This makes ELL a strong candidate for being the partner gene of MLL in this case. However, the possibility that MLL was fused to a different “cancer gene” cannot be ruled out. Work is in progress to map the breakpoints at the base pair level and to precisely determine the partner gene of MLL in this case.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1029W

CYTOGENETIC PROFILES OF 370 CASES WITH RENAL CELL CARCINOMA-UCLA EXPERIENCE. *P.H. BUI, P.N. RAO.* Pathology and Laboratory Medicine, David Geffen School of Medicine, UCLA, Los Angeles, CA.

Renal cell carcinoma (RCC) accounts for approximately 3% of adult malignancies with 90–95% of these arising from the kidney. Cytogenetic analyses have emerged as a powerful tool for diagnosis and classification of RCCs. From 2005 to 2012, tumors from 370 patients (124 females, 246 males), who underwent nephrectomy for RCCs, were cytogenetically analyzed from primary cultures following standard laboratory protocols. Various RCC subtypes were examined which included 247 clear cell (ccRCC- most common), 107 papillary (pRCC), and 16 chromophobe (cRCC) subtypes. From these analyses, we have identified recurrent and specific chromosomal aberrations associated with particular subtypes. 229 of the 247 ccRCCs showed abnormalities involving chromosome 3 (93%). These were deletions of the 3p region, monosomy 3, and unbalanced translocations resulting in deletions of 3p. These aberrations were not seen in other subtypes. In addition, monosomy 8/del 8p (24%), -6/del 6q (19%), -4/del 4q (13%), -10/del 10q (13%), -21 (9.3%), -17/del 17p (8.9%), -5/del 5q (6.9%), -12/del 12q (6.9%), -20 (3.2%), or gains of chromosomes 9 (4.9%), or 18 (4.5%), were more commonly observed in ccRCCs. Recurrent unbalanced translocations t(3;5)(p13;q15) (19%) were also specific to ccRCCs. Among the pRCCs, gains of chromosomes 12 (69%), 16 (65%), 7 (62%), 3 (42%), or 17 (41%), and translocations such as t(16;17)(q24;q11.2), t(1;14)(q32;q32), and t(3;21)(q12;p12) were more common. 10 of the 16 (62.5%) cRCCs exhibited concurrent monosomies of chromosomes 1, 2, 6, 10, 13 and 17. Besides these subtype-specific aberrations, we have also identified several other recurrent anomalies that were common to all RCCs. For instance, loss of Y was the most common aberration in males (55.7%), while loss of the X was observed in 26.6% females. Among autosomes, gains of 5, 8, 19, 20, or 22, or losses of 9, 11, 13, 14, 18, or 22 were found in all RCCs. Our results show that not only are specific cytogenetic abnormalities associated with the various subtypes but also can play an important role in the classification. In addition, such studies aid in providing valuable information of prognostic significance. Our studies, when correlated with the clinically available data, have shown that deletions of 3p and trisomy 5q are generally associated with good prognosis, whereas deletion of 4q/monosomy 4 and loss of chromosome 14 have been associated with a poorer outcome.

1030T

Identifying of cryptic genomic alterations by high-resolution microarrays in LLA-B and LLA-T Mexican patients. *C. Córdova-Fletes[1,2], B. Macias-García[1], A. Lugo-Trampe[1], N. Mendez-Ramírez[3], R. De la Rosa-Alvarado[4], H. Gutiérrez-Aguirre[3], D. Gómez-Almaguer[3], A. Hidalgo[7], R. Salazar-Riojas[3], P. Ruíz-Flores[3], P. Villarreal-Quiroga[2], O. González-Llano[6], A. Rojas-Martínez[1,2], H. Decanini-Arcaute[5], R. Ortiz-López[1,2].* 1) Dept. Bioquímica, CIDICS, Universidad Autonoma de Nuevo Leon, Monterrey, Nuevo león, Mexico; 2) Unidad de Biología Molecular, Genómica y Secuenciación, Centro de Investigación y Desarrollo en Ciencias de la Salud, Universidad Autónoma de Nuevo León, Monterrey, Nuevo León, México; 3) Servicio de Hematología, Centro Universitario contra el Cáncer, Universidad Autónoma de Nuevo León, Monterrey, Nuevo León, México; 4) Departamento de Genética, Facultad de Medicina, Universidad Autónoma de Nuevo León, Monterrey, Nuevo León, México; 5) Hospital CIMA Santa Engracia, Monterrey, Nuevo León, México; 6) Facultad de Medicina, UAC, Torreón, Coahuila; 7) INMEGEN, Cd. de México, México.

Acute lymphoblastic leukemia (ALL) is a malignant disorder that originates in progenitors of T or B lymphocytes. Virtually all patients with ALL acquire genetic changes, which somehow contribute to the formation and expansion of leukemic clones. The T (T-ALL) and (ALL-B) B lineage progenitors are characterized by recurrent numeric and structural alterations affecting cellular pathways in lymphoid differentiation, tumor suppression, apoptosis and lymphoid signaling through alteration of genes such as MLL, MYC, HOX11, HOX11L2, LYL1, TAL1, among others. However, experimental models have suggested that additional changes that include deletions in the genes CDKN2A/CDKN2B and NOTCH1 are required to generate leukemia (T-ALL). In this study, we performed a genetic screening by high-resolution microarrays (Affymetrix 500k platform) to identify frequent changes and characterize new copy number alterations in a group of Mexican patients with acute lymphoblastic leukemia of both subtypes. Interestingly, we observed that the CDKN2A/B deletion was present in 60% of our patients but from both ALL subtypes, suggesting a pathogenic role shared. In addition, we observed a PTPRU (Receptor-type tyrosine-protein phosphatase U) gene deletion in 60% of patients (which 83% were B-ALL). We also highlight interesting changes as cryptic microdeletions of the BCR (22q11.23), TCTA (at 3p21.31) and MLLT3 (in 9p21.3) genes not associated with translocation, whereas amplified regions involved some oncogenes such as NRAS (1p13.2), MYB (6q23.3), MYCl1 (1p34.2). Moreover, very recently some of these changes have been associated to other malignancies. Summarizing, this work is consistent with the hypothesis that implies alternative genomic changes for generating ALL including deletions of CDKN2A/B, TAL1/STIL, NOTCH1 genes or the deletion of 6q16.1 locus, and presents evidence to propose the PTPRU microdeletion as candidate for participate in B-ALL pathogenesis. This project was supported by CONACyT (convocatorias SALUD-2005-01-14125 and SALUD-2004-C02-026).

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1031F

An unusual case of APL in a 14 year old girl with t(15;17)(q22;q21) and t(9;22)(q34;q11.2). *V. Jobanputra[1], C. Freeman[1], S. Hussein[1], M.J. Macera[2], B. Levy[1], V.V. Murty[1], M.L. Sulis[3], G. Bhagat[1].* 1) Department of Pathology, Columbia University Medical Center, New York, NY; 2) New York Presbyterian Hospital, New York, NY; 3) Department of Pediatrics, Columbia University Medical Center, New York, NY.

Acute promyelocytic leukemia (APL) represents approximately 1% of all childhood leukemias and 10% of all childhood AML. Children presenting with APL have a greater degree of leukocytosis than adults and often present with severe bleeding diathesis. We report a case of APL in a 14 year old girl of Mexican descent who presented with new onset headaches and easy bruising, as well as leukocytosis and thrombocytopenia. Her smear showed 44% promyelocytes. Flow cytometry was consistent with a diagnosis of APL. Of interest, G-band karyotype analysis of her bone marrow revealed 46,XX,t(15;17)(q22;q21)[1]/46,idem,t(9;22)(q34;q11.2)[19]. FISH analysis with the PML/RARA probe was positive for the t(15;17)(q22;q21) in 96% of her cells and FISH with the BCR/ABL1 probe also confirmed rearrangement in 95% of her cells. The patient was started on Idarubicin and ATRA and is currently in remission post induction chemotherapy. The hallmark of APL is the presence of the t(15;17)(q22;q21) translocation. Generally, the t(15;17)(q22;q21) is the only abnormality present, however, additional abnormalities and variants involving the 17q21 breakpoint have been reported. These additional abnormalities usually appear to be secondary events. The presence of the t(9;22)(q34;q11.2) with the t(15;17)(q22;q21) is extremely rare. Rare cases with such double translocations have been reported in patients presenting with promyelocytic blast crisis of CML. However, only one previous case of primary APL has been reported in a 38 year old woman who received ATRA and arsenic trioxide, but died of intracranial hemorrhage the following day. We report the second case of primary APL with both a t(9;22)(q34;q11.2) and t(15;17)(q22;q21). Our patient has shown typical symptoms of childhood APL without any unusual clinical presentation and responded well to chemotherapy. Six weeks after her initial cytogenetic analysis only 2/1200 cells positive for PML/RARA were noted and all 2000 cells analyzed were BCR/ABL negative. The presence of a metaphase with the t(15;17)(q22;q21) as the sole abnormality, suggests that a cell line acquired the t(9;22)(q34;q11.2) representing clonal evolution of disease. It does not appear that the presence of the t(9;22) has affected the clinical course of this patient thus far as she responded to ATRA. Analysis of this case should provide valuable insight into this leukemia.

1032W

Integration of cytogenetics and FISH studies for accurate evaluation of chromosome abnormalities. *N. Qin, M. Hibbard.* Cytogenetics, Clarient, Alisoviejo, CA.

Fluorescence in situ hybridization (FISH) provides an important adjunct to conventional cytogenetics in the evaluation of chromosome abnormalities associated with hematologic malignancies, and vice versa. We present two interesting cases to show how the chromosome rearrangements were precisely defined by employing both cytogenetic and FISH tools. In one case, the initial FISH interphase study using the probe D20S108 showed no deletion detected in the long arm of chromosome 20. However, subsequent cytogenetic analysis revealed an obvious deletion of the long arm of chromosome 20. Only the combination of FISH analysis on metaphase cells using the same probe uncovered the discrepancy. In another case, cytogenetic analysis of a bone marrow specimen revealed what appeared to be a deletion of the long arm of chromosome 11 at q23 and additional material of unknown origin attached to the long arm of chromosome 2. Concurrent chromosome analysis on peripheral blood revealed a recognizable translocation between the long arms of chromosomes 2 and 11 at bands of 2q37 and 11q23 [t(2;11)(q37;q23)] due to improved morphology. Only five cases with such a translocation have been reported in the literature; we assumed that our case could be the sixth. To confirm the cytogenetic finding, FISH analysis using an MLL break-apart probe was performed. The FISH signal pattern did confirm the translocation of chromosome of 11 material to chromosome 2; however, the MLL gene was not split apart. The entire MLL gene had moved to chromosome 2, unlike the five reported cases in which an MLL rearrangement was confirmed by FISH. Chromosome karyotypes and FISH images in both cases will be presented. In the second case, comparison of the same abnormalities in different specimen types reveals the difference in morphology, which may be beneficial in the determination of chromosome 11 abnormalities differentiating between deletion and translocation since chromosome 11 has a multitude of translocation partners. In our cases, both chromosome and FISH studies were integral to defining the acquired chromosome abnormalities; neither methodology could be dispensed with.

1033T

Evaluation of FGFR1 gene amplification status and its association with clinicopathological features in breast cancer. *R.M. Rodrigues-Peres[1], J.K. Heinrich[1], L.O.Z. Sarian[2].* 1) CAISM - Women's Hospital, University of Campinas - UNICAMP, Campinas, SP, Brazil; 2) Department of Obstetrics and Gynecology, Faculty of Medical Sciences, University of Campinas - UNICAMP, Campinas, SP, Brazil.

Introduction: The FGFR1 gene, located on chromosome 8p12, is a member of the Fibroblast Growth Factor Receptor Family. It regulates cell proliferation, differentiation, apoptosis and is also known to be involved in angiogenesis and wound repair, which are in turn associated with tumor development. FGFR1 was shown to be amplified in approximately 7–17% of breast tumors. Furthermore, association between FGFR1 amplification and pathological features, as ERBB2 status, estrogen and progesterone receptors, lymph node metastasis, tumor size, grade and decreased survival, was described. However, the differences between the techniques used to assess these associations, which might exhibit disputant results and the lack of studies to entirely support these evidences are critical. Objectives: 1) To evaluate the amplification and aneuploidy status of FGFR1 in breast cancer samples; 2) To correlate the FGFR1 status to ERBB2, C-MYC and CCND1 status in order to access genomic instability of the samples; 3) To verify the association between FGFR1 status and tumor clinicopathological features. Materials and Methods: For this study, 52 cases were obtained from mastectomies. A TMA slide containing these samples was set and used in FISH assays. Homebrew probes were specifically designed for FGFR1 and applied onto samples concomitantly with a commercial probe for chromosome 8 centromere (CEP8, Kreatech). Clinicopathological data were obtained from patients' records. Results: We found 19% FGFR1 amplified cases, of which 7% presented high amplification. We detected a positive correlation between the FGFR1 status and C-MYC and CCND1 status (p = 0.006 and 0.04, respectively). A gradual increase in age was found to FGFR1 amplification when comparing to normal gene status (mean = 86 y.o. X 57.42 y.o.), as well as to lymph node metastasis (mean = 3 affected lymph nodes X 2.5 affected lymph nodes). The % ER status is higher in cases where FGFR1 is amplified, reversely to the PR status, which is higher in normal FGFR1 (ER mean = 83.75% and PR mean = 53.47%). The tumor size was higher in the FGFR1 amplified cases (mean = 4.9cm). Discussion: Our results corroborate previous findings of FGFR1 amplification in breast cancer samples and on the association to the clinicopathological features, also demonstrating its involvement in genomic instability. These results suggest that FGFR1 is an important gene involved in development of breast tumors and might also be a target to new therapies.

1034F

Large scale identification of chromosomal abnormalities in familial breast and ovarian cancer. *H.WH. Yin[1], L.L. Wang[3], E. Tsui[1], Y. Elshimali[2], D.C. Chen[4].* 1) La Sierra University, Riverside, CA; 2) UCLA/Charles Drew University, Los Angeles, CA; 3) Rite Aid Pharmacy, Fontana, CA; 4) Pathology Dept, UCIMC, Orange, CA.

The examination of biochemical pathways of hereditary breast and ovarian cancer has greatly increased our understanding of the pathogenesis of breast cancer. Currently, 22 breast cancer susceptible genes have been identified and can be found in the OMIM database. This list of genes provide candidates for a technology utilizing massive and parallel DNA sequencing in what may potentially be a more comprehensive genetic testing panel for both the familial and somatic forms of breast and ovarian cancers. One challenge for current technology in genetic testing is the relative inefficiency in identifying chromosomal structure abnormalities such as deletion, inversion and/or translocations. Non-allelic homologous recombination (NAHR) and Non-homologous end joining (NHEJ) are major mechanisms responsible for such chromosomal change events. Since the locations for potential chromosomal changes are non-random and limited by the respective location and presence of homologous sequence pairs, a targeted DNA sequencing approach for chromosomal abnormalities in these candidate genes may be feasible. To test our hypothesis, we conducted a bioinformatic analysis of the genomic sequences around all 22 potential breast cancer susceptible genes and globally identified the potential target pairs for NAHR and NHEJ events; we also employed a cross-species sequence comparison to reduce potential false positive results. Overall, 356 potential genomic sequence pairs have been identified and we have designed a solution phase genomic DNA capture system targeting these 356 sites. Specifically, hybridization probes conjugated to magnetic beads have been constructed. The system has been optimized for capturing genomic DNA of tumor specimens with the potential to participate in NAHR and NHEJ mediated events. Lastly, we have conducted parallel pair-ending sequencing of genomic DNA of tumor specimens in order to determine the feasibility of this strategy in identifying chromosomal rearrangement in tumor specimen.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1035W

Homozygous deletions of *CDKN2A* are present in all dic(9;20)(p13.2;q11.2)-positive BCP ALLs and may be important for leukemic transformation. *V. Zachariadis[1], J. Schoumans[2], G. Barbany[1], M. Heyman[3], E. Forestier[4], B. Johansson[5], M. Nordenskjöld[1], A. Nordgren[1].* 1) Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden; 2) Service de Génétique Médicale, Centre Hospitalier Universitaire Vaudois, CH-1011 Lausanne, Switzerland; 3) Department of Women's and Children's Health Karolinska Institutet, Stockholm, Sweden; 4) Department of Clinical Sciences, Pediatrics, University of Umeå, Umeå, Sweden; 5) Department of Clinical Genetics, University and Regional Laboratories, Skåne University Hospital, Lund University, Lund, Sweden.

In B-cell precursor acute lymphoblastic leukemia (BCP ALL) recurrent chromosomal aberrations define unique subtypes of the disease, each with its own prognostic and therapeutic implications. The chromosomal translocation dic(9;20)(p13.2;q11.2) is reported to be present in up to 5 percent of childhood BCP ALL and is associated with a worse survival compared with the most common standard risk aberrations. Previous molecular studies of dic(9;20) have revealed heterogeneous breakpoints on chromosomes 9 and 20, failing to produce a consistent gene fusion as a result. The genetic lesions underlying leukemic transformation and propagation in dic(9;20)-positive BCP ALL thus remains largely unknown. We present a detailed genomic characterization of six childhood BCP ALL harboring dic(9;20)(p13.2;q11.2). Six diagnostic and one “pre-leukemic” sample were investigated, using high resolution single nucleotide polymorphism array, and evaluated for the presence of additional genomic imbalances that may play a role in leukemogenesis. The breakpoints on chromosomes 9 and 20 were clustered, albeit non-identical; three cases had breakpoints within the *PAX5* gene on 9p13.2 and within *C20orf112* on 20q11.2. Importantly, all six cases harbored homozygous deletions of the *CDKN2A gene*, transcribing the two tumor suppressors *P16INK4* and *P14ARF* from alternative reading frames. The *CDKN2A* deletions on the non-rearranged chromosome 9 homologue ranged from 23 to 201 kilobases, thus escaping detection by fluorescence in situ hybridization. A shared proximal breakpoint (chr9: 21998026) of the *CDKN2A* deletions strongly indicates a common underlying mechanism. In addition, one case revealed an apparent progression from a hemizygous to a homozygous deletion of *CDKN2A* during leukemic transformation, indicating that homozygous *CDKN2A* deletions are important for the pathogenesis of dic(9;20)-positive ALL.

1036T

C-MYC and IGH amplification in patients with intermediate high grade B-cell lymphomas. *A. Zaslav[1], M. Bellone[2], F. Lan[3], T. Mercado[1], H. Lee[1], Y. Hu[2].* 1) Dept Pathology, Stony Brook Univ Hosp, Stony Brook, NY; 2) Department of Pathology, Stony Brook University Medical Center, Stony Brook, N.Y; 3) Blood and Bone Marrow Stem Cell Transplantation Program, Stony Brook University Medical Center, Stony Brook, N.Y.

High grade B-Cell lymphomas (BL) with histological and clinical features intermediate between Burkitt lymphoma and diffuse large B-cell lymphoma have recently been classified by the World Health Organization (WHO) (Jaffe, E.S. et.al. (eds.) IARC Press; Lyon, France, 2008. These tumors have large and medium size cells with greater irregulatory of nuclei and more prominent nucleoli than BL. Approximately 35% to 50% have C-MYC rearrangements with non-Ig variants. We present six patients with stage IV Intermediate high grade B-Cell lymphomas demonstrating C-MYC and/or IGH low-level amplification. Standard chromosome analysis, FISH, and/or paraffin-embedded FISH on archived tissue were performed on bone marrow, and/or tumor tissue. The following probes were used: LSI IGH(14q32)/ BCL2(18q21) Dual Color, and the LSI IGH(14q32/ MYC(8q24)/CEP 8 Tri-Color Dual Fusion probes. All cases demonstrated low-level amplification with three to four copies of either C-MYC, IGH or both C-MYC and IGH. Three of six cases also had a t(14;18)(q32;q21). Two of six patients expired shortly after diagnosis. One patient continued to demonstrate progressively worsening refractory disease. Two patients achieved remission after an expanded course of treatment. Gene amplification and the subsequent over expression of the amplified oncogene play an important role in tumor pathogenesis. Few cases of Intermediate tumor patients with C-MYC amplification and a t(14;18)(q32;q21) have been reported. All were associated with a poor prognosis. Our patients demonstrated low level C-MYC and/or IGH amplification. To our knowledge, IGH amplification has not been previously reported in patients of with Intermediate high grade lymphoma. Low-level IGH and C-MYC amplification maybe predictive of a poor outcome in these patients. FISH analysis of C-MYC, IGH and IGH/BCL2 appears to be significant and is advised when Intermediate tumors are detected. Clinical data on these aggressive Intermediate lymphomas has been sparse and therapeutic regimes are controversial. Our findings, as well as further studies may be useful in the assistance of determining the best clinical course for this patient population.

1037F

Two deletions in the APC gene (an inherited deletion in exon 1 and a de novo deletion in exon 12) and a de novo MUTYH deletion causing familial adenomatous polyposis. *G. A. Molfetta[1,2], O. C. Vincenzi[2], J. Huber[3], A. A. Marques[1], W. A. Silva, Jr.[1,2].* 1) Departament of Genetics, Medical School of Ribeirao Preto, University of Sao Paulo; 2) National Institute of Science and Technology in Stem Cell and Cell Therapy, Regional Blood Center of Ribeirao Preto; 3) Medical Genetics Unit, University Hospital of the Medical School of Ribeirao Preto, University of Sao Paulo.

Familial adenomatous polyposis (FAP) is a syndrome of colon cancer predisposition characterized by the development of hundreds to thousands adenomatous colonic polyps. It presents a wide spectrum of clinical manifestations, including the following phenotypes: classical FAP, attenuated FAP, Gardner's and Turcot's syndromes. The clinical diagnosis of classic FAP includes the presence of more than one hundred colorectal polyps or a reduced number of polyps together with affected close relatives. Molecular diagnosis relies on detecting mutations in the APC gene, a tumor suppressor gene with autosomal dominant inheritance, complete penetrance, and variable expression. Molecular genetic testing of APC detects disease-causing mutations in up to 90% of individuals with typical FAP and is most used in the early diagnosis of at-risk family members, as well as in confirming the diagnosis of classical or attenuated FAP in individuals with equivocal findings. We describe here a family presenting classical FAP and DNA sequencing of the entire coding region of the APC gene that showed no mutations. Multiplex Ligation-dependent Probe Amplification (MLPA®) was also employed to detect deletions in MUTYH and APC genes and revealed that the patient inherited from her mother a deletion of the entire exon 1 in the APC gene. In addition, the same patient also showed two *de novo* deletions: exon 1 in the MUTYH gene and another deletion encompassing exon 12 in the APC gene. RFLP analysis helped to characterize the deletion extension on approximately 150 bp. Microsatellite analysis confirmed paternity as patient's father has no clinical diagnosis of FAP as well as neither mutations nor deletions within the APC gene. Regarding clinical aspects of this family, the patient showed less polyps than her mother and other relatives affected by FAP but showed more compromised organs such as intestine and stomach as she reported gastric polyps. We raise the hypothesis that the inherited deletion was the first event and caused the appearance of the other deletion due to recombination and unequal crossing over. We also speculated whether the two deletions within the APC gene are cis deletions located at the same chromosome as this gene is an important tumor suppressor and the total absence of its protein in case of trans deletions would be incompatible to the embryo life. Financial Support: CTC, INCTC.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1038W

Mutation screening in adenomatous polyposis coli gene in patients clinically diagnosed as familial adenomatous polyposis. *O. C. Vincenzi[1], G. A. Molfetta[1,2], J. Huber[3], C. Ayres[1], W. A. Silva, Jr.[1,2]*. 1) National Institute of Science and Technology in Stem Cell and Cell Therapy, Regional Blood Center of Ribeirao Preto; 2) Department of Genetics, Medical School of Ribeirao Preto - University of Sao Paulo, Ribeirao Preto, Sao Paulo, Brazil; 3) Medical Genetics Unit, University Hospital of the Medical School of Ribeirao Preto, University of Sao Paulo.

Colorectal cancer is among the most important malignant neoplasms, being the third cause of death by cancer in Brazil. Its development involves storage of mutations in oncogenes and in tumoral suppressor genes, as well as epigenetic changes. One of these abnormal genes is the Adenomatous Polyposis Coli (APC), a tumor suppressor that presents mutations in its codifying sequence and these mutations may be associated to colorectal adenomas that develop into adenocarcinomas. Defects in this gene cause Familial Adenomatous Polyposis (FAP), an autosomal dominant pre-malignant disease that usually progress to malignancy. FAP usually begins by the age of 16 years old and its main characteristic is the large development of pre-cancerous colonic polyps in the intestines, which invariably evolve for the installment of cancer. Among the goals of this project are the screening for mutations in the APC gene (GenBank M74088), using High Resolution Melting (HRM)technique and the sequencing of DNA in patients who present variations in this gene together with the search for deletions in gene APC in patients who do not show mutations. In this research are included patients with clinical suspicion of FAP which were attended in the Medical Genetics Unit and have signed for the consented blood taken in the University Hospital. After the extraction of genomic DNA, we proceeded with HRM and DNA sequencing in patients who presented variations. We also performed the technique of Multiplex Ligation-dependent Probe Amplification (MLPA) to detect deletions in patients which do not presented mutations in the APC gene. Among the 17 patients in study, we described a new mutation c.del2157T, which creates a premature stop codon at position 726 of the protein. We also found out two pathogenic mutations: a *nonsense* one, creating a premature stop codon site in position 302 (Arg302Stop), and a *missense* one in position 1822 (Asp1822Val) together with a deletion, detected by MLPA, of the complementary allele characterizing this patient as a compound heterozygous. In addition, MLPA was also employed to detect deletions in MUTYH and APC genes and revealed a family sharing a deletion of the entire exon 1 in the APC gene. Furthermore, the proband also showed two *de novo* deletions, one in exon 1 in the MUTYH gene and another deletion in exon 12 in the APC gene. Genetic counseling was offered to other members of the affected families who agreed to be tested. Financial Support: FAPESP(2011/11456-0).

1039T

High-Precision CNV Analysis of FFPE Samples with the nCounter Analysis System. G. *Geiss.* Res, NanoString Technologies, Seattle, WA.

Background The nCounter Analysis System from NanoString Technologies is a digital technology capable of precisely quantifying up 800 nucleic acid targets in a single reaction. The nCounter System has gained a strong reputation for delivering high-quality data for gene expression and miRNA analysis of FFPE samples. Here, we describe the development and optimization of a protocol for analysis of Copy Number Variation in FFPE samples and present high quality data generated using the protocol. Method A variety of fresh frozen and FFPE samples were tested against an nCounter CodeSet designed to quantify 85 CNVs known to be associated with cancer. Multiple parameters were evaluated to optimize performance of the assay. Results Four experimental parameters were found to influence the accuracy of the assay: sample quality, number of probes per region, method of fragmentation, and reference sample selection. Greater than 90% accuracy can be expected when appropriate experimental design considerations are accounted for and in many cases 95% accuracy was observed in our experiments. As expected, the quality of DNA extracted from FFPE samples quality is critical for data quality, but even poor quality samples are expected to generate reliable data with the right experimental considerations. Conclusion The nCounter Analysis System is a reliable tool for analysis of CNVs in FFPE samples. It generates unbiased, highly reproducible results from FFPE samples and the workflow advantages enable studies of hundreds of specimens to be completed quickly and efficiently.

1040F

Germline copy number variation in high-risk African American families with prostate cancer. *E. Ledet[1], X. Hu[2], O. Sartor[3], W. Rayford[4], M. Li[5], D. Mandal[1]*. 1) Dept Genetics, Louisiana State Univ HSC, New Orleans, LA; 2) Cytogenetics, Sonora Quest Laboratories, Tempe, AZ; 3) Departments of Medicine and Urology, Tulane University School of Medicine, New Orleans, LA; 4) 4Department of Preventive Medicine, University of Tennessee Health Science Center, Memphis, TN; 5) Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, LA.

Prostate cancer is a complex multi-allelic disease and the most common malignancy in men. The incidence of prostate cancer in African American men is more than twice as high as that of any other race. Despite the high prevalence of prostate cancer amongst African American men, this population has been under represented in genetic studies of prostate cancer. Although genomic copy number variations (CNVs) have been detected in prostate tumors, this is the first study describing germline CNVs in African American hereditary prostate cancer families. Ten high-risk African American families with 3 or more affected individuals and with an early age of onset were recruited. From these families, 37 individuals, including 23 affected males and 14 unaffected males, were selected for CNV analysis. Array comparative genomic hybridization (aCGH) was used to characterize germline CNVs unique to African-American men with hereditary prostate cancer. Through common aberration analysis in affected family members, novel CNVs were identified at chromosomes 1 and 16. Differential analysis comparing affected and unaffected family members identified a 9.4 kb duplication on chromosome 14q which segregate with prostate cancer patients in these high risk families. Importantly, this region encompasses IGHG3 gene which has been shown to have both significant gains in copy number as well as overexpression in prostate tumors in African Americans. These CNVs may represent a component of genetic predisposition which contributes to the high prevalence and mortality of prostate cancer in African American men.

1041W

***hapLOH*: Powerful haplotype-based profiling of allelic imbalance in impure tumor samples.** *P. Scheet[1,2], R. Xia[3,1], S. Vattathil[2,1]*. 1) Dept. of Epidemiology, Univ. of Texas MD Anderson Cancer Center, Houston, TX; 2) Human & Molecular Genetics Program, Graduate School of Biomedical Sciences, Univ. of Texas at Houston, Houston, TX; 3) Div. of Biostatistics, Univ. of Texas School of Public Health, Houston, TX.

Tumor samples often contain both cancer cells and normal cells due to difficulties in tumor excision. In these cases, cancer-associated allelic imbalance events such as copy-neutral loss of heterozygosity (LOH), deletions and other aneuploidies will exist in a fraction of the sampled cells only. Multiple SNP-array based methods have been developed specifically for detecting allelic imbalance in impure tumor samples, but they lose sensitivity when the tumor proportion is below 10%, as the signal from allelic imbalance resembles noise when analyzed at the single-marker level. However, for many cancer types (e.g. breast, leukemias, myelomas) and due to incomplete dissection, analysis of cancer genomes at low levels of tumor purity is a reality.

To address these problems, we have created *hapLOH*, a hidden Markov model-based computational tool that applies to observed B allele frequencies and total intensity data from SNP microarrays. The key to our approach is that we address directly the dependence among frequencies within a region of allelic imbalance that results from loss or gain of a contiguous set of alleles on an individual's germline chromosome. To do so, we infer germline haplotypes to inform expected patterns in the B allele frequencies at heterozygous markers in regions of imbalance. We offer both parametric and nonparametric strategies to accommodate haplotype uncertainty and magnitude of signals.

Our methods are sufficiently powerful to allow inference of LOH as short as 10 Mb in tumor purities of 3–4% using a 317K SNP chip applied to serial dilutions of breast cancer cell line CRL-2324. With a hypothetical 5M SNP array, we estimate we can detect aberrant cells at purities of less than 1%, depending on the level of genomic aberration. Finally, our method substantially improves the estimation of the lost or gained haplotype within the tumor genome; at 5% purity, we observe an absolute 16% gain in accuracy over methods that fail to account for germline information. Our method can be used to profile allelic imbalance in samples for which efficient methods have not previously existed.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1042T

Protein Tyrosine Phosphatase Receptor Type J (PTPRJ) is implicated in lymphomagenesis of Follicular (FL) and Diffuse Large B-cell (DLBCL) lymphomas. *C.A. Aya-Bonilla[1], M.R. Green[1, 2], E. Camilleri[1], M. Benton[1], C. Keane[3, 4], P. Marlton[4], R. Lea[1], M.K Gandhi[3, 4], L.R. Griffiths[1].* 1) Genomics Research Centre, Griffith University, Gold Coast, QLD, Australia; 2) Department of Oncology, School of Medicine, Stanford University, Stanford, CA, USA; 3) Clinical Immunohaematology Laboratory, Queensland Institute of Medical Research, Brisbane, Australia; 4) Department of Haematology, Princess Alexandra Hospital, Ipswich Road, Woollongabba, Australia.

Diffuse Large B-Cell Lymphoma (DLBCL) and Follicular Lymphoma (FL) account for around 50% of all Non-Hodgkin lymphoma (NHL) cases. However, despite their high variability in morphology, tumour biology and clinical outcome, common genetic alterations across DLBCLs and FLs have suggested the existence of shared molecular mechanisms in the lymphomagenesis of these diseases, which could have biological/clinical relevance. The use of the Hidden Markov model (HMM) approach in the analysis of SNP array data from unpaired tumour DNA samples has allowed the detection of loss of heterozygosity (LOH) regions and possible tumour suppressor genes (TSGs) that might play a role in oncogenesis. A HMM algorithm was used for the analysis of loss of heterozygosity (LOH) on SNP array data from FLs (n=21) and DLBCLs (n=21). This revealed a high frequency of LOH over the chromosomal region 11p11.2, containing the gene encoding the protein tyrosine phosphatase receptor type J (PTPRJ), which was validated by significantly increased rates of homozygosity of PTPRJ-targeting microsatellites in NHL cases. Although, PTPRJ regulates components of key survival pathways in B-cells (BCR, MAPK and PI3K signaling), its role in B-cell development is poorly understood. LOH of PTPRJ has been described in several types of cancer but not in any hematological malignancy. Interestingly, FL cases with LOH exhibited down-regulation of PTPRJ, in contrast non significant variation of expression in DLBCL cases. Moreover, we undertook a sequence screening in exons 5 and 13 of PTPRJ to identify the G973A (rs2270993), T1054C (rs2270992), A1182C (rs1566734) and G2971C (rs4752904) coding SNPs (cSNPs) and showed that the A1182 allele was significantly more frequent in FLs and in NHL cases with LOH. Significant over-representation of the C1054 allele (rs2270992) and the C2971 allele (rs4752904) were also observed in LOH cases. A haplotype analysis, performed on the cSNPs, found a significant lower frequency of the GTCG haplotype in NHL cases, but it was only detected in cases without LOH. Conversely, the GCAC haplotype was over-represented in cases with LOH. Altogether, these results implicate inactivation of PTPRJ as a common lymphomagenic mechanism in DLBCL and FL. In addition, it also suggests that the detection of polymorphisms in the PTPRJ ectodomain, together with LOH events, could affect the ligand-mediated activation of this receptor-like protein tyrosine phosphatase in these B-cell lymphomas.

1043F

Computational Techniques to Detect BFB in Tumor Genomes. *S. Zakov, M. Kinsella, V. Bafna.* Computer Science and Engineering, University of California, San Diego.

Breakage-Fusion-Bridge (BFB) is a mechanism of genome instability characterized by increased copy numbers of chromosomal segments. BFB begins with the loss of a telomere followed by the fusion of telomere-lacking sister chromatids. During anaphase, as centromeres migrate to opposite ends of the cell, the fused chromatids are torn apart, resulting in one chromosome with a terminal inverted duplication and another with a terminal deletion. After several such BFB cycles, repeated inverted duplications can yield dramatic amplifications of chromosome segments.

Observing an ongoing BFB process directly is rare, so evidence for BFB is generally only available from the modified genome after BFB has ceased. But, the proper interpretation of such evidence is not obvious. Clearly, much information can be inferred from the modified genomic sequence itself. Unfortunately, the expected genomic amplifications create significant difficulties when applying current sequencing methods, due to high ambiguity during read reassembly. Additional BFB evidence can be inferred from inversions detected by paired-end sequencing, identification of chromosomal segment locations by FISH analysis, and examination of copy number variations of enhanced chromosomal segments.

In this work, we present computational techniques to aid the interpretation of the types of BFB evidence mentioned above. We first pose the *BFB Count Vector* problem: given a genome segmentation and corresponding segment copy numbers, decide whether BFB can yield a genome with the given segment counts. We present the first polynomial time algorithm for the problem, improving a previous exponential time algorithm. We also show, that for certain alternative rearrangement models, count vectors, inversion patterns, and FISH analyses are sufficient for detecting BFB, while for other models such analyses are not enough for distinguishing BFB from non-BFB modifications. We apply the presented techniques to SNP array data from cancer cell lines and osteosarcoma biopsies and paired-end sequencing data from pancreatic tumors, and demonstrate their practicality for BFB identification and rejection.

1044W

Evaluation of cytogenetic and gene polymorphic variants in workers occupationally exposed to Silica-stone quarry industry. *R. Chandirasekar[1], K. Sasikala[1], B. Lakshman kumar[2], R. Raichel Jacob[1], H. Kavitha[2].* 1) Unit of Human Genetics, Department of Zoology, Bharathiar University, Coimbatore, Tamil Nadu, India. 641 04; 2) Department of Biotechnology KonguNadu arts and science college, Coimbatore, India 641 046.

Abstract Stone workers are exposed primarily to silica, there is strong epidemiological evidence to support an association between occupational crystalline silica exposure and several diseases such as silicosis, lung cancer, pulmonary tuberculosis, and chronic obstructive pulmonary disease, autoimmune diseases (systemic sclerosis, systemic lupus erythematosus, and rheumatoid arthritis), and renal diseases. Aim of the present study was, to evaluate the chromosome aberrations of workers exposure to silica dust in stone quarry industry. The cases for this study include workers exposed to silica for a long term. They were divided into two groups in accordance with duration of exposure and tobacco smoking habits. We observed changes in the frequency of chromosomal aberrations and micronuclei among the exposed subjects when compare to age matched controls. In this context, the results were observed in significantly ($P < 0.05$) high. Occupational exposure to crystalline silica was associated with an increased risk of lung cancer. a higher number of micronuclei and chromosomal aberrations were recorded in the workers exposed to silica-stone quarry industry (group I non-smokers and group II smokers). When XRCC1 and Tp53 gene polymorphic variants were carried out, based on that homozygous polymorphic variants were found have increased level of chromosome and micronuclei frequency. From this study we conclude that homozygous polymorphic variants has an increased association with cytogenetic parameters among the long term exposed workers of stone quarry industry.

1045T

Differential expression of AURKA and AURKB genes in bone marrow stromal mesenchymal cells of myelodisplastic syndrome: Correlation with G-banding analysis and FISH. *F.M. Oliveira[1,2], A.R. Lucena-Araujo[1,2], M.C. Favarin[1], E.M. Rego[1,2], R.P. Falcão[1,2], B.P. Simões[1,2], D.T. Covas[1,2], A.M. Fontes[2].* 1) Internal Medicine - Hematology, Faculty of Medicine of Ribeirão Preto, Riberão Preto, São Paulo`,Brazil; 2) National Institute of Science and Technology in Stem Cell and Cell Therapy, Ribeirão Preto, Brazil.

It has been demonstrated that genomic alterations of cells in the hematopoietic microen-vironment could induce myelodysplastic syndromes (MDS) with ineffective hematopoi-esis and dysmorphic hematopoietic cells, and subsequent transformation to acute mye-loid leukemia. This investigation is the first attempt to correlate the gene expression profile of AURKA and AURKB in a cytogenetically stratified population of mesenchymal stem cells (MSCs) from MDS patients. We found that AURKA mRNA was expressed at significantly higher levels in MSCs even with normal/altered karyotype when compared with hematopoietic cells and healthy donors. Additionally, we found that the presence of chromosomal abnormalities (mainly aneuploidies) in hematopoietic cells/ MSCs was also associated with higher levels of AURKA. Different form previous investigations, our findings, regarding AURKA expression support the hypothesis that the presence of chromosomal abnormalities in MSCs from MDS is not a consequence of the method used for chromosome preparation. They may reflect the genomic instability present in the bone marrow microenvironment of MDS patients. This information is also supported by differences observed in the growth kinetics between MSCs from healthy donors (normal karyotype) and from MDS patients with abnormal karyotype. In summary, our results may not be considered evidence that MDS and MSCs are originated from a single neoplastic clone. In fact, both cells (hematopoietic and MSCs) may probably be altered in response to damage-inducing factors and the presence of genomic abnormalities in MSCs suggests that an unstable bone marrow microenvironment may facilitate the expansion of MDS/leukemic cells.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1046F

A new function for the VHL tumour suppressor protein: Prevention of genome instability through promotion of homologous recombinational DNA repair. *M.S. Meyn[1,2,3], J.L. Metcalf[4], P.S. Bradshaw[1], M. Komosa[1], S.N. Greer[4], W.Y. Kim[5], M. Ohh[4].* 1) Program in Genetics and Genome Biology, Hospital for Sick Children, Toronto, Ontario, Canada; 2) Department of Molecular Genetics, University of Toronto, Toronto, ON, Canada; 3) Department of Paediatrics, University of Toronto, Toronto, ON, Canada; 4) Department of Laboratory Medicine and Pathology, University of Toronto, Toronto, ON, Canada; 5) Department of Haematology Oncology, The Lineberger Comprehensive Cancer Centre, University of North Carolina, Chapel Hill, NC.

Germ line mutations in the *VHL* tumour suppressor gene cause Von Hippel Lindau syndrome, a condition associated with a high risk of renal carcinoma. Loss of VHL function also occurs in the vast majority of sporadic clear-cell renal carcinomas. One characteristic of these tumours is a high degree of genome instability. The molecular mechanism underlying the link between genome instability and loss of VHL is not easily explained by VHL's known function: targeting the transcription factor HIFα for ubiquitylation-mediated destruction. We now report that VHL is a member of the DNA Damage Response (DDR) network that coordinates DNA repair and other cellular functions following induction of DNA Double Strand Breaks (DSBs).

We find that SOCS1 (suppressor of cytokine signalling 1) promotes K63-ubiquitylation of VHL in response to DSBs. Loss of VHL results in impaired phosphorylation of H2AX, a key component of the platform of modified chromatin that develops around DSBs. VHL loss also causes attenuation of DSB-induced foci of the DDR sensor protein MRE11, decreased homologous recombination (HR) repair of DSBs and persistence of unrepaired DSBs. These effects on the DDR network are VHL K63 ubiquitylation-dependent but HIFα-independent, as a *VHL* mutant that cannot undergo K-63 ubiquitylation but retains its ability to target HIFα is deficient in these DDR responses.

Our results identify VHL as a novel component of the DDR network, a major defense against genome instability. Like the Fanconi anemia proteins FancD2 and FANCI, VHL must be ubiquitylated in order to facilitate HR repair of DSBs. Our observations support the idea that loss of VHL impairs both the cell's DDR network and HIFα-dependent apoptosis, defects that act synergistically to promote the survival of cells with increased DNA damage and genome instability, cells that likely contribute to the development of cancer.

1047W

A new role for the ATM protein: Homologous recombination repair of DNA damage. *P. Bradshaw[1, 2], M. Komosa[1], M.S. Meyn[1, 2].* 1) Program in Genetics & Genome Biology, Hospital for Sick Children, Toronto, ON, Canada; 2) Dept of Molecular Genetics, University of Toronto, Toronto, ON, Canada.

Ataxia-Telangiectasia (A-T), an autosomal recessive disorder caused by ATM mutations, is characterized by progressive cerebellar ataxia, immunological defects, radiation sensitivity, cancer predisposition and genomic instability. Following induction of DNA Double Strand Breaks (DSBs), the ATM protein kinase plays a critical early role in triggering the cellular DNA Damage Response (DDR), which leads to local establishment of a megabase region of modified chromatin surrounding the DSB, as well as activation of cell cycle checkpoints and stress responses, DDR functions that together determine cell fate (survival, senescence or death) following DNA damage. We now report a new role for the ATM protein: facilitating Homologous Recombination (HR) repair of DSBs. We find that ATM recruitment to DNA damage sites differs from those of other DDR proteins: ATM forms small discrete foci that do not diffuse with time but remain tightly confined to regions proximal to the DSB. Interestingly, ATM foci do not directly co-localize with foci of RPA, FANCD2 and Rad51, proteins implicated in HR repair. Rather, the kinetics and spatial dynamics of ATM and HR protein foci suggest that ATM may mark a subset of DSBs for subsequent HR-mediated repair. Consistent with these findings, ATM promotes HR repair by facilitating CtIP-mediated resection of DSBs, thereby allowing RPA access to DSBs, a key initial step in HR repair. Further, while MDC1 is required for the association of most DDR proteins with the megabase region of modified chromatin that surrounds a DSB, it is dispensable for ATM accumulation at DSBs and for ATM activation of the DDR, supporting the notion that ATM action occurs in small localized regions immediately proximal to DSBs. Our results demonstrate a direct role played by ATM in the repair of DSBs and suggest that ATM-mediated processing of DNA damage is restricted to those DSBs that are destined to undergo HR repair. Further, ATM-mediated signaling does not extend along megabase chromatic regions but is restricted to localized regions proximal to the DSB. The data represents a paradigm shift in our understanding of the role played by ATM in the repair of DSBs and activation of the DDR network.

1048T

A crosstalk between ATM and FA/BRCA proteins is necessary to preserve DNA integrity during G2/M phase. *L.C. Torres Maldonado[1], D. Sosa[1,2], M. Castañon[2], A. Rodriguez[1], B. Molina[1], C. Lopez Camarillo[2], S. Frias[1,3].* 1) Genetica Humana, Instituto Nacional de Pediatria, Mexico DF, Mexico DF, Mexico; 2) Posgrado en Ciencias Genomicas. Universidad Autonoma de la Ciudad de México. Mexico D.F. Mexico; 3) 3Departamento de Medicina Genomica y Toxicología Ambiental, Instituto de Investigaciones Biomedicas, Universidad Nacional Autonoma de México. Unidad INP. Mexico D.F. Mexico.

Fanconi Anemia (FA) is a recessive disease, characterized by chromosomal instability, pancytopenia, congenital abnormalities, predisposition to malignancy and hypersensitivity to DNA interstrand crosslinking agents. FA is caused by mutations in one of 15 genes involved in the FA/BRCA pathway of DNA repair. Recent evidence suggests that in order to preserve their genome integrity FA cells are dependent on alternative DNA monitoring and repair proteins, such as ATM. We have observed that hydroxyurea (HU), applied during G2, is able to enhance the DNA damage produced by MMC in FA-A cells. Therefore, we hypothesized that HU interferes with a DNA integrity checkpoint orchestrated by ATM and the FA/BRCA pathway that operates during G2. In order to corroborate this hypothesis we analyzed the chromosomal aberrations (CA), ATM mRNA levels by real time RT-PCR and ATM phosphorylation status by flow cytometry during G2 in FA-A and normal cell lines, cultured and treated with 1) MMC, 2) HU, and 3) MMC+HU, and 4) an untreated control. FA-A cells displayed a larger number of CA compared with normal cells in all treatments; but the highest frequency of CA was observed with the MMC+HU treatment. ATM mRNA levels were significantly elevated in FA-A cells treated with MMC or HU, being five times higher with the MMC+HU treatment in comparison to the FA-A control; in contrast normal cell lines showed no differences in any treatment. Finally, the phosphorylation of ATM was increased in FA-A cells even without treatment, predominantly during G2 regardless of the experimental condition, in comparison with normal cells, Our data suggest that normal cells with a proficient connection between ATM and the FA/BRCA proteins transit from G2 to M with low chromosome damage despite the presence of MMC or HU. Moreover the elevated levels of ATM mRNA in FA-A cells suggests that effectively these cells are over activating the ATM pathway to compensate their DNA repair deficiencies. However, the increased number of CA in FA-A cells induced by HU during G2, suggest that FANCA and ATM belong to a DNA damage checkpoint pathway that operates before the progression to M phase, and that the connection among these proteins is lost in the presence of FANCA mutations, rendering FA cells with an excess of chromosomal aberrations that otherwise should be eliminated. HU could interfere with this alternative pathway inhibiting downstream effectors of ATM. CONACYT 84259.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1049F

A dynamic model of the FA/BRCA DNA repair pathway. *A. RODRIGUEZ[1], L. TORRES[1], D. SOSA[1], E. CORTES[2], A.M. SALAZAR[3], R. ORTIZ[2], P. OSTROSKY[3], L. MENDOZA[3], S. FRIAS[1, 3].* 1) Laboratorio de Citogenética, Departamento de Investigación en Genética Humana, Instituto Nacional de Pediatría, Mexico; 2) Departamento de Ciencias de la Salud, Universidad Autónoma Metropolitana-Iztapalapa; 3) Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México.

The FA/BRCA pathway repairs DNA interstrand-crosslinks (ICLs). Mutations in the genes coding for this pathway cause Fanconi anemia (FA), a chromosome instability syndrome with increased spontaneous chromosomal aberrations and high sensitivity to ICL inducing agents. Despite the presence of DNA damage, FA cells are able to progress throughout the cell cycle, suggesting that the FA/BRCA pathway deficiency affects the function of the DNA integrity checkpoints and that FA cells could process the damage by using alternative error-prone DNA repair pathways. Given the complexity of the ICL repair process, in this work we have developed a Boolean network model of the FA/BRCA pathway to generate some predictions about the functioning of the pathway and its mutants. To obtain the model we have thoroughly analyzed the FA literature, reconstructed the simplest network that contains all the relevant information and generated a Boolean network model using the logic operators AND, OR and NOT. The program BoolNet has been used to perform computational simulations of the model. The in silico simulations suggest that FA cells: 1) Activate primarily the non-homologous end-joining pathway as an alternative DNA repair process, followed by the homologous recombination repair to deal with the double strand breaks generated during ICL processing. 2) Continue throughout the cell cycle, even with unrepaired DNA double strand breaks, activating the mechanism known as Checkpoint Recovery. These proposals were partially addressed in the lymphoblastoid cell lines FA- A VU817 and NL-49, exposed to mitomycin C. To validate proposal no. 1, the presence of the phosphorylated DNA-PKcs protein was evaluated by flow cytometry, a critical protein in the non-homologous end-joining pathway. Experimental results and simulations concur that the non-homologous end-joining pathway is the main alternative DNA repair pathway in FA-A cells. To validate proposal no. 2, the presence of the phosphorylated CHK1 protein was evaluated, this protein is involved in the cell cycle arrest when the DNA has been damaged. Moreover, the entrance and exit from mitosis has been evaluated by using the MPM2 antibody and a cytokinesis block essay. Both, the model and our experimental approach concur that despite FA-A cells efficiently activate the CHK1 protein and accumulate on the G2 phase of the cell cycle due to the large amount of DNA damage, they are still able to divide and proceed throughout the cell cycle.

1050W

Fanconi anemia pathway proteins act together and independently to regulate homologous recombination and synthesis of telomeric DNA in ALT-immortalized human cells. *H. Root[1,2], M. Komosa[1], A. Larsen[1,3], D. Bazett-Jones[1,3], M.S. Meyn[1,2,4].* 1) Program in Genetics & Genome Biology, Hospital for Sick Children, Toronto, ON, Canada; 2) Dept Molecular Genetics, University of Toronto, Toronto, ON, Canada; 3) Dept of Biochemistry, University of Toronto, Toronto, ON, Canada; 4) Dept of Paediatrics, University of Toronto, Toronto, ON, Canada.

Fanconi anemia (FA) is an inherited disorder characterized by bone marrow failure, congenital malformations and cancer. To understand how FA proteins function in replication and genetic recombination, we are studying their role in ALT, the recombination-dependent telomere maintenance pathway used by the 10–15% of human cancers which are telomerase negative.

We find that FANCD2, FANCJ, FAN1 and FANCP/SLX4 form nuclear foci that colocalize with telomeric foci in cells that utilize ALT telomere maintenance. This colocalization typically occurs during S/G2 within ALT associated PML bodies (APBs). Our electron spectroscopic and immuno-FISH imaging demonstrates that APBs primarily contain extrachromosomal telomeric repeat (ECTR) DNA that is partially single stranded in nature. SiRNA depletion of FANCD2, FAN1 or FANCP in ALT cells leads to a rapid and dramatic increase in ECTR DNA synthesis, while depletion of the DNA helicases FANCJ or BLM leads to a decrease in ECTR DNA synthesis. Using siRNA co-depletion we find that ECTR DNA synthesis in FANCD2-depleted cells requires BLM, FANCJ, and ATR expression, but is independent of RAD51 and POLH. To characterize the ECTR DNA within individual cells, we have developed a novel diffusion based FISH approach and find that the ABPs in FANCD2, FAN1 or FANCP depleted-ALT cells contain tangled complexes of hundreds to thousands of individual G and C strand ECTR DNA ranging from <1 to >50 kb in length. Production of excess ECTR DNA occurs within the first few doublings post protein depletion, consistent with a mechanism involving rolling circle replication. This ECTR DNA is not the result of rapid telomere deletion events, as there is no increase in telomere free ends in FANCD2-depleted ALT cells. Instead, interactions between telomeric DNAs increase upon FANCD2 depletion, with cells showing an increase in telomere sister chromatid exchanges and more frequent or stable interactions between telomeres and APBs.

Although siRNA depletion of FANCD2, FAN1 or FANCP all results in the common phenotype of ECTR DNA amplification, siRNA depletion of any one of these three proteins does not reduce the localization of the other proteins to APBs. This, taken together with our FANCJ results, suggests that FA proteins do not function in a single linear pathway, but rather, work both together and independently to regulate genetic recombination.

1051T

Transcriptome sequencing of tumor subpopulations reveals a spectrum of therapeutic options in squamous cell lung cancer. *C.L. Barrett[1,2], R.B. Schwab[1,4], H. Jung[1], B. Crain[1], D.J. Goff[5], C.H.M. Jamieson[5], P.A. Thistlethwaite[4], O. Harismendy[1,2,3], D.A. Carson[6], K.A. Frazer[1,2,3,7].* 1) Moores Cancer Center, UCSD, La Jolla, CA; 2) Department of Pediatrics and Rady Children's Hospital, UCSD, La Jolla, CA; 3) Clinical and Translational Research Institute, UCSD, La Jolla, CA; 4) Division of Cardiothoracic Surgery, UCSD, La Jolla, CA; 5) Department of Medicine, Stem Cell and Moores Cancer Center, UCSD, La Jolla, CA; 6) Sanford Consortium for Regenerative Medicine, La Jolla, CA; 7) Institute for Genomic Medicine, UCSD, La Jolla, CA.

The cancer stem cell (CSC) hypothesis predicts that the tumor subpopulation responsible for chemotherapy resistance is comprised of CSCs, making CSC-targeted therapies a priority. Here, we sorted a squamous cell non-small cell lung carcinoma into CSC and non-CSC subpopulations based on the CD133 marker and performed whole-genome analysis and transcriptome sequencing (RNA-seq) to gain insights into the molecular mechanisms underlying their phenotypic differences and to identify new potential therapeutic targets. We determined that approximately half of the genome in both the CD133+ and CD133- subpopulations was highly deranged by copy number alterations in a largely indistinguishable manner. Transcriptome analysis identified highly expressed cell surface (CD molecule, GPCR, and ion channel) mRNA isoforms in the tumor cells whose protein products are actively being pursued as drug targets for other cancers. Interestingly, many of these gene isoforms are differentially expressed between the CD133+ and CD133- subpopulations and potentially contribute to their phenotypic differences. Expression analysis suggests that both subpopulations proliferate due to very high levels of c-Myc and the TRAIL long isoform (TRAILL) and that the normal apoptotic response to this was likely blocked by very high levels of Mcl-1L, Bcl-xL and c-FlipL —isoforms for which targeted therapies are now in clinical development. We also detected expression of two oncofetal isoforms in the tumor subpopulations. Transcriptome sequencing data for 221 squamous cell lung cancers confirmed the generality of our findings for this disease. Our study demonstrates that transcriptome sequencing can provide deep insights into the phenotypic differences and similarities of CSCs and non-CSCs and, importantly, can be leveraged to identify new potential therapeutic options for cancers beyond what is possible with DNA sequencing.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1052F

A novel bi-allelic MSH2 mutation associated with constitutional mismatch repair deficiency syndrome, and review of the clinical phenotype. *P-Y.B. Au[1], R. Perrier[1], E.G. Puffenberger[3], S. Hume[4], R. Anderson[5], L. Lafay-Cousin[5], D. Strother[5], B. McInnes[1], J. Parboosingh[1,2], F. Bernier[1,2].* 1) Medical Genetics, Alberta Children's Hospital, Calgary, Alberta, Canada; 2) Alberta Children's Hospital Research Institute for Child and Maternal Health; 3) The Clinic for Special Children, Strasburg, Pennsylvania; 4) Department of Medical Genetics, University of Alberta, Edmonton, Alberta, Canada; 5) Department of Hematology and Oncology, Alberta Children's Hospital, Calgary, Alberta, Canada.

DNA mismatch repair (MMR) is important for genomic stability and DNA replication fidelity. Lynch syndrome is due to mono-allelic mutations of MMR genes. However, bi-allelic mutations that affect MMR genes result in constitutional mismatch repair deficiency syndrome (CMMR-D). Patients with CMMR-D frequently develop malignancy in childhood, often within the first decade. Malignancies include childhood leukemia, lymphoma, and brain tumours. Lynch related tumours such as colorectal cancer are also possible, and can present in adolescence. Interestingly, these patients have skin findings similar to neurofibromatosis type 1 (NF1). We report a novel mutation in MSH2 causing CMMR-D in several children in a large extended consanguineous Southeast Asian family. The index patient presented at 4 years with T cell lymphoblastic lymphoma, and then developed gliobastoma multiforme at age 10. His sister and his first cousin once removed also developed glioblastoma at ages 7 and 6 respectively. Another cousin developed a primitive neuroectodermal tumour at 7 years. All 4 children had multiple café au lait macules and axillary freckling. However, unlike NF-1, many of their café au lait macules were atypical. SNP array homozygosity mapping identified a region identical by descent on chromosome 2, corresponding to both MSH2 and MSH6 loci. MSH2 and MSH6 were absent on immunohistochemical staining of tumour tissue. Sequencing of MSH6 revealed no mutation. MSH2 was sequenced, and the mutation hMSH2 c.[1276+3A>C] was identified on both alleles in the index patient. The other affected children of this extended family were also homozygous for this mutation. This mutation affects the MSH2 intron donor sequence and is predicted to be pathogenic. Mutations affecting the +1 and +2 donor positions of this donor site have been described for Lynch syndrome and are known to affect splicing, but this mutation at +3 is novel, particularly in the context of bi-allelic mismatch repair deficiency. PMS2 is the most common MMR gene associated with CMMR-D. However, MLH1, MSH6 and MSH2 related CMMR-D have also been described, with MSH2 being more rare. There are now at least 10 cases of MSH2 related CMMR-D in the literature, involving 5 families. We will review the literature around CMMR-D and MSH2 mutations, and describe the skin pigmentary abnormalities seen in these children. The skin phenotype has not been well characterized for CMMR-D and exhibits some differences from NF-1.

1053W

Clinical and histopathologic characteristics between familial and sporadic melanoma in Barcelona, Spain. *C. Badenas[1], P. Aguilera[2], C. Carrera[2], J. Malvehy[2], JA. Puig-Butille[1], E. Martinez-Barrios[1], J. Palou[2], S. Puig[2].* 1) Biochemistry and Molecular Genetics Unit, Hospital Clinic-IDIBAPS. CIBERER, Barcelona, Spain; 2) Dermatology Department. Melanoma Unit. Hospital Clínic. IDIBAPS. CIBERER. Barcelona. Spain.

Background: About 6 to 14% of melanoma cases occur in a familial setting. In 20 to 40% of melanoma families germline mutations in *CDKN2A* are detected. **Objective**: to further characterize the clinical and histopathologic characteristics of familial melanoma to provide more information to clinicians and also contribute to understand the complex interplay of genetic and environmental factors in the pathogenesis of melanoma. **Methods**: Clinical, histopathologic and immunohistochemical characteristics of 62 familial melanomas were compared with 127 sporadic melanomas from our melanoma unit. **Results**: The most representative variables associated to familial melanoma were earlier age at diagnosis (OR 1.036; 95% CI 1.017–1.055), lower Breslow thickness (OR 1.288; 95% CI 1.013–1.683) and *in situ* melanomas (OR 2.645; 95% CI 1.211–5.778). And the most representative variables associated to *CDKN2A* mutation carriers were earlier age at diagnosis (OR 1.060; 95% CI 1.016–1.105), in situ melanomas (OR 6.961; 95% CI 1.895–25.567), the presence of multiple melanomas (OR 8.920; 95% CI 2.399–33.166) and the immunopositivity of the tumors for cytoplasmic survivin (OR 9.072; 95% CI 1.025–85.010). **Limitations**: the study is limited to patients from the Mediterranean area. **Conclusions**: familial occurrence of cutaneous melanoma was significantly associated with the earlier age of onset, lower Breslow thickness and a higher proportion of in situ melanomas; and also with the higher risk of multiple melanomas and the cytoplasmic survivin immunostainig in the case of carriers of *CDKN2A* mutations.

1054T

A survey of stromal responses and their clinical associations defined by the fibromatosis signature in different types of carcinomas. *JL. Chen[1], I. Espinosa[1,2], AY. Lin[3,4], OY. Liao[5], M. van de Rijn[1], RB. West[1].* 1) Department of Pathology, Stanford University Medical Center, Stanford, California, United States of America; 2) Department of Pathology, Hospital de la Santa Creu i Sant Pau, Institute of Biomedical Research (IIB Sant Pau), Autonomous University of Barcelona, Barcelona, Spain; 3) Department of Medicine, Santa Clara Valley Medical Center, San Jose, California, United States of America; 4) Department of Medicine, Stanford University Medical Center, Stanford, California, United States of America; 5) Department of Statistics, Stanford University, Stanford, California, United States of America.

Purpose: The fibromatosis signature developed from desmoids-type fibromatosis (DTF) was able to define a robust stromal response in breast carcinoma and identify a subset of breast tumors with better clinical outcome. The purpose of this study is to examine (1) this DTF fibromatosis signature in determining stromal responses and (2) its prognostic value in different types of malignant carcinomas. Experimental Designs: Gene expression profiles were collected from five ovarian, five lung, two colon and three prostate cancer datasets constituted by the total of 1127 ovarian, 519 lung, 493 colon and 279 prostate tumors. DTF fibromatosis signature was applied to each dataset to identify the subset of tumors with strong DTF stromal responses. Additionally, two different tissue microarrays of 204 ovarian tumors and 140 colon tumors were constructed for immunohistochemistry studies to examine the expression of protein markers of DTF fibromatosis signature, SPARC and CSPG2, in these tumors. Clinical data from each tumor were also collected to determine the correlation of high DTF stromal response with the survival of patients. Results: DTF fibromatosis signature defined robust stromal responses in ovarian, lung, and colon cancer. Transcriptome analysis and immunohistochemistry both showed that there were statistically significant worse survival associations of the subset of ovarian tumors with strong DTF stromal response. Conclusion: The results suggest that DTF fibromatosis signature can successfully identify the subset of ovarian, lung and colon carcinomas with strong stromal responses. The subset of tumors identified by this signature in breast and ovarian cancer, however, had opposite survival associations.

1055F

Anticipation in Hereditary Paraganglioma and Pheochromocytoma Syndromes. *H. Druker[1,2], D. Malkin[1,2], S. Meyn[1,2].* 1) Cancer Genetics Program, Hospital Sick Children, Toronto, ON, Canada; 2) Univeristy of Toronto, Toronto, ON, Canada.

Introduction: Paragangliomas and pheochromocytomas are rare in children. Germline mutations of *VHL, RET, SDHx, NF1, TMEM127*, or *MAX* has been identified in 59% of children. Genetic anticipation in these conditions has yet to be described. Four kindreds are presented with two conditions, VHL and PGL4, where anticipation is observed. This observation suggests the need for earlier initiation of surveillance. VHL Case 1: A three year old presented with bilateral pheochromocytomas. The average age at diagnosis in his parents' generation was 17 and in his grandfather's generation was 45. VHL Case 2: A nine year old presented with bilateral pheochromocytomas. Of the six VHL carriers in his generation, the average age at diagnosis was 11. The average age at diagnosis in their parents' generation was 45, and their grandmother remains unaffected at 78 years of age. PGL4 Case 1: A 15 year old presented with a paraganglioma of the bladder. His father was found to carry the same mutation at 47 years of age and is tumour-free. PGL4 Case 2: A 13 year old initially presented with bilateral pheochromocytomas. His mother was subsequently found to carry the same mutation at 67 years of age and is tumour-free. Discussion: Genetic anticipation describes the phenomenon in a genetic disorder of earlier age of onset with increasing symptom severity in subsequent generations. Molecular mechanisms to support clinical observations have been described in trinucleotide repeat disorders such as Huntington's disease and in telomere attrition disorders such as hereditary cancer predisposition syndrome, Li-Fraumeni Syndrome. The molecular mechanisms of both VHL and PGL4 are to be elucidated although preliminary evidence suggests that increasing genomic variation in subsequent generations contributes overall to their genetic instability. The possibility of earlier presentation of tumours in these conditions points to the need for earlier initiation of surveillance in subsequent generations.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1056W

Evaluation of MtDNA Dloop polymorphisms in Iranian Breast Cancer patients. *M. Ghaffarpour[1,2,5], S. Eslamizadeh[3], M. Arabzadeh[3], B. Motahari[4], N. Moazami[2], M. Houshmand[1,5].* 1) Medical Genetics Dep, National Institute of Genetic Engineering and Biotechnology, Tehran, Tehran, Iran; 2) Iranian Research Organization for Science and Technology; 3) Biology Dep, Faculty of Science, Ahar Azad University, Tabriz,Iran; 4) Biology Dep, Faculty of Science, Complex Azad University, Tehran,Iran; 5) Medical Genetics Dep,Special Medical Center, Tehran, Iran.

Introduction: Mitochondrial genome alterations have been suggested to play an important role in carcinogenesis. Its genome is also more susceptible to mutations because of high levels of reactive oxygen species and limited repair mechanisms. The D-loop of mitochondrial DNA (mtDNA) contains essential transcription and replication elements, and mutations in this region might alter the rate of DNA replication. the mtDNA D-loop region which is highly polymorphic, containing two hyper variable regions: HV1 (16024–16383) and HV2 (57–333) has been reported as somatic mutation “hot spot” in many types of cancer, including breast cancer. The aim of this study was evaluation of frequency of mononucleotide repeat (poly-C) between 303 and 310 nucleotides(HV1) and C16069T, T16189C, C16261T and T16311C polymorphisms(HV2) as somatic mutations “hot spots in Iranian patients with breast cancer. Materials and Methods: patients; 50 paired Fresh tumor and adjacent normal samples were obtained from consecutive patients with BC from IRANIAN National Tumor Bank, National Cancer Institute, Imam Khomeini Hospital Complexes, Medical Tehran University, and Tehran, Iran. Blood samples from 404 individuals of 17 ethnicities and 100 random in divide from Tehran special Medical Center (mixed ethnicities) were used as healthy controls. Mutational analysis: fresh tumors and their adjacent and also bloods samples as a control were extracted for genomic DNA using the QIAamp Mini Kit .We searched for mtDNA alterations of mononucleotide repeat (poly-C) between 303 and 310 nucleotides and C16069T, T16189C, C16261T and T16311C polymorphisms were analyzed by means of PCR sequencing. Result and conclusion: our result indicated that the frequency of mononucleotide repeat (poly-C) between 303 and 310 nucleotides were 11 of 27;40.72%, 9 of 27;33.33%,7 of 27;25.95% and 0 of 27;0% for C7,C8,C9 and C10 respectively for patients cases and124 of 504 ;38.3%,156 of 504; 48.1%,42 of 504 ;13% and 2 of 504;0.6% respectively for control group.The frequency of C16069T, T16189C,A16261G and A16311G polymorphisms were 11 of 50;22%, 6 of 50;12%,7 of 50;14% and 12 of 50;24%respectively for patients group and were 13 of 504;2.6%, 95 of 504;18.8%,53 0f 504;10.5%and 78 of 504;15.5% respectively for normal group. In conclusion, our study indicated that C16069T mtDNA D-Loop mutation may significantly play a role in the etiology of breast cancer. Further studies are warranted to confirm these findings.

1057T

Does high HERV-K expression participates in acute leukemia pathogenesis in children? *D. Januszkiewicz-Lewandowska[1,2], K. Nowicka[1], J. Rembowska[1], J. Nowak[1].* 1) Institute of Human Genetics, Polish Academy of Sciences, Poznan, Poznan, Poland; 2) Department of Pediatric Oncology, Hematology and Bone Marrow Transplantation of Medical University.

Human endogenous retroviruses - HERV are dispersed in the genome and constitute about 8% of the human genome. It is believed that HERV sequences derived from exogenous retroviruses, which in the course of evolution were incorporated into human genome thousands years ago. Most HERVs are non-functional due to numerous mutations and deletions. Some HERVs contain sequences capable of transcription. Patients with leukaemia show antibodies against HERV-K, which suggests increased expression of HERV-K sequences in leukemic cells. Due to small number of published research devoted to the expression of endogenous retroviruses in leukemias in children, studies were undertaken which purpose was to assess the expression of the env gene of HERV-K and HERV-W in the cells of acute lymphoblastic leukaemia and acute myeloid leukaemia in children. Material for the study was bone marrow blasts cells or peripheral blood of patients diagnosed with acute leukemia and normal peripheral blood lymphocytes. Study group consisted of 170 children with acute lymphoblastic leukemia (ALL) and 38 with acute myeloid leukemia (ANLL) hospitalized in the Clinic of Pediatric Oncology, Hematology and Transplantation Medical University in Poznan. The control group consisted of 30 healthy volunteers. Expression of the env gene of retroviruses HERV-K and HERV-W and the control gene beta-actin was studied by real-time PCR using specific oligonucleotide primers and SYBR Green marker. Expression was evaluated on the basis of absolute thresholds cT, normalized expression relative to beta-actin and double normalized expression to beta-actin reference gene and reference cells - normal peripheral blood lymphocytes. Absolute env gene expression of HERV-K and HERV-W (measured by cT - threshold values) was significantly higher in peripheral blood lymphocytes in comparison to ALL and ANLL. Only in ANLL a significantly higher, normalized to beta actin gene, relative expression of env gene of HERV-K has been found. Double normalized to beta actin gene, and normal peripheral blood lymphocytes relative expression of the env gene of HERV-K was significantly higher in ANLL compared to ALL. No statistically significant difference between double-normalized expression of the env gene of HERV-W between ALL and ANLL has been demonstrated. It can be concluded, that high normalized expression of HERV-K env gene in ANLL suggests that the gene may participate in the pathogenesis of ANLL.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1058F

SMAD4 loss-of-function mutations cause wide range of clinical disorders with unpredictable age-related penetrance. *N.M. Lindor[1], K.E. Wain[2], M.S. Ellingson[3], J. McDonald[4], A.D. Gammon[5], M.E. Roberts[6], D.L. Riegert-Johnson[7].* 1) Health Science Research, Mayo Clinic AZ, Scottsdale, AZ; 2) Laboratory Genetics, Mayo Clinic, Rochester, MN; 3) Medical Genetics, Mayo Clinic, Rochester, MN; 4) Department of Pathology, University of Utah, Salt Lake City, Utah; 5) Huntsman Cancer Institute, Salt Lake City, Utah; 6) Hematology/Oncology, Mayo Clinic, Jacksonville, FLA; 7) Gastroenterology, Mayo Clinic, Jacksonville, FLA.

Background: Haploinsufficiency of *SMAD4*, a key gene in the BMP/TGF-beta signaling pathway, leads to the syndrome of juvenile polyposis (JP) and hereditary hemorrhagic telangiectasia (HHT), now referred to as the JP-HHT syndrome. Clinical information on SMAD4 mutation carriers is extremely limited with the largest clinically detailed series describing only 15 individuals. Thus, the full extent of clinical variability and frequency of features is not well-understood. In this study, we sought to better define the penetrance for JP and HHT manifestations. **Methods:** Medical records of 24 mutation carriers from 14 families were abstracted from three clinical practices. **Results:** Ten different *SMAD4* mutations were found. The mean age was 33.7 years (range .1–65) at last follow up. First colon exam was conducted at a mean age of 23 (range 4–51) for indications including anemia (n=7), GI bleeding+/− anemia (n=8), JP suspected (n=7), obstructive symptoms (n=1), and unknown (n=5). The most recent colon exams were conducted at a mean age of 35 years. The number of colonic polyps ranged from 0 to over 1000, were generally pan-colonic, and was not associated with age. Gastric and small bowel imaging, performed at a mean age of 38 years (range 4–68), showed extensive gastric polyposis in seven individuals and low numbers of small bowel polyps, mostly in the duodenum. HHT findings included epistaxis (n=15, 63%), mucocutaneous AVMs (n=12, 50%), iron deficiency anemia at some point (n=17,71%), liver AVMs (n=5, 38%), brain AVM (n=1, 6%), lung AVM (n=2, 17%), and an intrapulmonary shunt detected by cardiac bubble study (n=12, 71%). Additionally, three individuals had mildly enlarged aortic roots and one died following repair of ascending aortic dissection, one had three incidental brain aneurysms, one had mild aortic and mitral insufficiency, and one had bilateral spontaneous retinal detachments. Neoplasms included two colorectal cancers, two Hurthle cell thyroid tumors, and bilateral malignant ovarian teratomas, jejunal cancer, adenocarcinoma of the esophagus, lymphoma found in one individual each. **Conclusion:** Carriers of *SMAD4* mutations require management for both juvenile polyposis and hereditary hemorrhagic telangiectasia as risk for findings of both are high and quite unpredictable. A larger series of patients will be necessary to further explore the significance of some of the occasional findings in this series.

1059W

Automatic Endometrial Tumor Screening for Lynch Syndrome. *M. Myers[1], P. Conrad[1], J. Rabban[2], J. Terdiman[1], K. Loranger[1], A. Blanco[1], L. Chen[2].* 1) UCSF Gastrointestinal Cancer Prevention Program, San Francisco, CA; 2) UCSF Helen Diller Comprehensive Cancer Center, San Francisco, CA.

Lynch syndrome (LS), also known as Hereditary Non-Polyposis Colorectal Cancer (HNPCC), is a hereditary cancer syndrome associated with up to a 75% lifetime risk of developing colon cancer and up to a 70% lifetime risk of endometrial cancer. Historically, Lynch syndrome has been diagnosed using clinical and family history criteria (Amsterdam and Bethesda criterion); however, it remains under diagnosed due to the lack of sensitivity of this approach. Tumors caused by LS gene mutations are known to display molecular features (microsatellite instability (MSI) and loss of mismatch repair protein expression detected by immunohistochemistry (IHC), but these molecular tests have only recently begun to be utilized in a systematic way to screen patients for LS. In July of 2010, the pathology department at UCSF began automatic screening for LS on endometrial tumors that met specific criteria and were surgically excised at UCSF. This automatic screening included MSI and IHC but did not include MLH1 promoter methylation. Between June 2010 and May 2012, 130 endometrial tumors met inclusion criteria for automatic tumor screening for LS of which 30 tumors had abnormal IHC results. Of the 30 endometrial tumors that had abnormal IHC results, 5 had a loss of MSH6 protein expression, 2 had a loss of MSH2/MSH6 protein expression, 4 had a loss of PMS2 protein expression and 19 had a loss of MLH1/PMS2 protein expression. These abnormal IHC results were followed up with by a genetic counselor who contacted the patient to explain their results, obtain a family history and discuss the role of further genetic testing to rule out LS with germline testing and hypermethylation analysis of the MLH1 promoter. Of the 19 patients who had a loss of MLH1/PMS2, 10 underwent further genetic testing and 9 refused any further evaluation due to lack of interest or insurance restrictions. Five patients underwent hypermethylation analysis of the MLH1 promoter All MLH1 hypermethylation testing was sent to an outside institution and all five were identified to have hypermethylation of the MHL1 promoter, thus ruling out LS. Based on this limited data, we feel MLH1 promoter methylation testing should be included as part of the automatic endometrial tumor testing protocol. This would streamline the process in properly identifying what patients are in need of further follow up by the genetic counselor.

1060T

Detection of EGFR mutations by TaqMan® Mutation Detection Assays powered by Competitive Allele-Specific TaqMan® PCR Technology (castPCR™). *N. Normanno[1,2], C. Roma[1], R. Pasquale[1], A.M. Rachiglio[1], C. Esposito[1], F. Fenizia[1], M.L La Porta[1], A. Iannaccone[1], F. Bergantino[1], S. Costantini[1], A. Rico[3], R. Petraroli[4].* 1) Centro di Ricerche Oncologiche di Mercogliano, 83013 Mercogliano, Avellino, Italy; 2) Cell Biology and Biotherapy Unit INT-Fondazione Pascale Via M. Semmola 80131 Naples Italy; 3) FALCON Team, Life Technologies, Parc Technologique, 91190 Saint Aubin, France; 4) FALCON Team, Life Technologies, via Tiepolo 18, 20900, Monza, Italy.

The discovery of driver mutations in key genes involved in regulating proliferation and survival of cancer cells is leading to remarkable successes in translational medicine. Different molecular techniques have been developed in order to deal with some limitations of tumor samples: the poor quality of the DNA extracted from formalin fixed paraffin embedded (FFPE) tissues, the low quantity of DNA available and the contamination of tumor sample by non-transformed cells. castPCR™ technology has a higher sensitivity as compared with other molecular techniques. EGFR mutations were assessed by castPCR™ in 40 FFPE tumour samples including 34 NSCLC and 6 colorectal carcinoma (CRC). The sensitivity and specificity of this method was compared with routine diagnostic techniques including fragment analysis, Real-Time PCR (allelic discrimination) and the EGFR Therascreen RGQ kit. By using routine diagnostic techniques, EGFR mutations were detected in 12/34 NSCLC patients, including 6 samples with deletions in exon 19; 4 samples with a L858R point mutations in exon 21; 2 samples with respectively a G719S and a G719A point mutation in exon 18. No EGFR mutations were detected in 6 CRC specimens. We next analyzed the 28 EGFR mutation negative samples with castPCR™ in order to identify cutoffs that ensure high specificity of the assay in FFPE samples. We found that 27/28 samples were also negative by cast PCR™ when excluding samples with a Ct>37 or with a deltaCt>7. We found in one of the negative samples a L858R point mutation that was not detected by using other methods. This mutation was confirmed by using a Peptide Nucleic Acid clamping-based Real-Time PCR technique. castPCR™ was also able to identify the same 12 mutations detected by the other methods. Our preliminary data suggest that there is an excellent correlation between castPCR™ and other molecular methods and that castPCR™ is highly specific to detect EGFR mutations in FFPE samples. In addition, castPCR™ might reveal EGFR mutations that are not detectable by other methods and this might have important implications for clinical decisions. These findings are being confirmed in a larger cohort of specimens in which TMDA will be also compared with Next Generation Based-methods using an AmpliSeq™ Custom Panel designed on predictive mutations for colon and lung cancer on the Ion Torrent™ PGM™.“For Research Use Only. Not for human or animal therapeutic or diagnostic use.”.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1061F

Germline Mutations and Microsatellite Instability Status in a Population-based Study of Women with Ovarian Cancer. *T. Pal[1], A. Mohammad[2], P. Sun[2], J. Lee[1], J. Fulp[1], Z. Thompson[1], A. Doty[1], D. Coppola[1], S. Nicosia[3], T. Sellers[1], J. McLaughlin[4], H. Risch[5], B. Rosen[6], J. Schildkraut[7], S. Narod[2].* 1) Interdisciplinary Oncology, Moffitt Cancer Center, Tampa, FL; 2) Womens College Hospital, Womens College Research Institute, Toronto, Canada; 3) Pathology and Cell Biology, University of South Florida, Tampa, FL; 4) Samuel Lunenfeld Research Inst., Dalla Lana School of Public Health, Toronto, Canada; 5) Epidemiology and Public Health, School of Public Health, Yale University, New Haven, CT; 6) Gynecology-Oncology, Princess Margaret Hospital, Toronto, Canada; 7) Community and Family Medicine, Duke University, Durham, NC.

Introduction: The frequency of germline mutations in the mismatch repair (MMR) genes, which results in Lynch Syndrome, remains incompletely defined in epithelial ovarian cancers. The purpose of this study was to estimate the frequency of germline MMR mutations in a multicenter population-based study of epithelial ovarian cancers. **Methods**: The analysis included 1893 women with epithelial ovarian cancer ascertained from three population-based studies. DNA was obtained from blood samples and full sequencing of the coding regions was performed on three HNPCC genes, MLH1, MSH2 and MSH6. In addition, there were a total of 967 women in whom microsatellite instability (MSI) testing results were available on tumor samples and normal DNA, and in a subset of those with a MSI-H phenotype, comprehensive rearrangement testing was performed through Multiplex Ligation-dependent Probe Amplification (MLPA) in the three genes. **Results**: In the 1893 women, 9 clearly pathogenic mutations were identified through sequencing (0.5%): 5 in MSH6, 2 in MLH1 and 2 in MSH2. An additional 60 patients had predicted pathogenic missense variants identified (3.1%). Thus, the combined frequency of pathogenic and predicted pathogenic mutations in the entire sample was 3.6% (69/1893) and 6.3% (27/431) among women with endometrioid or clear cell tumors. Of the 967 women who had MSI testing results, 133 had an MSI-H phenotype, of which additional germline DNA was available in 88 cases. In these samples, MLPA for the MMR genes was performed, and 3 additional mutations (large deletions) were identified in MSH2. Of note, 4 of the 9 mutation carriers identified based on sequencing results had MSI testing, all of whom were MSI-H, thus MLPA was not performed on these samples. Therefore, the combined frequency of pathogenic mutations detected either by sequencing or MLPA in those with an MSI-H phenotype was 7.6% (7/92). **Conclusion**: A small but potentially important proportion of women with epithelial ovarian cancer harbor mutations in genes predisposing to Lynch Syndrome. Furthermore, MSI-H tumor phenotype and presence of non-serous histology appear to be of utility in identifying those with germline MMR mutations. Ultimately, identification of Lynch Syndrome has the potential to impact screening and treatment decision in these women and may facilitate testing for their at risk family members.

1062W

BRCA2 mutations might be detected at a high rate in a subset of cutaneous melanoma patients with pancreatic cancer in the familial history. *N. Soufir[1, 2], E. Rouleau[3], C. Derouet[3], A. Riffault[1], HH. Hu[1,2], S. Caputo[3], P. Hammel[4], V. Descamps[5], N. Basset-Seguin[2], P. Saiag[6], A. Bensussan[2], M. Bagot[2], R. Lidereau[3], B. Grandchamp[1].* 1) Department of Genetic, Hopital Bichat, Paris, France; 2) Inserm U976, Centre de Recherche sur la Peau, Hôpital Saint Louis, AP-HP, Paris, France; 3) Oncogenetic laboratory, Institut Curie, 35 rue Dailly 92210 Saint Cloud, France; 4) Gastro-Enterology department, Hôpital Beaujon, APHP, Clichy, France; 5) Dermatology department, Hôpital Bichat Claude Bernard, APHP, Paris, France; 6) Dermatology department, Hôpital Ambroise Paré Bernard, APHP, Boulogne Billancourt, France.

Three major melanoma predisposing genes have been characterized. The most frequently involved is *CDKN2A*, a cell cycle regulator, that is mutated in up to 20–30 percent; of melanoma prone families, and 5–10 percent; multiple primaries melanoma. The second is *CDK4*, an oncogene inhibited by *CDKN2A*. More recently, a third melanoma predisposing gene, *BAP1*, are described to predispose to ocular and cutaneous melanoma and also mesothelioma. In addition, it has been reported an excess of melanoma in members of *BRCA2* mutated families. Interestingly, *BRCA2* also predisposes to pancreatic cancer, a cancer which also shares common gene predisposition with melanoma. Few reports estimate the prevalence of BRCA2 variants in melanoma patients. Patients : genetic alterations in the BRCA2 gene were studied in 48 melanoma patients having no mutation in *CDKN2A* or *CDK4* and which had at least to one of the following criteria: breast cancer in the proband (15 cases) or at least 2 breast cancers (BC) in related (5 cases), pancreatic cancer (PC) in first or second related (20 cases), ovary cancer in first related (4 cases), prostate cancer in the proband (7 cases). Method: the entire coding sequence of the BRCA2 gene were screened with EMMA prescreening method and all the abnormal profiles were sequenced with Sanger sequencing. Results: One causal *BRCA2* variant (c.6037 A>T, p.K2013X) was identified in a multiple primary melanoma proband with a PC in 2nd degree relative. Two unclassified and rare *BRCA2* variants were identified in 2 patients: c.7505G>A, p.Arg2502His (primary melanoma and PC in first degree relative), c.8545A>G, p.Lys2849-Glu (primary melanoma and PC in first degree relative) and c.9256+28T>A (primary melanoma and 3PC in first and second degree relatives). A partial segregation in three patients was found with the variant c.8545A>G in two patients with melanoma. Discussion and conclusion : *BRCA2* causal or rare variants can be detected in melanoma patients with pancreatic cancer in the kindred. Melanoma patients with a familial history of pancreatic cancer should be tested for *BRCA2* mutation.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1063T

SUFU germline mutation in a patient with Gorlin syndrome identified by next generation sequencing. *J.B. Geigl[1], E. Heitzer[1], I. Wolf[2], L. Cerroni[2], M.R. Speicher[1].* 1) Institute of Human Genetics, Medical University of Graz, Graz, Austria; 2) Department of Dermatology, Medical University of Graz, Graz, Austria.

Gorlin Syndrome (GS; OMIM #109400), also termed Basal Cell Nevus Syndrome, is characterized by the presence of a variety of developmental anomalies and predisposition to a range of cancers. A major and problematic feature of this disorder is the development of multiple basal cell carcinomas (BCCs) beginning to appear from the early teens, varying in number from only some to thousands. About 5% of the patients develop medulloblastoma mostly presenting as desmoplastic variant. Gorlin syndrome is inherited in an autosomal-dominant manner with near complete penetrance and extremely variable expressivity. Birth incidence is estimated to be 1 in 19.000 individuals. Between 50% and 85% of patients harbor germ line mutations in the PTCH1 gene, a key regulatory component in the Sonic Hedgehog (SHH) signaling pathway. However, rare causative germline mutations, such as PTCH2 and SUFU, were also identified in other genes in the SHH signaling pathway. Here we present a 55-year old patient with a 20-year history of multiple BCCs involving face, back, and chest. Examination of the dental status revealed missing teeth in the lower jaw, paradontopathia, and atrophia of the alveolar process. Ocular anomalies were described as pigmentary changes of uvea and excavation of the optic papilla. Ectopic calcification of the falx, which is present in more than 90% of affected individuals, was detectable. Family history revealed further tumor affected relatives. We started our diagnostic work-up with Sanger sequencing and MLPA analysis of the PTCH1 gene, which did not reveal a disease associated mutation. Because of the growing evidence that mutations in other genes of the SHH signaling pathway may rarely cause GS, we enriched 372 tumor associated genes (TruSeq Custom Enrichment), including SMO, HIP, and SUFU, and sequenced them on the Illumina MiSeq® platform. The analysis identified a novel splice mutation c.454+1 G>A in SUFU. The splice site mutation was confirmed by mRNA analysis. SUFU is a tumor suppressor gene, which had previously mainly been associated with medulloblastoma. To the best of our knowledge, our report represents just the second germline SUFU mutation causative for GS. New sequencing strategies allow now novel strategies in the work-up of hereditary cancer, especially the identification of rare, causative germline variants. This significantly extends our genetic counseling options and may improve genotype-phenotype correlations in familial cancer.

1064F

Li-Fraumeni Syndrome: A Clinical, Genetic, and Epidemiologic Cohort Study. *P.L. Mai, J.A. Peters, L. Hoskins, F. Walcott, R. Bremer, S.A. Savage.* Clinical Genetics Branch, National Cancer Institute, Rockville, MD.

Li-Fraumeni syndrome (LFS) is a highly-penetrant autosomal dominant cancer predisposition syndrome associated with a wide range of cancers occurring at younger-than-expected ages. Specific diagnostic criteria for classic LFS and Li-Fraumeni-like syndrome (LFL) are based on the pattern of cancers within families. Germline *TP53* mutations are found in ~70% of LFS and ~40% of LFL families. Limited data exist on genotype-phenotype correlations and on the effect of environmental exposures on cancer risk in LFS. Also, very few studies of early cancer detection have been conducted. A recent study employing whole body MRI and other organ-specific screening showed promise, but the results need confirmation in larger studies. We are conducting a detailed, IRB-approved natural history study of LFS (NCT01443468), which opened to accrual in August 2011, after a successful recruitment pilot study. Eligibility criteria include individuals and family members with 1) classic LFS; 2) LFL; 3) a known germline *TP53* mutation; 4) individuals with choroid plexus (CPC) or adrenocortical carcinoma (ACC); or 5) individuals with 3 or more primary cancers. Participants complete family history and individual information questionnaires. Medical records are reviewed, and biospecimens are collected. Eligible *TP53*-mutation positive individuals may elect to undergo comprehensive, protocol-defined cancer screening at the NIH Clinical Center. Energy balance and physical activity are also being studied. In order to understand the human factors influencing uptake of genetic tests, adherence to screening/surveillance, and issues of individual, family, and community impact of living with LFS, we are undertaking a full battery of quantitative and qualitative individual and family evaluations. To date, 109 eligible unrelated families have been identified, including 76 with classic LFS or a germline *TP53* mutation, 22 LFL, and 11 individuals with multiple primaries, ACC, or CPC. After obtaining consent, additional selected family members are invited to participate. Genetic counseling and testing is underway for families and individuals with unknown *TP53* mutation status. Through this study and in collaboration with the newly organized international LFS Exploration (LiFE) Consortium, important questions on clinical management, genotype-phenotype correlations, and gene discovery in mutation-negative families will be conducted in order to improve LFS-related quality of life and survival.

1065W

Early detection of cystic pleuropulmonary blastoma and cystic nephroma in an asymptomatic child with *DICER1* mutation — Screening considerations. *F. Plourde[1], V. Larouche[2], N. Ferguson[2], M. St-Amant[2], I. Harvey[3], N. Taherian[4], N. Sabbaghian[1], J.R. Priest[5], W.D. Foulkes[1, 4, 6].* 1) Department of Medical Genetics, Segal Cancer Centre, Lady Davis Institute, Jewish General Hospital, Montreal, QC, Canada; 2) Centre Hospitalier de l'Université Laval, 2705 Laurier blvd, Quebec City, QC, G1V 4G2, Canada; 3) Centre Hospitalier Universitaire de Québec - Pavillon Hôtel-Dieu de Québec, 11 Côte du Palais, Quebec City, QC, G1R 2J6, Canada; 4) Montreal General Hospital, Department of Medical Genetics, 1650 Cedar Ave, Montreal, QC, H3G 1A4, Canada; 5) Minneapolis, MN; 6) Program in Cancer Genetics, Department of Oncology and Human Genetics, McGill University, 546 Pine Ave West, Montreal, QC, H2W 1S6, Canada.

DICER1 is an endoribonuclease involved in the generation of miRNAs which modulate gene expression at the posttranscriptional level with a crucial role in embryogenesis and early development. Heterozygous germline mutations in *DICER1* (14q32.13) cause the DICER1 syndrome (OMIM #601200). This syndrome is characterized by mainly low penetrance for benign and malignant, mostly rare, tumors in children and young adults. Pleuropulmonary blastoma (PPB, before age 6 years) is the most frequent tumor and occurs in a highly treatable early cystic form which may progress to much-less-treatable advanced disease. Other phenotypes include cystic nephroma (CN), ovarian Sertoli-Leydig cell tumor (SLCT), multinodular goiter (MNG), various rhabdomyosarcomas, pituitary blastoma, nasal chondromesenchymal hamartoma, ciliary-body medulloepithelioma and other, rarer phenotypes. Cystic PPB is diagnosed from birth to age ≈2 years with a cure rate >90%; advanced Types II and III PPB occur from ages ≈2 to 6 years and are cured in ≈50% of cases. If not detected and treated, early cystic PPB may progress to advanced disease. We report here radiographic screening resulting in early detection of both cystic PPB and CN in an asymptomatic 13.5-month-old *DICER1* mutation carrier. The child was genotyped and screened because the large extended family had *DICER1* mutation (c.2379T>G, predicted to result in p.Y793X) and several members with SLCT, MNG, CN, and other PPBs. We believe this is the first child for whom knowledge of carrier status allowed focused early screening with the potential benefit of eradicating early PPB. Compared to PPB, other *DICER1*-associated phenotypes occur less frequently, occur in many organ systems over many years and have a better prognosis; taken together, these factors make most *DICER1* phenotypes less compelling targets for detection by pre-symptomatic screening. However, the natural history of PPB (presentation as highly curable disease in a well-defined age group with potential progression to an aggressive malignancy) offers the opportunity of early detection and eradication via pre-symptomatic screening. The potential benefits of early detection must be weighed against potential risks of iatrogenic harm due to early radiographic screening.

1066T

Impact of gynecological screening in Lynch syndrome carriers with an *MSH2* mutation. *S. Stuckless[1], P. Parfrey[1], L. Dawson[2], B. Barrett[1], J. Green[3].* 1) Clinical Epidemiology Unit, Faculty of Medicine, Health Sciences Centre, St. John's, NL, Canada; 2) Department of Obstetrics & Gynecology, Faculty of Medicine, Health Sciences Centre, St. John's, NL, Canada; 3) Department of Genetics, Faculty of Medicine, Health Sciences Centre, St. John's, NL, Canada.

Background & Aims: The lifetime risk of developing endometrial cancer in Lynch syndrome (LS) mutation carriers is very high and females are also at an increased risk of developing ovarian cancer. The aim of the study was to determine the impact of gynecological screening in *MSH2* mutation carriers. **Methods**: The study included 174 female carriers of three different founder *MSH2* mutations. Fifty four females entered a gynecological screening program and 120 females did not. Gynecological cancer incidence and overall survival was compared in female mutation carriers who received gynecological screening (cases) and in matched controls. Controls were randomly selected from non-screened mutation carriers who were alive and disease-free at the age the case entered the screening program. One matched control was selected for each case. **Results**: Endometrial or ovarian cancer occurred in 14 of 54 (26%) women in the gynecological screened group. Median age to diagnosis of gynecological cancer was 54 years in the screened group compared to 56 years in controls (p=0.50). Stage I or II cancer was diagnosed in 91% of screened patients compared to 71% in the control group (p=0.355). Two of three deaths in the screened group were the result of ovarian cancer. Mean survival in the screened group was 79 years compared to 69 years in the control group (p=0.11), likely associated with concomitant colonoscopy screening. **Conclusions**: In female LS carriers, gynecological screening did not result in earlier gynecologic cancer detection nor did it prevent interval cancers from occurring. The occurrence of both ovarian cancer deaths at a young age in screened patients and gynecological cancer deaths due to early stage disease in matched controls suggests that prophylactic hysterectomy with bilateral salpingo-oophorectomy should be performed in female mutation carriers who have completed childbearing.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1067F

Evaluation of mismatch repair gene sequence variants in the Colon Cancer Family Registry using multifactorial likelihood analysis. *B.A. Thompson[1,2], D.E. Goldgar[3], M. Clendenning[1], R. Walters[1], M.T. Parsons[1], S. Gallinger[4], R.W. Haile[5], J.L. Hopper[6], M.A. Jenkins[6], L. LeMarchand[7], N.M. Lindor[8], P.A. Newcomb[9], S.N. Thibodeau[10], J.P. Young[1], D.D. Buchanan[1], S.V. Tavtigian[11], A.B. Spurdle[1], Colon Cancer Family Registry.* 1) Queensland Institute of Medical Research, Brisbane, Australia; 2) School of Medicine, University of Queensland, Brisbane, Australia; 3) Department of Dermatology, University of Utah School of Medicine, Salt Lake City, Utah; 4) Cancer Care Ontario, Department of Surgery, Familial Gastrointestinal Cancer Registry, University of Toronto, Toronto, Ontario, Canada; 5) Department of Preventive Medicine, University of Southern California, Los Angeles, California; 6) Centre for Molecular, Environmental, Genetic and Analytic Epidemiology, University of Melbourne, Melbourne, Australia; 7) Cancer Research Center of Hawaii, University of Hawaii at Manoa, Honolulu, Hawaii; 8) Department of Health Science Research, Mayo Clinic Arizona, Scottsdale, Arizona; 9) Prevention Program, Fred Hutchinson Cancer Research Center, Seattle, Washington; 10) Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota; 11) Department of Oncological Sciences, Huntsman Cancer Institute, University of Utah.

Sequence variants of unknown significance constitute a challenge in both clinical and research settings. Mutations in the mismatch repair (MMR) genes (MLH1, MSH2, MSH6, PMS2) confer an increased risk of colorectal and endometrial cancers. Unclassified variants constitute over half of the unique MMR gene variants reported, and include rare nucleotide changes predicted to cause missense substitutions, small in-frame deletions, or possible alterations in splicing. We are developing a multifactorial approach to MMR variant evaluation by developing a likelihood model to provide a quantitative estimate of variant pathogenicity. A calibrated combined MAPP and custom PolyPhen-2 output has been developed to estimate a prior probability of pathogenicity for MMR missense substitutions based on the evolutionary conservation and physicochemical properties of amino acid alterations. This sequence analysis-based prior can be used in the context of a quantitative model for clinical classification of MMR gene missense substitutions. Analysis of variant co-segregation with disease within families to assess evidence for causality has been previously established and applied to MMR variant data. We have now derived likelihood ratios (LR) of variant causality for colorectal tumor characteristics, such as microsatellite instability (MSI) and somatic BRAF tumor mutation status, by drawing on comprehensive reference data from the Colon Cancer Family Registry (CFR). The LR in favor of pathogenicity was estimated to be ~12-fold for a colorectal tumor with a BRAF mutation-negative MSI-H phenotype. For 31 of the 44 (70%) variants already assessed, the posterior probabilities of pathogenicity were such that altered clinical management would be indicated. Strategies to incorporate LRs based on tumor MLH1 methylation status and family history data from the Colon CFR are under development. Estimation of the posterior probabilities for the remaining 245 variants detected in Colon CFR probands, for which assessment of tumor pathology and variant association with disease in families is available, is currently underway. Results from this analysis could alter the recommended clinical management of these and other MMR variant-carrying families.

1068W

Development of a Custom Targeted AmpliSeq Multiplex Sequencing Assay for the Detection of Tumor Derived Somatic Mutations and Application as Part of a Clinical Trial. *C.-J. Lih[1], D.J. Sims[1], M.G. Mehaffey[1], M. Dindinger[2], E. Levandowsky[2], C.C. Lee[2], T.T. Harkin[2], P. Brzoska[3], B. Conley[4], S. Kummar[5], P.M. Williams[1,4].* 1) Molecular Characterization Clinical Assay Development Laboratory, Frederick National Laboratory for Cancer Research, Frederick, MD; 2) Life Technologies, 500 Cummings Center, Suite 2450, Beverly, MA; 3) Life Technologies, 850 Lincoln Centre Drive, Foster City, CA; 4) National Cancer Institute, Cancer Diagnosis Program Division of Cancer Treatment and Diagnosis, Bethesda, MD; 5) National Cancer Institute, Division of Cancer Treatment and Diagnosis, Bethesda, MD.

Molecular targeted therapies are increasingly important in treating cancer patients, but robust analytically validated clinical assays are required for patient selection in early stage clinical trials. Here, we describe the development of a custom assay for somatic mutation detection using Ion Torrent AmpliSeq™ technology. The AmpliSeq™ technology enables generating hundreds of amplicons in a multiplex PCR pool from low amount of input DNA without the complications of primer dimers and cross amplification. A panel of 378 amplicons covering a total of 42.3 Mb regions was designed to detect 1152 mutation loci, including common therapeutically actionable targets, in 63 genes. Analytical performance studies using genomic DNA samples from 14 cancer cell-lines and a normal hapmap individual (CEPH) demonstrated the assay was reproducible using 30 ng input of DNA. The assay demonstrated high sensitivity and specificity (less than 1 percent). Using a series of positive control plasmid spikes mixed into a normal reference CEPH DNA at pre-defined copy number ratios, we verified the assay is sensitive to detect variants at 5 percent (or less) allelic frequency with a minimum 500 X coverage. We subsequently applied this assay to characterize proficiency samples whose mutations status have been previously reported by other analytically validated clinical assays (pyrosequencing or SNAPshot). Our data demonstrated this assay recapitulated all previously identified mutations in a semi-quantitative way. We conclude this assay is accurate, specific, sensitive, and reproducible with rapid turn-around time, making it a powerful tool for supporting clinical studies. The application of this assay for planned clinical studies will be discussed.

1069T

A 30 year perspective of The International Fanconi Anemia Registry (IFAR). *A.D. Auerbach, The IFAR Collaboration Team.* Human Gen & Hematology, Rockefeller Univ, New York, NY.

The IFAR was founded in May 1982, by Dr. Arleen D. Auerbach, in collaboration initially with Dr. Traute Schroeder-Kurth of Germany. Schroeder-Kurth, in 1964, discovered the chromosome instability in Fanconi anemia patients and Auerbach, in 1981, developed the diepoxybutane (DEB) chromosome breakage test. Patients with a positive DEB test were eligible to participate in the IFAR study. The purpose of the IFAR is to study a large number of patients exhibiting the full spectrum of diverse features of FA and to study genotype/pheonotype correlations. Studies of IFAR patients led to the discovery that the clinical features of the syndrome were much more varied than originally thought, and to the recognition of the need for a more timely diagnosis of the disease. The IFAR Tissue Repository of specimens from patients with a positive DEB test aided researchers to identify new genes for FA and to find a large spectrum of variants in these genes that caused the disease. Selected highlights from clinical IFAR publications are: Giampietro PF et al. The need for more accurate and timely diagnosis in FA. *Pediatr. 91:1116,1993*; Butturini A et al. Hematologic abnormalities in FA. *Blood 84:1650,1994*; Giampietro PF et al. Diagnosis of FA in patients without congenital malformations. *Am J Med Genet 68:58,1997*; Wajnrajch MP et al. Evaluation of growth and hormonal status in patients referred to the IFAR. *Pediatr. 107:744,2001*; Kutler DI et al. A 20 year perspective of the IFAR. *Blood 101:1249,2003*; Berwick M et al. Genetic heterogeneity among Fanconi anemia heterozygotes and risk of cancer. *Cancer Res 67:9591, 2007*. The IFAR is currently based in The Rockefeller University Laboratory of Genome Maintenance, headed by Dr. Agata Smogorzewska: http://lab.rockefeller.edu/smogorzewska/ifar/. There are presently 1167 patients enrolled in the IFAR, 557 females and 610 males. Approximately 40% of the patients, mostly alive at the time of registration, are known to be deceased. We have now developed an enhanced second generation IFAR database. Included in the IFAR are all genotypic data available from our current collaboration with NHGRI and the most up-to-date, streamlined, ontology driven phenotypic information. There are new display screens and navigational tools which should make it easier to learn more about genotype/phenotype correlations, the role of the FA proteins in damage repair, the role of FA proteins in cancer, and the implications of mosaicism.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1070F

Histology and prognostic markers are powerful tools for predicting BRCA mutation status. *C.D. Delozier[1, 2,3], A. Rahman[4], T. Liu[5], C.J. Curry[1,2,3]*. 1) Genetic Medicine Central California, Central California Faculty Medical Group, Fresno, CA; 2) Dept Pediatrics, UCSF-Fresno, Fresno, Ca; 3) Saint Agnes Medical Center, Fresno, CA; 4) Dept Internal Medicine, UCSF-Fresno, Fresno, CA; 5) Dept Biology, UCSD, San Diego, CA.

Five to seven percent of breast cancer is associated with a constitutional mutation in BRCA-1 or BRCA-2, the principle genes whose mutation results in “monogenic inheritance” of breast and ovarian cancer. The best single predictor of BRCA mutations in breast cancer patients is pre-menopausal age at diagnosis (especially for BRCA1); personal and/or family histories of multiple breast primaries, ovarian cancer and/or male breast cancer are also characteristic of BRCA-positive families. However, many patients have unknown/truncated family histories, and cancer diagnostic information as related by families may be inaccurate. In patients with a “borderline” probability of a BRCA mutation, the decision to test (or not) is often determined by insurance coverage. Although numerous publications in the late 1990s attested to a “typical histological phenotype”, particularly in BRCA1-associated breast cancers, histology (ductal vs. lobular phenotype) and prognostic markers (ER,PR,Her2/neu) have been displaced as major determinates in genetic risk assessment. Our experience in the BRCA 1–2 testing of over 300 breast cancer patients at “high-risk” for hereditary breast and ovarian cancer syndrome strongly suggests that histology (ductal vs. lobular vs. rarer histologies) and prognostic marker status (triple negative status, her2/neu overexpression) are key results for evaluating the likelihood of BRCA mutations. Our BRCA-positive patient population includes 26 women with BRCA-1 mutations and 17 with BRCA-2 mutations; these women had 48 breast cancers. *No BRCA-1 mutations were found in patients with tumors of lobular histology or in tumors with Her-2/neu overexpression*. Although the catagorization was not as clear with BRCA-2, lobular histology and Her-2/neu overexpression were also rare in BRCA2-associated breast cancers. One patient (of 17) who had a premenopausal ductal carcinoma had a contralateral post-menopausal lobular primary. In another BRCA2 carrier with bilateral breast, the second, postmenopausal ductal cancer was Her2/neu positive. These proportions are to be compared to histological findings in all breast cancers, where 10–20% of tumors are lobular carcinomas and about 20% show Her2/neu overexpresion. We have used histology - performed on every tumor and thus available for assessment in all breast cancer patients - to develop a scoring tool for predicting BRCA mutation status.

1071W

Variants of uncertain significance in BRCA testing: Surgical decisions, risk perception, cancer distress, and evaluation of the genetic cancer risk assessment process. *J. O. Culver[1], C. D. Brinkerhoff[2], J. Clague[1], K. Yang[1], K. E. Singh[2], S. R. Sand[1], J. N. Weitzel[1]*. 1) City Hope, Duarte, CA; 2) UC Irvine, Irvine, CA.

Incidental identification of variants of uncertain significance (VUS) is one of the greatest challenges of new clinical genomic testing approaches. Previous studies suggest that patients carrying a VUS in BRCA1 or BRCA2 can have lingering confusion regarding results interpretation and healthcare recommendations. This may be due in part to inappropriate recommendations made by inexperienced or misinformed healthcare providers. In standard clinical genetics practice, VUS results in BRCA genes are considered uninformative, and thus not used to determine risk or make recommendations-similar to uninformative negative (UN) results. This study sought to compare VUS patients and UN patients receiving genetic cancer risk assessment (GCRA) in high risk genetics clinics with respect to surgical decisions, risk perception, distress levels, and opinions about the GCRA process. Follow-up questionnaires from 71 VUS patients were compared to 715 UN patients, all having undergone GCRA at City of Hope Clinical Cancer Genetics Community Research Network. Statistical analyses were performed with Pearson's Chi-square and Fisher's Exact test, as appropriate. Ten percent of VUS carriers and 6% of UN patients opted for risk reducing mastectomy (p=0.15) and 5% of VUS and 3% of UN patients opted for risk reducing bilateral salpingo-oophorectomy (p=0.44). Of those able to recall their breast cancer risk, 15% of VUS patients and 10% of UN patients perceived themselves at high risk (p=.31); similar results were observed for ovarian cancer risk. UN patients showed a greater reduction in concern about cancer after GCRA than did VUS patients (36% versus 23%, p=.043). Among the minority of patients (7% of VUS and 3% of UN) that did not find the counseling process helpful, more VUS patients attributed this to a lack of new knowledge gain (100% versus 48%, respectively, p=.033). Overall, results suggest that patients receiving VUS and UN results in a GCRA clinic environment react similarly with respect to preventive surgical decisions (VUS results did not cause excessive surgery or exaggerated risk perception), while patients with a VUS found counseling somewhat less informative and less reassuring. Given that VUS patients receiving GCRA yielded similar clinical outcomes to UN patients, skilled cancer risk counseling will continue to be a critical component of genomic testing strategies involving multi-gene panels and exomic testing given the high likelihood of VUS results.

1072T

THE CONTRIBUTION OF PREDISPOSITION POLYMORPHISMS FOR BODY SIZE TO THE INHERITED RISK OF COLORECTAL CANCER. *L. Carvajal-Carmona[1], G. Casey[2], O. Sieber[3], G. Montgomery[4], J. Young[5], P. Baird[6], D. Kerr[7], R. Houlston[8], M. Dunlop[9], I. Tomlinson[1,10], The Colon Cancer Family Registry, The COGENT Consortium.* 1) Wellcome Trust Ctr Human Gen, Univ Oxford, Oxford, United Kingdom; 2) USC/Norris Comprehensive Cancer Center, University of Southern California, Los Angeles, California, USA; 3) Ludwig Colon Cancer Initiative Laboratory, Ludwig Institute for Cancer Research, Royal Melbourne Hospital, Parkville, Victoria, Australia; 4) Genetic and Molecular Epidemiology Laboratories, Queensland Institute of Medical Research, Herston Q4006, Australia; 5) Familial Cancer Laboratory, Queensland Institute of Medical Research, Herston Q4006, Australia; 6) Centre for Eye Research Australia, University of Melbourne, 32 Gisborne Street, East Melbourne, VIC 3002, Australia; 7) Department of Oncology, Oxford University, Old Road Campus Research Building, Oxford, OX3 7DQ, United Kingdom; 8) Section of Cancer Genetics, Institute of Cancer Research, Sutton SM2 5NG, United Kingdom; 9) Colon Cancer Genetics Group, Institute of Genetics and Molecular Medicine, University of Edinburgh and MRC Human Genetics Unit, Edinburgh EH4 2XU, United Kingdom; 10) NIHR Comprehensive Biomedical Research Centre, Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford OX3 7BN, United Kingdom.

Background. Despite recent successes, most of the heritability of colorectal carcinoma (CRC) remains unexplained. One possible contributor is common genetic variants with individually small effects which act through variation in normal traits, such as height and obesity, that are associated with CRC risk. Methods. We studied five series of northern European ancestry, comprising a total of 5,380 CRC cases and 6,034 controls. We identified from the literature common autosomal SNPs that have been associated (Passoc<5×10-8) in large studies with adult height (N=134), body mass index (BMI, N=29) and waist-hip ratio (WHR, N=13). We extracted genotypes for these SNPs or excellent proxies (r2>0.8) and performed single-SNP and SNP-set association analyses with CRC risk. Results. No individual SNP was significantly associated with CRC risk, although there was an overall over-representation of associations in the expected direction for increased body size. SNP-set analyses, particularly the allele score test, provided good evidence that SNP alleles associated with obesity (WHR and BMI) were also associated with an increased risk of CRC. We found no evidence of similar effects for height. Conclusions. WHR- and BMI-increasing SNP alleles are associated with raised CRC risk, consistent with epidemiological data. Effect sizes are very modest, consistent with an intermediate phenotype model in which obesity causally affects CRC risk. The absence of associations with height SNPs suggests that the link between height and CRC may not be causal. Obesity SNPs explain a small, but important, part of CRC risk, and normal, heritable traits may have similar influences.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1073F

The combined effect of the functional variants in cell death pathway genes on cervical cancer. *K. Chattopadhyay[1], C. Dandara[2], A. Hazra[3], A.L. Williamson[4, 5].* 1) School of Life Sciences, College of Agriculture, Engineering and Science, University of KwaZulu-Natal, Durban, Republic of South Africa; 2) Division of Human Genetics, Faculty of Health Science, University of Cape Town, Cape Town, Republic of South Africa; 3) School of Statistics and Actuarial Sciences, University of KwaZulu-Natal, Durban, Republic of South Africa; 4) Division of Medical Virology and Institute of Infectious Disease and Molecular Medicine (IIDMM), University of Cape Town, Cape Town, Republic of South Africa; 5) National Health Laboratory Service, Groote Schuur Hospital, Observatory, Cape Town, Republic of South Africa.

Background: Cervical cancer is one of the most important cancers worldwide with a high incident and mortality rate. Even though many sexually active women get infected with human papillomavirus (HPV), only a small fraction of them progress to cervical cancer suggesting the important roles of the additional risk factors in development of the disease including host genetic factors. Since cellular apoptosis plays a crucial role in controlling the spread of virus-infections in cells, functional variations in cell death pathway genes might alter the apoptotic mechanism, thereby influencing its ability to clear virus-infections. Functional polymorphisms in FasR (-1377G>A and -670A>G), FasL (-844T>C) and CASP8 (-652 6N ins/del) genes are all know to alter the mechanism of apoptosis by modifying the level of expression of their correspondent proteins. These polymorphisms have been investigated in cervical cancer by different groups including our lab previously. **Objective:** In the present study, we attempted to investigate the combined effects of CASP8 polymorphism with any (and all) of the FasR/FasL polymorphisms in cervical cancer, pre-cancerous lesions and HPV infection. **Materials and Methods:** Participants were 442 South African women of black African and mixed-ancestry origin with invasive cervical cancer and 278 control women matched by age, ethnicity and domicile status. FasR and FasL polymorphisms were genotyped by TaqMan and CASP8 polymorphism by PCR-RFLP. The study was approved by the University of Cape Town Human Research Ethics Committee (study approval number, REC REF: 075/2009) and was performed in accordance with the ethical standards of the Declaration of Helsinki. The results were analysed using haplo.stats software version 1.5.2. All results were adjusted for ethnicity and smoking. **Results:** No significant association with any of the combinations of the polymorphisms was found with cervical cancer. Stratifying the data for smoking and HSV positivity status also did not show any significance. Hybrid capture, pap smear, smoking and HSV positivity status were also investigated among only control group which showed no significant association. **Conclusion:** This is, to our knowledge, the first time the combined effects of the functional variants in cell death pathway genes have been investigated in cervical cancer. Our results show that the combined effect of the functional variants in cell death pathway genes do not play a role in cervical cancer.

1074W

Genetic ancestry is associated to changes in DNA Repair Capacity in non-melanoma skin cancer patients from Puerto Rico. *H.J. Diaz-Zabala[1], L. Morales[2], J. Matta[2], J. Dutil[1].* 1) Biochemistry Department, Ponce School of Medicine, Ponce, PR; 2) Pharmacology Department, Ponce School of Medicine, Ponce, PR.

Among Caucasians and light-skinned individuals, skin cancer occurs with higher frequency. In the Hispanic population of Puerto Rico, the contribution of genetic ancestral factors to non-melanoma skin cancer risk is unknown. Previous studies have shown that the DNA repair capacity (DRC) is reduced in skin cancer patients. A study population of 83 BCC and SCC skin cancer patients and 89 (non-cancerous) controls were recruited from private dermatology practices in Ponce, Puerto Rico. A panel of 106 ancestry-informative markers was genotyped using iPLEX Sequenom technology. Individual genetic ancestry was estimated using a maximum likelihood approach. The DRC levels were measured using a modified host-cell reactivation assay. African ancestry proportion was significantly associated with dark-skinned color ($P < 0.01$) and the absence of freckles in the body ($P = 0.023$). European ancestry was significantly associated with phenotypes such as light-skinned color ($P = 0.001$), hair color ($P = 0.037$), presence of freckles ($P = 0.013$) and susceptibility to sunburn ($P = 0.046$). There was no significant difference in the average European ($P = 0.822$), African ($P = 0.395$), or Native American ($P = 0.451$) ancestry proportions in non-melanoma skin cancer cases compared to non-cancer controls. A linear regression analysis of genetic ancestry and DRC suggests that African ancestry is positively correlated ($P = 0.042$) to DRC, whereas European ancestry is negatively correlated ($P = 0.053$) to DRC. Our data indicate that genetic ancestry is not associated with skin cancer risk. However, African and European ancestries were associated to changes in DRC levels, skin and hair pigmentation, and the susceptibility to sunburn.

1075T

Genetic ancestry and breast cancer survival in US Latinas. *L. Fejerman[1,2,3], D. Hu[1], S. Huntsman[1], C. Eng[1], E. Gonzalez Burchard[1,2,4], E.M. John[5,6], M.C. Stern[7], E. Perez-Stable[1], E. Ziv[1,2,3].* 1) Dept Med, Univ California San Francisco, San Francisco, CA; 2) Institute of Human Genetics, Univ California San Francisco, San Francisco, CA; 3) Comprehensive Cancer Center, Univ California San Francisco, San Francisco CA; 4) Dept of Biopharmaceutical Sciences and Medicine, Univ California San Francisco, San Francisco CA; 5) Cancer Prevention Institute of California, Fremont, CA; 6) Dept. of Health and Research Policy, Stanford University, Palo Alto, CA; 7) Dept. of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA.

Latinas in the US are diagnosed with more advanced stage breast cancer and have poorer survival than non-Latina White women. Latinas are an admixed population with European, Indigenous American and African genetic ancestry. We evaluated the association between genetic ancestry and survival in Latinas from Northern California to test the hypothesis that the observed difference in mortality might be partly due to population specific genetic factors. We used a Cox proportional hazards model to assess the association of Indigenous American genetic ancestry with breast cancer survival among 927 women with breast cancer who participated in two population-based studies. Genetic ancestry was estimated from SNPs from an Affymetrix 6.0 array. We further evaluated whether the relationship between genetic ancestry and breast cancer specific survival was explained by differences in socioeconomic status, education, age at diagnosis or tumor characteristics. Cases diagnosed between 1995 and 2004 were followed for a total of 8,619 person-years during which 76 died from breast cancer. The median follow up was 9 years (range 2–16 years). In the unadjusted model, an increase in Indigenous American ancestry of 25% was associated with an increased risk of mortality [Hazard Ratio (HR): 1.55; 95%CI: 1.06–2.26, p=0.02]. Adjustment for education and socioeconomic status slightly attenuated the association (HR: 1.49, 95%CI: 1.01–2.21, p=0.05). Further adjustment for age at diagnosis reduced the HR to 1.37 (95%CI: 0.93–2.03, p=0.11). In a subset of cases with complete information on tumor characteristics (601 cases, 48 events), age at diagnosis, stage and grade, but not hormone receptor status attenuated the HR for genetic ancestry. In this reduced set of cases, the unadjusted and adjusted associations with genetic ancestry were no longer statistically significant. Our results suggest that the observed disparity in breast cancer specific mortality between US Latinas and non-Latina Whites might be partly due to differences in patient's age, tumor stage and grade at diagnosis. Further work exploring possible risk factors for early age at diagnosis, advanced stage and grade in US Latinas, including genetic predisposition, should be of high priority since it could lead to a reduction in breast cancer outcome disparities.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1076F

Genetic polymorphisms in candidate inflammation-related genes and risk of esophageal squamous cell carcinoma. *A. Golozar[1,2], T. Beaty[1], P. Gravitt[1], I. Ruczinski[1], Y.L. Qiao[3], J.H. Fan[3], T. Ding[3], Z.Z. Tang[3], A. Etemadi[2], N. Hu[2], S.M. Dawsey[2], N.D. Freedman[2], C.C. Abnet[2], A. Goldstein[2], P.R. Taylor[2].* 1) Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD; 2) Division of Cancer Epidemiology and Genetics, NCI, Bethesda, MD; 3) Chinese Academy of Medical Sciences, Beijing, China.

Wound healing mechanisms are thought to play some role in cancer development in multiple organs, likely by creating a chronic inflammatory milieu. This study sought to assess the role of common genetic polymorphisms in 13 inflammation-related genes involved in wound healing in the etiology of esophageal squamous cell carcinoma (ESCC). The selected genes, include genes in the IL-1 family (*IL-1a, IL-1b, IL-1 Receptor type I (IL-1Ra) and type II (IL-1Rb)), TNF-α, TNFRSF1A, NF-kB1, NF-kB2, iNOS, COX-2, HIF-1α, VEGFA and P-53*). We used the data from the Nutritional Intervention Trial (NIT) cohort and the Shanxi Case-Control Study, both conducted in northcentral China between 1985–2005. From the NIT, 410 ESCC cases and 377 controls frequency matched for age and sex, and from Shanxi, 546 cases and 556 controls individually matched for age, sex and neighborhood were genotyped for 125 tag single nucleotide polymorphisms (SNPs). Tag SNPs were located within the 20 kb upstream and 10 kb downstream of each gene. All SNPs had a minor allele frequency of ≥5%; and a pairwise linkage disequilibrium of $r^2 \geq 0.8$. Cox proportional-hazard models with robust variance estimator and conditional logistic regression models were used for individual SNP analyses for the NIT and Shanxi, respectively. The empirical significance of the enrichment pattern of these findings was assessed using 5000 permutations under the null hypothesis. Fisher's inverse test statistics followed by permutation tests was used to obtain gene-level significance. Of the 125 tag SNPs, 8 SNPs (4 in *VEGFA*, 1 in *TNF-α*, and 3 in *IL-1R1*), and 5 SNPs (1 in *iNOS*, 2 in *TNFRSF1A*, and 3 in *VEGFA*) showed nominally significant associations with risk to ESCC in the NIT and Shanxi, respectively. None of these SNPs retained statistical significance after strict Bonferroni correction. However, in the NIT, the observed p-values for the tested tag SNPs for *VEGFA* gene were mostly concentrated in the lower tail of the distribution (unlike what is expected under the null hypothesis). Permutation test confirmed the significance of this enrichment pattern. At the gene level, *VEGFA* yielded an empiric significance equivalent to 0.023 in the NIT. Given the role of VEGFA in the development and progression of early precancerous lesions of esophagus, the empirically significant association between markers in the *VEGFA* gene and ESCC observed in the NIT provides further evidence for some role of VEGFA in ESCC development.

1077W

Frequency of VNTR polymorphism in IL-1Ra gene in colorectal cancer. *I. A. Gutíerrez[1,3], E. Cortes[2], A. M. Puebla[4], M. P. Gallegos[3].* 1) Genética Molecular, CIBO, IMSS, Guadalajara, Jalisco, Mexico; 2) Departamento de Oncocirugia, UMAE, HE, CMNO; 3) División de Médicina Molecular, CIBO, IMSS, Guadalajara, Jalisco, México; 4) Laboratorio de inmunofarmacologia, UdeG.

The IL-1ra protein (IL-1Ra), has been widely investigated in and found to be associated with different types of cancers. In the second intron of the IL-1Ra gene, are localized a polymorphism VNTR type, which is characterized by participated in the expression regulating of gene, immune response and cancer risk. We analyzed this VNTR of IL-1Ra in 92 colorectal cancer patients and 118 controls from Guadalajara, Jalisco. Mexico, and found a genotype frequency of 45% (41/92) and 36% (42/118) of A1/A1; 17% (16/92) and 22% (26/118) of A2/A2; 38% (35/92) and 42% (50/118) of A1/A2 in cases and controls respectively, without to be statistically different (p>0.05). These results suggest that the VNTR IL-1RN polymorphism in this sample is not associated with Mexican colorectal cancer patients.

1078T

Individual ancestry influences the gene expression profile in breast cancer patients of admixed populations. *D. Hu[1], E. Ziv[1,2,3].* 1) Department of Medicine, University of California San Francisco, San Francisco, CA; 2) Helen Diller Family Comprehensive Cancer Center, University of California San Francisco, San Francisco, CA; 3) Institute for Human Genetics, University of California San Francisco, San Francisco, CA.

Gene expression is a significant predictor of survival from breast cancer and is currently used clinically to select therapy for estrogen receptor (ER)-positive breast cancer tumors. The degree to which gene expression profiles might be altered by the genetic ancestry of the patient and how that may affect the association between expression profile and survival is unknown. We addressed the question of how gene expression is affected by genetic ancestry in breast cancer samples from The Cancer Genome Atlas (TCGA). We downloaded microarray and genotype data of 481 samples from TCGA Data Portal. These 481 samples include 39 African Americans, 32 Asians, 6 Hispanics, 313 Caucasians, and 91 patients from unknown ethnic groups. Gene expression was analyzed using Agilent G4502A platform. Genetic variation was measured by Affymetrix Genome-Wide Human SNP Array 6.0. To infer genetic ancestry, we calculated principal components using the genotype data. We used the top 3 principal components to represent the patient's global ancestry. We focused on gene expression profiles from patients with ER-positive tumors (N=364). Gene expression profile was studied by hierarchical clustering. We first performed clustering analysis on the original expression values. We then conducted linear regression on gene expression and the top 3 principal components. Residuals from the linear regression were used in the clustering analysis again. We compared the clustering results from the original gene expression values and the residuals. We found changes in clustering results after individual's ancestry was used in the analysis. Our results suggest that gene expression profile might be influenced by the global ancestry of breast cancer patients. The degree to which these differences affect the association between gene expression and survival should be evaluated in large cohort studies with diverse populations.

1079F

Risks of colorectal and other cancers following endometrial cancer in Lynch syndrome. *M.A. Jenkins[1], N.M. Lindor[2], R.W. Haile[3], P.A. Newcomb[4], L. Le Marchand[5], S. Gallinger[6,7], J.L Hopper[1], A.K. Win[1].* 1) Centre for Molecular, Environmental, Genetic and Analytic Epidemiology, The University of Melbourne, Parkville, Victoria, Australia; 2) Department of Health Science Research, Mayo Clinic Arizona, Scottsdale, Arizona, USA; 3) Department of Preventive Medicine, University of Southern California, Los Angeles, California, USA; 4) Cancer Prevention Program, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA; 5) University of Hawaii Cancer Center, Honolulu, Hawaii, USA; 6) Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, Ontario, Canada; 7) Cancer Care Ontario, Toronto, Ontario, Canada.

Background: Lynch syndrome is an autosomal dominantly inherited disorder of cancer caused by germline mutations in DNA mismatch repair genes. Previous studies have shown that mismatch repair gene mutation carriers are at increased risk of first primary colorectal, endometrial and several additional cancers. We aimed to estimate the risks of primary cancers following endometrial cancer for mutation carriers to determine if those with initial endometrial cancer might have higher subsequent risks for new primary cancers. Methods: We obtained data from the Colon Cancer Family Registry for a cohort of 196 women carrying a pathogenic mutation in a mismatch repair gene (56 MLH1, 111 MSH2, 26 MSH6 and 3 PMS2) who had a previous diagnosis of primary endometrial cancer. We estimated the age-, sex-, country- and calendar period-specific standardized incidence ratios (SIRs) of subsequent primary cancers following endometrial cancer, compared with the general population. Kaplan-Meier method was used to estimate 10- and 20-year cumulative risk (penetrance) for each cancer. Results: For carriers of mismatch repair gene mutations who had a previous diagnosis of endometrial cancer, we observed significant increased risks of colorectal cancer (SIR 22.80; 95% confidence interval 15.80–31.89), small intestinal cancer (56.93, 12.51–123.45), kidney, renal pelvis or ureteric cancer (23.02, 10.63–37.26), urinary bladder cancer (22.19, 9.66–36.90), hepatobiliary cancer (12.97, 2.93–27.35), and breast cancer (2.99, 1.80–4.42). The cumulative risks at 20-year after endometrial cancer were as follow: colorectal cancer (35%, 24–46%), small intestinal cancer (2%,0–6%), kidney, renal pelvis or ureteric cancer (9%, 4–15%), urinary bladder cancer (7%, 2–13%), hepatobiliary cancer (2%, 0–6%) and breast cancer (13%, 7–20%). Conclusions: For Lynch syndrome patients, the estimated risks of subsequent primary cancers may inform appropriate cancer screening and risk reducing strategies following an endometrial cancer diagnosis.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1080W

Prevalence and clinical features of *HOXB13* mutation carriers in prostate cancer patients from Germany. *C. Maier[1], J. Xu[2], S.L. Zheng[2], M. Luedeke[1], A.E. Rinckleb[1], W. Vogel[3], S.N. Thibodeau[4], K.A. Cooney[5], W.B. Isaacs[6] on behalf of the ICPCG.* 1) Department of Urology, University of Ulm, Ulm, Germany; 2) Center for Cancer Genomics, Wake Forest University School of Medicine, Winston-Salem, NC; 3) Institute of Human Genetics, University of Ulm, Ulm, Germany; 4) Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN; 5) Department of Internal Medicine, University of Michigan Medical School, Ann Arbor, MI; 6) Department of Urology, Johns Hopkins Medical Institutions, Baltimore, MD.

Although prostate cancer (PrCa) has the highest heritability among common cancers, specific high risk genes responsible for familial clustering of PrCa have remained undiscovered for decades. Recently, by deep sequencing of a particular hereditary prostate cancer linkage region at 17q21–22, a recurrent missense mutations (G84E) in the homeobox gene *HOXB13* was identified in cases of European American descent. In order to assess the role of *HOXB13* in the European subpopulation of Germany, case-control comparisons were performed on PrCa cases with (n = 379) and without (n = 317) family history and 490 geographically matched population controls. All subjects were genotyped for G84E in the course of a multi-SNP analysis conducted by the International Consortium for Prostate Cancer Genetics (ICPCG) using an iPLEX MassArray. Among the present study cohort carriers of *HOXB13* G84E were rare in controls (0.02%) and showed increased frequencies in both familial (3.2%) and sporadic PrCa cases (1.9%). Estimated risks of OR = 16.0 (p = 0.008) for familial cases and OR = 9.4 (p = 0.013) for sporadic cases were similar to results from initial reports, consistent with the high penetrance role of *HOXB13* in PrCa susceptibility. Predisposition of *HOXB13* G84E carriers for aggressive PrCa was considered by subgrouping for high histological grade (Gleason score ≥7), advanced tumor stage (>T2, or N+, M+), high PSA at diagnosis (> 20ng/ml) or lethal PrCa. Patients with parameters of aggressive disease showed elevated frequencies of G84E mutation carriers in both familial (3.8%) and sporadic cases (2.5%). The strongest association was observed between G84E and tumor stage. Among patients diagnosed with locally advanced or metastatic disease, 4.4% of familial (OR = 22.7; p = 0.004; as compared to controls) and 3.7% of sporadic cases (OR = 18.5; p = 0.008; as compared to controls) were found to carry the *HOXB13* mutation. In conclusion, the percentage of familial clustering in the German study cohort explained by *HOXB13* is comparable to the observations in US-American populations of European ancestry. However, as an association between G84E and advanced tumor stage was also apparent in sporadic cases, a larger fraction of *HOXB13* mutation carriers may be found among aggressive PrCa cases regardless of family history, rather than within high risk PrCa pedigrees.

1081T

Histologic Types and Risk Factors in Familial Lung Cancer Cases from Southern Louisiana. *D. Mandal[1], M. Haskins[1], A. Bencaz[1], J. Hutchinson[1], J. Chambliss[1], H. Rothschild[1], J.E. Bailey-Wilson[2].* 1) Dept Gen, LSU Hlth Sci Ctr, New Orleans, LA; 2) National Human Genome Research Institute, National Institutes of Health, Baltimore, MD.

The association between lung cancer (LC) and smoking is well known. However, only 15% of smokers are diagnosed with LC. In addition, about 10% of LC cases (22,000 cases per year in the U.S.) have at least one relative affected with LC. So, family history is a significant risk factor. In a case control study in Germany, adenocarcinoma was most prevalent in never smokers and in women; squamous cell carcinoma was most prevalent in male smokers. In the 1990's in the population in Louisiana, squamous cell carcinoma was also observed to be the most frequent type in LC cases in general(39.3%), with nearly equal numbers of adenocarcinoma (25.0%) and small cell subtypes (25.5%). While some studies have shown similar proportions of cases with a family history in all histologic subtypes of LC, others have suggested that a higher proportion of patients with squamous cell carcinoma have a family history of LC. However, less is known about the proportions of each histologic type in LC cases with a strong family history.The objective of the present study was to analyze histologic subtypes and their association with smoking behaviors and other risk factors of interest among familial LC cases from southern Louisiana. Eligible subjects (N=148) with ≥2 relatives affected with primary lung cancer were recruited from southern Louisiana. Diagnosis of primary LC was confirmed through medical records, and histologic subtype (N=114) was abstracted from pathology reports. About 81% of cases had non-small cell LC, with adenocarcinoma (40%) being the most common histologic subtype, followed by squamous cell at 28% and small cell at 18%. Histologic type of LC was most strongly related to pack-years of cigarette smoking and age at diagnosis in these familial cases, with squamous cell associated with higher mean pack-years of smoking and older age. Preliminary results indicated significant difference in the age of diagnosis between cases with non-small cell LC and those with small cell LC; mean number of pack years was about twice as high in cases with non-small cell LC. The frequency of adenocarcinoma in these familial cases was higher and of squamous cell carcinoma was lower than previously reported for the Louisiana LC population. This is consistent with a higher risk of adenocarcinoma in persons with a family history, which may be related to our previous observations that less smoking appears to be necessary for LC in persons with a familial risk of LC.

1082F

Using an 18 locus genotype to predict risk of breast cancer, and the relationship between age, estrogen receptor status and genetic risk. *C. Merrick[1], C. Purdie[2], L. Jordan[2], C. Palmer[3], R. Tavendale[3], A. Ashfield[4], P. Quinlan[4], P. Armory[1], A. Thompson[4], J. Berg[1].* 1) Genetics, University of Dundee, Dundee, United Kingdom DD1 9SY; 2) Department of Pathology, Tayside University Hospitals Trust, Ninewells Hospital, Dundee, UK; 3) Biomedical Research institute, University of Dundee, Dundee; 4) Surgery and Molecular Oncology, University of Dundee, Dundee, UK.

The majority of the inherited risk of breast cancer is caused by multiple low penetrance polymorphisms. Whilst the effect of each polymorphism is individually too small to be clinically useful, combining information from many loci is expected to provide improved risk estimation. We investigated the performance of an 18 locus genotype in a cohort of 815 women affected with breast cancer in Tayside from the Tayside Tissue Bank, and 960 matched controls from the Generation Scotland GS:3D repository. DNA samples were analysed using Sequenom Massarray technology. For each individual, genotype data from all 18 loci were combined into a single genetic risk score using a log additive model. Data were analysed using Area Under the Receiver Operating Characteristic Curve (AUROC) analysis. Overall, log genetic risk was normally distributed and significantly higher in cases than controls, AUROC 0.605 (95% CI 0.58–0.63). Age, estrogen receptor (ER) status and genotype risk was available for 506 cases. In these cases, the genotype risk was higher for ER positive cases, AUROC 0.611 (95% CI 0.58–0.64) than ER negative cases, AUROC 0.585 (95% CI 0.528–0.642). When stratified into 10 year age bands, the mean of the log relative risk decreased with increasing age. Those aged 35–44 had the highest log genetic relative risk with a mean of 0.0472 (95% CI 0.0069–0.0875), whereas those aged 75–85 had the lowest log relative risk with a mean of 0.0225 (95% CI -0.0122–0.0572). Pearson's correlation coefficient demonstrated a similar modest trend of -0.028. We have shown that an 18 locus genotype predicts risk of breast cancer, for both ER +ve and ER -ve cancers, although with a stronger association for ER +ve cancers. The genotype risk score is higher in women affected at a younger age, as expected for a classical multifactorial trait. Combination of multiple low penetrance genetic loci into a single score, therefore, shows promise as an easy and objective means of estimating risk of breast cancer, especially in younger women. However, to be clinically useful, a greater degree of risk prediction is required, either with the use of additional genetic loci, or by combining genetic risk with other established risk factors such as family history or mammographic density into a multi-modal risk model.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1083W

Association between Prostate Cancer Occurrence and Y-Chromosomal STRs. *M. Miri Nargesi[1], A. Nazemi[1], P. Amini[2], P. Ismail[3], R. Vasudevan[3], P. Pasalar[4], A.H. Abdul Razack[5].* 1) Department of Genetics, Tonekabon branch-Islamic Azad University, Tonekabon, Mazandaran, Iran; 2) Department of Pathology, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, Malaysia; 3) Department of Biomedical Science, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, Malaysia; 4) Department of Biochemistry, School of Medicine, Tehran University of Medical Sciences, Iran; 5) Department of Surgery, Faculty of Medicine, University of Malaya, Malaysia.

Purpose of Study: Different studies have shown that there are variations in the geographic and racial distribution of cancer of prostate. This disease is influenced by a complex genetics, and it supposed that a similarity in genetic factors may affect the susceptibility or resistance to develop this disease. This study aimed to compare prostate cancer patients and normal control males based on four Y-linked polymorphic microsatellite loci including DYS388, DYS435, DYS437, and DYS439. Materials and Methods: The study group included a total of 376 subjects comprising 182 patients and 191 healthy individuals. Samples were collected from different populations including different ethnic groups in Iran and Malaysia. Ethics clearance for Malaysian and Iranian samples were acquired from University of Malaya (Reference number: 703.55) Tonekabon Islamic Azad University (Reference number: 22899), respectively. A multiplex PCR was optimized to co-amplify DYS388, DYS435, DYS437, and DYS439 loci in a single reaction tube. All samples were genotyped for alleles of four DYS loci using a Genetic Analysis System CEQ8000 (Beckman Coulter). Results: The comparison between allelic frequencies of DYS loci between patients and controls examined in this study revealed that the frequencies of allele 12 (D) of DYS388 locus (OR, 3.314; 95% CI, 2.147–5.116; P < 0.01) and allele 14 (F) of DYS439 locus (OR, 8.736; 95% CI, 1.082–70.552; P = 0.015) were significantly higher in patients than healthy control group, while alleles 10 (B) and 13 (E) of DYS388 locus had a significantly lower frequency in patients than healthy controls (OR, 0.152; 95% CI, 0.057–0.400; P < 0.01 for allele 10, and OR, 0.430; 95% CI, 0.221–0.838; P = 0.011 for allele 13). However, the alleles of DYS435 and DYS437 loci showed no significant differences in frequency among patient and control groups studied. Conclusion: It is likely that men who belong to Y-lineages with either allele 12 (D) of DYS388 or allele 14 (F) of DYS439 are more susceptible to develop prostate cancer, while those who belong to lineages with allele 10 (B) or 13 (E) of DYS388 are more resistant to this disease. However, the comparison between results obtained from different ethnic groups examined in this study showed that alleles 10 (B) and 12 (D) of DYS388 are the only alleles associated with a respectively lower and higher risk of prostate cancer in all populations studied.

1084T

The Clinical Phenotype of Endometrial Cancer in Women of Newfoundland and Labrador with a MSH2 Lynch Syndrome Mutation. *A.H.G. Nichols[1], L.M. Dawson[2], J. Green[3], E. Dicks[1], P. Parfrey[1].* 1) Patient Research, Memorial University of Newfoundland, St. John's, NL, Canada; 2) Eastern Health, Health Sciences Center, St. John's, NL, Canada; 3) Clinical Genetics, Memorial University of Newfoundland, St. John's, NL, Canada.

Abstract Endometrial Cancer (EC) is the most common extra-colonic malignancy in Hereditary Nonpolyposis Colorectal Cancer (HNPCC) or Lynch Syndrome (LS), and in most cases is the first or `sentinel' cancer. In comparison to colonic cancers, the clinical phenotype of EC has not been as rigorously investigated and the natural history of the disease not well known. The purpose of this study is to examine clinically important variables in known *MSH2* mismatch repair gene mutation carriers to that of non-mutation carriers in general population with sporadic EC. **Methods** Clinical data was abstracted retrospectively from the medical charts of 46 women with a known *MSH2* mismatch repair mutation determined through genetic sequencing and that have been previously diagnosed with EC. Medical charts were examined for clinically relevant variables of EC and then compared to sporadic endometrial cancers from the general population. **Results** The median age of diagnosis of EC found in the cohort of mutation positive individuals was 45.00 years of age. The stage at diagnosis consisted of 64.1% at stage I and 17.4% of this population had papillary serous cell type. **Conclusions** We conclude that endometrial cancers in HNPCC are being diagnosed at an earlier age and stage of cancer when compared to the general population. This cohort is also presenting with more aggressive cell types resulting in a worse prognosis. Further research regarding the carcinogenic pathway and molecular profile of Lynch Syndrome based endometrial cancers is needed.

1085F

A unique African Prostate Cancer Cohort (AfPCC) to investigate genetic and environmental risk factors of aggressive disease within Africa. *E.A. Tindall[1], M.S.R. Bornman[2], S. van Zyl[3], D.C. Petersen[2], A.M. Segone[3], L.R. Monare[3], P.A. Venter[4], V.M. Hayes[1,4].* 1) J. Craig Venter Institute, San Diego, CA; 2) University of Pretoria, Department of Urology, Steve Biko Academic Hospital, Pretoria, South Africa; 3) University of Limpopo, Department of Urology, Medunsa Campus, South Africa; 4) University of Limpopo, Department of Medical Sciences, Turfloop Campus, South Africa.

A major clinical need exists for early detection of men at risk of aggressive prostate cancer (PCa). Identification of genetic markers of susceptibility provides a means to fulfill this requirement. Although African ancestry is a significant risk factor, PCa association studies are typically European-biased. We present an ongoing collection of non-migrant Southern African Bantu men with and without prostate cancer, namely the African Prostate Cancer Cohort (AfPCC), as a unique resource for assessing genetic and environmental influences of aggressive prostate cancer within Africa. As well as comparing disease phenotypes in the AfPCC to “White” and “Black” men enlisted in publicly available US based registries, we investigate the impact of lifestyle and environmental factors on PCa occurrence in the AfPCC. Furthermore, we investigate the impact of pre-defined PCa risk alleles, achieving genome-wide significance predominantly in European populations, on predicting PCa in the AfPCC. We demonstrate a disproportionate rate of aggressive PCa in the AfPCC compared to “White” and “Black” men within the US. Epidemiological data on the current content of AfPCC (456 cases and 280 controls), indicates a possible association with overall risk and a family history of PCa, the presence of diabetes, aspirin consumption, balding pattern and erectile dysfunction. The association with family history and balding pattern, as well as a reduced age, also translates to an association with aggressive disease phenotypes. Genetic analysis of an initial subset of the AfPCC (297 cases and 206 controls) revealed no significant association with previously reported PCa risk alleles and genetic risk profiles failed to accurately predict or enhance the predictive power of serum PSA levels in the AfPCC. Despite the assertion that African ancestry contributes to an increased risk of the disease, there is very little information pertaining to the impact of prostate cancer within non-migrant African populations. The AfPCC provides one of the largest and most comprehensive resources for investigating genetic and environmental influences on PCa associated with an African heritage. The limited availability of PSA testing in Southern Africa and the subsequent presentation of an extremely aggressive disease phenotype in the AfPCC, contributes to this study being a vital resource for much-needed research into risk factors of clinically relevant PCa.

1086W

Insulin-like growth factor biomarkers and genetic polymorphisms are associated with adenomatous polyp risk. *C.B. Vaughn[1], H. Ochs-Balcom[1], J. Nie[1], Z. Chen[2], C.L. Thompson[2], L. LI[2].* 1) Social and Preventive Medicine, University at Buffalo, Buffalo, NY; 2) Family Medicine and Community Health, Case Western Reserve University, Cleveland, OH.

Obesity is an established risk factor for colon cancer. Colorectal adenomas are precursors to colorectal cancer that have been previously linked to insulin resistance. We hypothesize that inherited genetic variation the insulin-like growth factor (IGF) axis may mediate the association of insulin resistance and risk of adenomatous polyps. In this study we examined the association of circulating levels of IGF-1 and its main binding protein, IGFBP-3, as well as single nucleotide polymorphisms (SNPs) in the IGF1R and IGF2R genes with risk of adenomatous polyps in the Case Transdisciplinary Research in Cancer and Energetics (TREC) Colon Polyps Study. Our study includes 1480 patients who underwent routine colonoscopy. At colonoscopy, 410 incident and pathologically confirmed adenoma cases were identified, as were 1070 polyp-free controls. Levels of the circulating biomarkers were categorized into tertiles based on the distribution in the controls. Logistic regression modeling was used to adjust for age, sex, body mass index, smoking status, NSAID use and family history of colorectal cancer to test for associations with circulating IGF-1 and IGFBP3, as well as SNPs in candidate genes. Among cases, Caucasians had statistically significant higher IGF-1 and IGFBP3 compared to African Americans. In logistic regression models stratified by race, we identified higher odds of adenomatous polyps with high circulating IGF-1 in African Americans only; the adjusted odds ratios for tertiles 2 and 3 compared to the lowest tertile were 1.68, 95% CI: 1.06–2.68 and 1.68, 95% CI: 1.05–2.71, respectively. We observed no associations of IGFBP-3 levels and odds of adenomas. Of ten tag-SNPs genotyped in the IGF1R gene, one SNP, rs4966011, was found to be associated with reduced odds of polyps in log additive genetic models adjusted for covariates and this was statistically significant in Caucasians only. The per (minor) allele OR adjusted for covariates is 0.73, 95% CI: 0.47–0.93. Similarly, one of six SNPs in the IGF2R gene was statistically significant in Caucasians only, rs3777404, with an adjusted per allele OR of 1.53, 95% CI: 1.10–2.14. Our study suggests that there may be race-specific differences in the association on the IGF pathway and risk of adenomatous polyps, however further studies are needed to replicate our findings.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1087T

An Unusual BRCA Mutation Distribution in a High Risk Cancer Genetics Clinic. *A.C. Nelson-Moseke[1], J.M. Jeter[2,3], H. Cui[3], D. Roe[3,4], S.K. Chambers[1,3], C.M. Laukaitis[2,3].* 1) Obst.-Gynecology, University of Arizona, AZ, 85724; 2) Dept. of Medicine, University of Arizona, AZ; 3) University of Arizona Cancer Center, Tucson, AZ; 4) College of Public Health, Univ. of Arizona, Tucson, AZ.

Purpose: The Database of Individuals at High Risk for Breast, Ovarian, or Other Hereditary Cancers at the Arizona Cancer Center in Tucson, Arizona, analyzes cancer risk and outcomes in patients with a family history of cancer or a known genetic mutation. We analyzed the subset of clinic patients who carry deleterious BRCA gene mutations to understand why mutations in BRCA2 outnumber those in BRCA1. Methods: Medical, family, social, ethnicity and genetic mutation history were collected from consenting patients' electronic medical records. Differences between BRCA1 and BRCA2 patients from this database were analyzed for statistical significance and compared to published analyses. Results: A significantly higher proportion of our clinic patients carry mutations in BRCA2 than BRCA1, compared with previous reports of mutation prevalence. This is similar in the Hispanic subgroup. Patients with BRCA2 mutations were significantly more likely than their BRCA1 counterparts to present to the high risk clinic without a diagnosis of cancer. Other differences between the groups were not significant. Six previously unreported BRCA2 mutations appear in our clinic population. Conclusion: The increased proportion of patients carrying deleterious BRCA2 mutations is likely multifactorial. It may reflect the family history of patients presenting for care and Tucson's unique ethnic heritage.

1088F

Functional Characterization of the Pancreatic Cancer TERT-CLPTM1L risk locus on Chr5p15.33. *J. Jia[1], A. Thompson[1], A. Bosley[2], H. Parikh[1], Z. Wang[3], I. Collins[1], G. Petersen[4], C. Westlake[5], T. Andresson[2], L. Amundadottir[1].* 1) Laboratory of Translational Genomics, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD 20877, USA; 2) Laboratory of Proteomics and Analytical Technologies, SAIC-Frederick, Inc, Frederick National Laboratory for Cancer Research, Frederick, MD, United States; 3) Frederick National Laboratory for Cancer Research, Advanced Technology Program, SAIC-Frederick Inc., NCI-Frederick, Frederick, MD, USA; 4) Department of Health Sciences Research, Mayo Clinic, Rochester, MN 55905, USA; 5) Laboratory of cell and Developmental Signaling, NCI-Frederick, Frederick, MD, USA.

Pancreatic cancer is a highly lethal cancer with few well established risk factors. PanScan, an ongoing genome wide association study (GWAS) of pancreatic cancer has identified four pancreatic cancer susceptibility loci in populations of European ancestry. One of these is in a multi cancer locus on chr5p15.33 where the most significant SNP from the GWAS (rs401681, P=3.7×10-7, ORAllele=1.19) that contains two genes, TERT and CLPTM1L. The TERT gene encodes the catalytic subunit of telomerase, well known for its essential role in maintaining telomere ends. The function of CLPTM1L is not as clear, although it has been proposed to play a role in apoptosis. It is predicted to encode a protein with 6 transmembrane (TM) domains and two large hydrophilic domains: a loop of 253 aa between the first and second TM domains, and a C-terminal tail of 89 aa. We have performed imputation using the 1000G and DCEG reference datasets and thereby fine-mapped the signal to a SNP three orders of magnitude more significant than the GWAS SNP (PImputed=1.4×10-10, ORAllele=1.30). As this SNP is located in the CLPTM1L gene we have performed a series of experiments to investigate the function of the CLPTM1L gene and its encoded protein. Immunofluorescence analysis in pancreatic cancer (PANC-1) and mouse kidney (IMCD-3) cells indicates that it localizes to the endoplasmic reticulum. Affinity purification and mass spectrometry (HEK-293T, hTERT-HPNE and PANC-1 cells) identified potential interacting proteins involved in intracellular transport, secretion and cytokinesis that are currently being validated by co-immunoprecipitation and co-localization. To examine if the protein product of this gene plays a role in growth control, we created stable PANC-1 cell lines overexpressing the full length CLPTM1L gene as well as two deletions, a C-terminal deletion (89 aa) and loop deletion (253 aa) and assayed growth in vitro and in vivo. Cell lines overexpressing full length CLPTM1L grow faster in vitro and in vivo as compared to cells containing empty vector. Interestingly, the two CLPTM1L mutants abolish this effect. Our results indicate that CLPTM1L may play a role in the control of cell growth and oncogenesis in the pancreas. Our current efforts is to characterize the function of the CLPTM1L gene and protein in greater detail and correlate pancreatic cancer risk variants on 5p15.33 to molecular phenotypes to attempt to explain the underlying biology of the risk.

1089W

A comprehensive approach to evaluating the impact of sequence variants of uncertain significance in 24 breast cancer genes on transcript splicing. *S. Casadei, T. Walsh, C.H. Spurrell, A.M. Thornton, J.B. Mandell, S.M. Stray, M.K. Lee, M.C. King.* Medical Genetics & Genome Sciences, University of Washington, Seattle, WA.

In addition to altering amino acids in the encoded proteins, sequence variation in exons and introns may affect mRNA splicing by altering splice sites, splicing regulatory elements (SREs), or mRNA secondary structure. Of loss-of-function mutations, perhaps 15% impact canonical splice sites and up to 50% impact splicing when additional splicing regulatory elements in exons and introns are considered. For familial breast cancer patients with no conventional mutations in BRCA1 or BRCA2, we tested possible effects on splicing of all variants detected by complete genomic sequence of all known breast cancer genes. We integrated in silico predictions of aberrant splicing with experimental studies of gene transcripts from blood-based RNA and splicing reporter minigenes. We have comprehensively evaluated all sequence variants in exons and introns of 24 breast cancer genes included in our targeted resequencing panel. We identified 11 sequence variants in BRCA1, 25 in BRCA2, 18 in PALB2 and 3 in BARD1 that were predicted by in silico tools to alter splicing. These 57 variants were evaluated experimentally by RT-PCR analysis of blood RNA. Aberrant splicing was confirmed for 18 mutations: 5 in BRCA1, 10 in BRCA2, 2 in PALB2 and 1 in BARD1. Consequences of the mutations at the transcript level were partial exonic deletion, exon skipping, intron inclusion and complex splices. For the remaining 39 sequence variants with negative RT-PCR analysis, we are generating splicing reporter minigenes to further test the predicted effect on splicing and to identify exonic and intronic elements that enhance or repress splicing. Rapid advances in high throughput DNA sequencing methods are leading to identification of more disease-related genes and a wider spectrum of loss-of-function mutations, but also to more variants of unknown functional and clinical significance than previously encountered. Bioinformatics tools can help predict the consequences of mutations potentially altering splicing and transcript stability and prioritize experimental studies of mRNA splicing. In our view, an integrated approach combining the use of in silico prediction tools with experimental studies on blood RNA and splicing reporter minigenes will maximize our ability to detect aberrant splicing events.

1090T

Functional analysis of germline *TP53* mutations in lymphocytes from Li-Fraumeni patients reveals a drastic biological impact of missense mutations. *J-M. Flaman[1], Y. Zerdoumi[1], C. Durambure[2], J. Laury-Andas[1], G. Bougeard-Denoyelle[1], T. Frebourg[1], The French LFS working group.* 1) Inserm, U1079, University of Rouen, Institute for Research and Innovation in Biomedicine (IRIB), Rouen, Normandy, France; 2) Inserm, U905, University of Rouen, IRIB, Rouen, Normandy, France.

The Li-Fraumeni syndrome (LFS, MIM 151623), resulting from *TP53* mutations, is a Mendelian predisposition to cancer affecting children and adults and characterized by a remarkably wide tumour spectrum including soft-tissue sarcomas, osteosarcomas, premenopausal breast cancers, brain tumors, adrenocortical tumors, plexus choroid tumors, leukemias and lung cancers. As demonstrated by the functional assays of *TP53* cDNA that we developed in yeast, all germline *TP53* mutations result into a decrease of p53 transcriptional activity. We reported that germline *TP53* missense mutations are associated with an earlier age of tumor onset, suggesting a more deleterious effect of these mutations, as compared to the other types of mutations. In order to assess the biological impact of the heterozygous *TP53* mutations in the genetic context of patients, we developed a new functional assay of the p53 pathway in EBV-immortalized lymphocytes derived from controls or LFS patients. This assay is based on the activation of the p53 pathway by exposure of the lymphocytes to doxorubicin, a DNA damaging agent, and then on global transcriptome analysis of the p53 target gene expression. In wild-type *TP53* EBV-immortalized lymphocytes, we identified almost 400 genes whose expression was induced more than two fold following DNA damage exposure. Fifty percent of the genes were already known to be p53 target genes, which validates the assay. We then perform a global transcriptomic analysis of lymphocytes carrying different germline *TP53* mutations to compare the effect of missense mutations to that of the other types of mutations. These analyses revealed that the number of induced genes was dramatically reduced selectively in lymphocytes with missense mutations, demonstrating the remarkable deleterious impact of *TP53* missense mutations. Measurement of known p53 target gene induction, using RT-QMPSF, showed that the defect of induction associated with the missense mutations was above 50% supporting that, in lymphocytes, the corresponding mutant proteins have a transdominant activity over the wild-type. These result explain why germline *TP53* mutations, like somatic mutations, are predominantly missense mutations and why these mutations are usually associated with an earlier age of tumor onset.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1091F

Investigation of the bladder cancer association within TP63 gene region. *Y.-P. Fu[1], I. Kohaar[1], J. Earl[2], W. Tang[1], J.D. Figueroa[3], N. Malats[4], M. Garcia-Closas[3,5], N. Chatterjee[3], M. Kogevinas[6-9], P. Porter-Gill[1], D. Baris[3], D. Albanes[3], M.P. Purdue[3], A. Carrato[10], A. Tardón[8,11], C. Serra[12], R. García-Closas[13], J. Lloreta[14], A. Johnson[15], M. Schwenn[16], M.R. Karagas[17], A. Schned[17], W.R. Diver[18], S.M. Gapstur[18], M. Thun[18], J. Virtamo[19], D.T. Silverman[3], N. Rothman[3], F.X. Real[2,12], L. Prokunina-Olsson[1]*. 1) Laboratory of Translational Genomics, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892, USA; 2) Epithelial Carcinogenesis Group, Molecular Pathology Programme, Centro Nacional de Investigaciones Oncológicas (CNIO), Madrid 28029, Spain; 3) Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892, USA; 4) Genetic and Molecular Epidemiology Group, Spanish National Cancer Research Center, Madrid 28029, Spain; 5) Division of Genetics and Epidemiology, The Institute of Cancer Research, London SW7 3RP, UK; 6) Centre for Research in Environmental Epidemiology (CREAL), Barcelona 08003, Spain; 7) Municipal Institute of Medical Research, Barcelona 08003, Spain; 8) CIBER Epidemiología y Salud Pública (CIBERESP), Barcelona; 9) National School of Public Health, Athens 11521, Greece; 10) Ramón y Cajal University Hospital, Madrid 28034, Spain; 11) Instituto Universitario de Oncología, Universidad de Oviedo, Oviedo 33003, Spain; 12) Departament de Ciències Experimentals i de la Salut, Universitat Pompeu Fabra, Barcelona 08003, Spain; 13) Unidad de Investigación, Hospital Universitario de Canarias, La Laguna 38320, Spain; 14) Hospital del Mar-Institut Municipal d'Investigació Mèdica (IMIM), Universitat Pompeu Fabra, Barcelona 08003, Spain; 15) Vermont Cancer Registry, Burlington, Vermont 05401, USA; 16) Maine Cancer Registry, Augusta, Maine 04333, USA; 17) Dartmouth Medical School, Hanover, New Hampshire 03755, USA; 18) Epidemiology Research Program, American Cancer Society, Atlanta, Georgia 30303, USA; 19) National Institute for Health and Welfare, Helsinki 00271, Finland.

BACKGROUND: Genome-wide association studies (GWAS) have identified a single nucleotide polymorphism (SNP), rs710521 near the tumor protein p63 (TP63) gene, as a risk factor of bladder cancer. Here we explored this region to map genetic variants contributing to bladder cancer susceptibility, and to investigate their molecular phenotypes and potential roles in cancer biology. METHODS: Based on 1000 Genomes and HapMap 3 project data, we imputed 1715 SNPs within +/− 100kb of TP63 region (GRCh37, chr3: 189,249,216-189,715,068) among 3,532 bladder cancer cases and 5,120 healthy controls of European ancestry. Alternative splicing of TP63 was evaluated by RNA-Seq in 6 pairs of tumor-normal bladder tissue samples. TP63 mRNA expression was measured with custom-designed TaqMan assays in 40 bladder tumor and 40 adjacent normal tissue samples, also genotyped by Illumina Omni1-Quad BeadChip and then imputed using the same reference panels. Logistic regression and generalized linear models were used to test for the association between SNPs and bladder cancer risk, and between SNPs and TP63 mRNA expression levels, respectively. All models were assumed the log-additive SNP effects and adjusted for study sites, age, gender and smoking habit when applicable. RESULTS: A total of 813 well-imputed and 126 GWAS-genotyped SNPs were used for analyses. Based on our RNA-Seq results, we measured expression of 5 main splicing forms of TP63 in 40 normal and 40 tumor bladder tissue samples. An imputed SNP rs28512356, located 400 bp downstream from TP63, showed the strongest association with mRNA expression of TP63-alpha splicing form in bladder tumors (p=0.018) but not in normal bladder tissue. TP63-alpha encodes a protein with the Sterile Alpha Motif (SAM) domain which regulates activity of TP53 gene. The original GWAS signal rs710521 had a per-allele odds ratio (OR) of 1.16 (95%CI=1.08-1.25, p= 4.6E-05) for bladder cancer risk, which is comparable to rs28512356, with a per-allele OR of 1.19 (95%CI=1.09–1.29, p=6.1E-05; r2=0.628, D'=0.998 with rs710521). CONCLUSION: We identified SNP rs28512356 as a possible variant for bladder cancer risk within the TP63 region, and associated with mRNA expression of the TP63-alpha splicing form in bladder tumors. Future studies are warranted to validate our findings in additional samples and investigate the possible functional mechanisms of this genetic association in bladder cancer.

1092W

Genetic variant as a selection marker for anti-PSCA immunotherapy of bladder cancer. *I. Kohaar[1], P. Porter-Gill[1], P. Lenz[2], Y. P. Fu[1], A. Mumy[1], W. Tang[1], A. B. Apolo[3], N. Rothman[4], D. Baris[4], A. R. Schned[5], K. Ylaya[6], M. Schwenn[7], A. Johnson[8], M. Jones[9], M. Kida[10], D. T. Silverman[4], S. M. Hewitt[6], L. E. Moore[4], L. Prokunina-Olsson[1]*. 1) Laboratory of Translational Genomics, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, 8717 Grovemont Circle, Bethesda, MD, 20892-4605, USA; 2) Clinical Monitoring Research Program, National Laboratory for Cancer Research, SAIC-Frederick, Maryland, USA; 3) Medical Oncology Branch, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, USA; 4) Occupational and Environmental Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, USA; 5) Department of Pathology, Dartmouth Medical School, Hanover, New Hampshire USA; 6) Laboratory of Pathology, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, USA; 7) Maine Cancer Registry, Augusta, Maine, USA; 8) Vermont Cancer Registry, Burlington, Vermont, USA; 9) Department of Pathology and Laboratory Medicine, Maine Medical Center, Portland, Maine, USA; 10) Department of Pathology, University of Vermont College of Medicine, Burlington, Vermont, USA.

A single nucleotide polymorphism (SNP) rs2294008 within prostate stem cell antigen (*PSCA*) gene on chromosome 8q24.3 has been identified by a genome-wide association study (GWAS) as a risk factor for bladder cancer [OR=1.13 (1.09–1.17), p=4E-11, Rothman *et al*, Nat Gen, 2010]. Previously, we found higher *PSCA* mRNA expression in bladder tumors compared to adjacent normal tissue and in both tissue types the expression was specifically increased in carriers of the risk T allele of rs2294008 [Fu *et al*. PNAS, 2012]. Here we explored mRNA expression of *PSCA* in relation to rs2294008 and bladder cancer risk. Analysis of mRNA sequencing (RNA-seq) data from 6 normal and 6 tumor bladder tissue samples heterozygous for rs2294008 showed evidence of strong allelic expression imbalance (AEI), which is reflected in deviation from the 50:50% allelic ratio expected in heterozygous transcribed variants. The average allelic ratio evaluated by counts of RNA-seq reads was 90:10 % for rs2294008 T:C, instead of the expected 50:50 % ratio. A similar pattern was observed for 11 transcribed *PSCA* SNPs, all of which were in high linkage disequilibrium with rs2294008, but not for SNPs located in neighboring genes, *Ly6K* and *JRK*. These results suggest that the AEI is specific for PSCA and not a result of AEI of the whole region. RNA-seq results were validated by allele-specific expression analysis in bladder tissue samples. We further showed that rs2294008 strongly predicted PSCA protein expression in bladder tumors examined by immunohistochemistry (p=6.46E-11, n=278). The pattern of rs2294008-associated protein expression was similar in non-muscle invasive tumors - stages Ta (p= 3.10E-05, n=173) and T1 (p=2.64E-05, n=60), and muscle-invasive tumors - stages T2 (p=1.02E-02, n=23) and T3/4 (p=2.64E-02, n=22). Monoclonal anti-PSCA antibody has emerged as a novel cancer therapy, currently being tested in clinical trials for prostate and pancreatic cancer. Thus, we suggest that genetic testing for rs2294008 could be used to select bladder cancer patients with high PSCA expression and therefore, suitable for anti-PSCA immunotherapy. Up to 75% of bladder cancer patients of European ancestry carry the rs2294008 risk TC and TT genotypes associated with high PSCA expression, hence these patients might benefit from this treatment.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1093T

Tumor microenvironment and genetic association with colorectal cancer risk. *V. Peltekova[1], M. Lemire[1], Q. Trinh[1], A. Qazi[1], R. Bielecki[2], L. Hodgson-Jensen[1], D. D'Souza[1], S. Zandi[3], T. Chong[1], R. De Borja[1], L. Timms[1], J. Rangrej[1], M. Volar[1], M. Chan-Seng-Yue[1], T. Beck[1], J. Kwan[1], K. Kozak[1], C. Ash[2], L.D. Stein[1,4], J.E. Dick[3,4], J.D. McPherson[1,4], B.W. Zanke[5], A. Pollett[2], S. Gallinger[2,4], T.H. Hudson[1,4].* 1) Ontario Institute for Cancer Research, Toronto, Ontario, Canada; 2) Samuel Lunenfeld Research Institute, Toronto, Canada; 3) Campbell Family Institute for Cancer Research, Ontario Cancer Institute, University Health Network, Toronto, Canada; 4) University of Toronto; 5) Ottawa Hospital Research Institute, Ottawa, Canada.

We used microarray-based target selection methods coupled with next generation sequencers to interrogate 2.3 Mb of DNA at 11 colorectal cancer loci. The dataset refined the signal of association at four loci: rs16892766/8q23.3, rs6983267/8q24, rs3802842/11q23, and rs9929218/16q22. The associated 11q23 region includes three uncharacterized protein-coding genes (C11orf53, C11orf92, and C11orf93) and POU2AF1. We characterized the gene expression levels of the four transcripts in colon tissues. Risk alleles correlate with decreased gene expression of C11orf92 and C11orf93 in benign adjacent colonic tissues as well as in tumors. C11orf92 and C11orf93 are arranged head-to-head on opposite strands, share a putative bidirectional promoter and may use common regulatory elements. Western blot analyses of C11orf92 in benign adjacent tissues from eight individuals show stronger C11orf92 in the lamina propria of all biopsies, but not in normal epithelium or epithelium-derived tumor cells. C11orf92 expression appeared cytoplasmic and often seen as part of granular structures that are both intra- and extra-cellular. Triple immunofluorescence on colon tissue cryosections showed co-localization of C11orf92 with cell-specific markers for eosinophils, mast cells, neutrophils, macrophages and dendritic cells. Four lines of evidence link C11orf92 and C11orf93 with colorectal cancer: 1) genetic association at 11q23; 2) decreased RNA (and likely protein) expression for both genes correlating with risk alleles; 3) C11orf92 protein expression in many mucosal immune cells of the gut implicated in tumor immunity; and 4) physical proximity between C11orf92-containing immune cells and/or extracellular granules with colon cancer cells. It will be interesting to further understand the mechanistic link between this risk locus, microenvironmental factors and human cancer predisposition.

1094F

Evaluation of the functional effects on cell signaling caused by allelic forms of prostate stem cell antigen (PSCA). *P. Porter-Gill, A. Mumy, W. Tang, I. Kohaar, L. Prokunina-Olsson.* Laboratory of Translational Genomics/Division of Cancer Epidemiology and Genetics, NCI/NIH, Gaithersburg, MD.

Genome-wide association studies (GWAS) have identified a single nucleotide polymorphism (SNP), rs2294008, within the prostate stem cell antigen (PSCA) gene at 8q24.3 as a risk factor for bladder cancer. A PSCA transcript with the non-risk C allele of rs2294008 encodes a protein of 114 amino acids (aa), including a leader peptide of 11 aa. The risk T allele creates a novel upstream translation start site (aCg->aTg) that extends the N-terminal leader peptide by 9 aa, to generate a protein of 123 aa, with a leader peptide of 20 aa. Previously, we have shown that the risk T allele is associated with increased PSCA mRNA and protein expression in bladder tumors. To further evaluate the functional role of the two PSCA protein isoforms, we examined their effects on 45 signaling pathways (Cignal 45-Pathway Reporter Array, Qiagen) in two different cell lines (HeLa and HTB-5). Thirteen of 45 pathways showed differential activation by PSCA allelic constructs. Each of these pathways will be individually tested in more detail. In addition, we are examining the effect of the PSCA protein isoforms on global transcriptome of HeLa and HTB-5 cell lines by RNA-sequencing of cells transfected with allelic constructs and performing cell proliferation and viability assays. Evaluation of the functional properties of the two PSCA protein isoforms should improve understanding of the molecular mechanism of bladder cancer.

1095W

Discovery and Functional Characterization of Somatic Mutations in the Novel Tumor Suppressor DEAR1. *J. Reuther[1,2], N. Chen[1], S. Balasenthil[1], A. Killary[1,2].* 1) Genetics, MD Anderson Cancer Center, Houston, TX; 2) Human and Molecular Genetics, University of Texas Graduate School of Biomedical Science, Houston Texas.

Ductal Epithelium Associated Ring Chromosome 1 (DEAR1) is a novel tumor suppressor gene that is both mutated and homozygously deleted in breast cancer. Loss of expression of DEAR1 by immunohistochemistry correlated with local recurrence in 123 cases of early onset breast cancer. Genetic complementation of wild type DEAR1 into a breast cancer cell line 21MT which contained a missense mutation in codon 187 restored acinar morphogenesis in three dimensional culture, indicating that DEAR1 is a novel, regulator of acinar morphogenesis. Furthermore, DEAR1 expression is downregulated or lost at the earliest stage of pre-invasive breast cancer, Ductal Carcinoma *In Situ*. DEAR1's relevance as an underlying “driver' event in the generation and/or progression of breast cancer is underlined in DEAR1's involvement in restoration of acinar morphogenesis, its loss at the early stages of breast cancer, and its prediction of local recurrence free survival. Identification and functional characterization of novel somatic mutations in DEAR1 in epithelial cancers is a major goal in our lab. Currently, we have identified 16 different non-synonomous mutations in DEAR1 in a variety of epithelial cancers including breast, pancreatic and colon cancer. Mutation constructs have been generated for a number of these mutations and have been tested for their functional significance by genetic complementation in various breast cancer cell lines. We have evidence of additional functional mutations that abrogate DEAR1's functions. The long term goal of these studies is to determine a mutation signature that could be used to predict prognosis in breast and other epithelial cancers in which DEAR1 is mutated or allow for stratification of patients for new targeted therapies aimed at the pathways regulated by DEAR1.

1096T

rs28381943 and rs2032586 SNPs of ABCB1 gene may be the reason of mRNA stabilization which may lead to gene overexpression. *M.S Shahbazi[1], M.M Mansoori[1, 2], M.G Golalipour[3], Sh.A Alizadeh[4], A.J Jahangirerad[4], A.T Tahmasebifar[2], S.R.K Khanduzi[4].* 1) biochemistry & Genetics, Medical Faculty, Medical Cellular & Molecular Research Center, Talghani Children Hospital of Golestan University of Medical Sciences, Bolv Janbazan, 4916668197, Gorgan, Iran; 2) Medical Cellular & Molecular Research Center, Talghani Children Hospital of Golestan University of Medical Sciences, Bolv Janbazan, 4916668197, Gorgan, Iran; 3) Golestan university, Medical Cellular & Molecular Research Center, Talghani Children Hospital of Golestan University of Medical Sciences, Bolv Janbazan, 4916668197, Gorgan, Iran; 4) Oncology Department, 5th Azar hospital of Golestan University of Medical Sciences, Gorgan, Iran.

Introduction. One of the major mechanisms for drug resistance is associated with altered anticancer drug transport, mediated by human-adenosine triphosphate binding cassette (ABC) transporter superfamily proteins. The overexpression of MDR1 by multidrug-resistant cancer cells is a serious impediment to chemotherapy. In our study we have studied the possibility that the mechanism of MDR1 overexpression is caused by structural SNPs. Materials and Methods. A total of 101 MDR1 cases and 100 controls were genotyped with SSP-PCR (Sequence Specific Primer). Gene expression was evaluated for 70 MDR1 cases and 54 controls by Real Time PCR. The correlation between the two groups was based on secondary structures of RNA were predicted by bioinformatics tool (Rn) (Vienna RNA Fold Server). Results. The results of genotyping showed that among 3 studied SNPs, rs28381943 and rs2032586 have significant difference between MDR1 cases and control group but there were no differences in the two groups for C3435T. The results of Real Time PCR showed overexpression of ABCB1 when we compared our data with each of the genotypes in average mode. Prediction of secondary structures in the existence of 2 related SNPs (rs28381943 and rs2032586) showed that the amount of ΔG for original mRNA is higher than the amount of ΔG for the two mentioned SNPs. Discussion. We have observed that 2 of our studied SNPs (rs283821943 and rs2032586) may affect the overexpression of MDR1 gene, and the observed overexpression may be due to mRNA stability while this was not the case for C3435T.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1097F

Identification and Functional Characterization of a Novel Splicing Form of the TPCN2 Gene on 11q13.3 Within a Region Associated with Prostate Cancer Risk. *N. Sikdar[1], C. Chung[1], J. Fang[1], M. Tarway[1], W. Tang[1], S. Mukherjee Dey[2], S. Ambs[3], L. Prokunina-Olsson[1], S. Chanock[1].* 1) Laboratory of Translational Genomics, NCI, Gaithersburg, MD, USA; 2) Medical oncology branch, Center for cancer Research, NCI, Bethesda, MD, USA; 3) Laboratory of Human Carcinogenesis, center for Cancer research, NCI, Bethesda, MD, USA.

Prostate cancer (PrCa) is the most frequently diagnosed cancer in males in developed countries. Genome-wide association studies (GWAS) have identified over 45 loci, including 3 independent PrCa susceptibility loci on chr 11q13.3. By fine-mapping analysis, we identified, rs10896438, as one of the informative markers on 11q13.3. This marker resides 20 kB from EST DA379985 and 92kB telomeric of the TPCN2 gene. We observed DA379985 expression in normal and tumor prostate as well as normal bladder, breast, cervix, ovary and testis tissue samples. In 71 primary prostate tissue samples, DA379985 expression was significantly lower in PrCa tissue compared to normal prostate tissue (p=0.034), but we did not observe association between rs10896438 and mRNA expression. Immunohistochemical data also suggest TPCN2.1 expression was lower in prostate cancer tissue samples with respect to normal prostate tissue section. Rapid amplification of cDNA ends (RACE) and RNA-seq data showed that this EST is a part of a novel splicing form of TPCN2 (which we refer to as TPCN2.1). TPCN2 encodes a 752 aa protein and the new splice variant, TPCN2.1, encodes a protein of 520 aa. Unlike TRPM8 endoplasmic reticulum Ca2+ channel, TPCN2.1 expression is independent of androgen regulation. Confocal imaging of cells transiently transfected with expression constructs for TPCN2 and TPCN2.1 showed cytoplasmic lysosomal expression for both proteins. Our recent findings suggest that siRNA silencing of TPCN2.1 leads to autophagy induced cell death, but not apoptosis, in PC3 cell line. A significant percentage of PC3 and TsuPro1 cells were accumulated in G2/M phase of cell cycle after siRNA silencing of TPCN2.1. siRNA-mediated downregulation of TPCN2.1 resulted in decreased proliferation of prostate cancer cells in vitro. These results indicate that TPCN2.1 is a novel spliced variant of voltage-gated TPCN2 Ca2+ channel expressed in prostate tissue. TPCN2.1 is an important determinator of Ca2+ homeostasis in prostate epithelial cells and may be a potential target for the action of drugs in the management of prostate cancer.

1098W

An integrated analysis of genomic variation and transcriptome sequencing to identify novel susceptibility genes for breast cancer. *E. Wagner[1], S. Clare[2], M. Radovich[2], Y. Liu[3], B. Schneider[4], A.M. Storniolo[4], C. He[1].* 1) Dept. of Public Health, Indiana University School of Medicine, Indianapolis, IN; 2) Dept. of Surgery, Indiana University School of Medicine, Indianapolis, IN; 3) Dept. of Medical & Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN; 4) Dept. of Medicine, Indiana University School of Medicine, Indianapolis, IN.

Recent genome-wide association studies (GWAS) have identified more than 20 novel loci that were associated with breast cancer. However, the biological functions of these genetic variants in the development of breast cancer have not been characterized. We hypothesize that the associated genetic variants influence tissue-specific gene expression that in turn leads to breast cancer. We performed an integrated analysis of genomic variation and transcriptome sequencing to elucidate expression quantitative trait loci (eQTL) and to identify breast cancer susceptibility genes whose expression were influenced by the candidate genetic variants. The current study utilized a unique source of paired blood and normal breast tissue from 20 healthy women free of diagnosed breast cancer. We genotyped 38 SNPs that were associated with breast cancer in previous GWAS reports, and assessed expression of more than 24,000 gene transcripts in normal breast epithelium using transcriptome sequencing data. Analysis of variance (ANOVA) was performed to assess the differences of gene expression across genotypes. After account of age, menstrual cycle days and contraceptive use, we found 301 unique genes whose expression in normal breast epithelium were correlated with the 38 breast cancer candidate SNPs (all P<0.001). The majority of the identified genes (>90%) were associated with the candidate SNPs in *trans*, i.e. gene expression were influenced by SNPs on a different chromosome. We have identified known genes (for example, *PARK7*, *LGALS1*, *MPDZ*, *CCL14*, *COL1A1*, *CTCF*) that showed differential expression in breast tumor vs. adjacent normal tissue in published microarray analyses. In addition, we have identified novel putative breast cancer genes including *LST1*, *GLIS2*, *BAG6*, *CELA2A*, and *TLX1*. An ingenuity pathway analysis of the associated genes suggested four functional networks. Network 1, related to “Cell Cycle”, includes genes *NEFL*, *TCEB3*, *SEPT9*, *PLG*, and *EPOR*; network 2, related to “Cell Death”, included genes *GSK3B*, *FAIM*, *GRIN1*, and *WNT1*; network 3, related to “Cellular Growth and Proliferation”, includes genes *CSF3*, *HTR4*, *POU3F2*, and *RAG2*; and network 4, related to “Cell to Cell Signaling and Interaction”, includes genes *LOX*, *NPPB*, *CNTNAP1*, *CNGA3*, and *NAT8L*. Our data identify known and novel putative breast cancer susceptibility genes whose expression were influenced by breast cancer candidate SNPs from previous GWAS, showing the impact of germline variants on the transcriptome.

1099T

Effects of breast cancer associated *PALB2* c.1592delT founder mutation at the cellular level. *R. Winqvist[1], J. Nikkilä[1], K. Pylkäs[1], H. Peltoketo[1], B. Xia[2], H. Pospiech[3, 4].* 1) Lab. of Cancer Genetics and Tumor Biology, Dept. of Clinical Genetics and Biocenter Oulu, University of Oulu, Oulu University Hospital, Oulu, Finland; 2) Dept. of Radiation Oncology, The Cancer Institute of New Jersey, University of Medicine and Dentistry of New Jersey, Robert Wood Johnson Medical School, New Brunswick, NJ, USA; 3) Leibniz Institute for Age Research - Fritz Lipman Institute, Jena, Germany; 4) Dept. of Biochemistry, University of Oulu, Oulu, Finland.

Besides *BRCA1* and *BRCA2* also constitutional defects in their interaction partner *PALB2* have proven to be important for hereditary susceptibility to breast cancer (BC). Disruption of BRCA1-PALB2-BRCA2 interactions results in defective homologous recombination (HR) and increases cellular sensitivity to DNA damaging agents. BRCA2 and PALB2 are required for G2 checkpoint maintenance and that may contribute to their tumor suppressor functions. Heterozygous *PALB2* mutation carriers have a ~2–6 fold increased risk for BC. Biallelic truncating mutations in *PALB2* again sensitize cells to mitomycin-C (MMC) treatment, which is a hallmark of Fanconi anemia. In addition, we have recently shown that mice lacking proper Palb2 function are embryonically lethal and display defective mesoderm differentiation. We have previously identified a heterozygous *PALB2* c.1592delT founder mutation associated with a 40% risk of BC by the age of 70. It results in a truncated protein with decreased BRCA2-binding affinity and that fails to support HR in PALB2-knockdown cells or to restore crosslink repair in a cellular PALB2-deficient model system. Therefore we investigated how a heterozygous *PALB2* defect influences on checkpoint functions, MMC and PARP inhibitor sensitivity, and genomic stability maintenance before and after treatment with DNA damaging agents. For that purpose lymphoblastoid cell lines were established from eight mutation carriers and six healthy controls. The mutation cell lines did not show any difference in IR-induced apoptosis or γH2AX phosphorylation states in comparison to the control lines. In cell survival assays the sensitivity of our heterozygous *PALB2* mutation cell lines to MMC or to PARP inhibitors did not differ from that of the controls. More importantly, after DNA damage *PALB2* carrier cell lines displayed a decrease in G2/M-phase cell ratio, suggesting that mutation carriers have a defective checkpoint, in spite of one copy of intact *PALB2*, and this was confirmed with histone H3 staining. Together with the current functional analysis, we will also work on modeling the influence of defective PALB2 function in mammary tissue development and tumorigenesis.

1100F

Global Correlation of Expression with Prostate Cancer Associated Variants Identifies IRX4 as a Prostate Cancer Expression Trait Locus. X. Xu[1], W.M. Hussain[2], N. Kitabayashi[2], M.A. Rubin[2], F. Demichelis[3,4], R.J. Klein[1]. 1) Department of Cancer Biol & Genetics, Memorial Sloan-Kettering Cancer Center, New York, NY; 2) Department of Pathology and Laboratory Medicine, Weill Cornell Medical College, New York, NY; 3) Institute for Computational Biomedicine, Weill Cornell Medical College, New York, NY; 4) Centre for Integrative Biology, CIBIO, University of Trento, Italy.

Background: Functional follow-up to prostate cancer genome-wide association studies (GWAS) have been sorely lacking. One mechanism by which a genetic locus increases risk for prostate cancer may be through regulating expression of a nearby gene, also known as a *cis*-expression quantitative trait locus (*cis*-eQTL). *Cis*-eQTLs are often highly tissue-specific and a global *cis*-eQTL study has yet to be performed in prostate tissue.

Objective: The objective of this study was to determine if any of the known prostate cancer associated genetic loci also act as *cis*-eQTL in prostate tissue, as a means of generating testable hypotheses for the functional mechanism underlying the genetic association.

Methods: 50 prostate cancer cases were genotyped for single nucleotide polymorphisms (SNPs) from genomic DNA and had RNA-seq data generated from paired primary prostate tumors. Prostate cancer associated SNPs were extracted from the NHGRI catalog of published GWAS results. *Cis*-eQTL analysis of transcripts within a 1Mb flanks of each SNP was performed, testing for an allelic effect of transcript levels versus genotypes with false discovery rates (FDR) computed to account for multiple testing.

Results: Out of 586 total transcript-SNP associations tested, 27 had an FDR less than or equal to 10%. These 27 putative eQTL coalesced into 13 distinct genetic loci. The most significant result was correlation of the IRX4 transcript with the prostate cancer risk SNP rs12653946, tagged by rs10866528 in our study ($p=4.91\times10^{-5}$, FDR= 0.00468).

Conclusions: We identified IRX4 as a prostate tissue eQTL through global *cis*- expression correlation with prostate cancer risk variants. *Cis*-eQTL analysis in the relevant tissue, even with a small sample size, can be a powerful method to generate hypotheses to move the functional follow-up of GWAS studies forward.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1101W

Interplay of tumor suppressor genes and oncogenes in ovarian cancer. *M. Zhao[1], S. Sun[1], Z. Zhao[1,2,3,4].* 1) Department of Biomedical Informatics, Vanderbilt University School of Medicine, Nashville, Tennessee, United States of America; 2) Department of Cancer Biology, Vanderbilt University School of Medicine, Nashville, Tennessee, United States of America; 3) Department of Psychiatry, Vanderbilt University School of Medicine, Nashville, Tennessee, United States of America; 4) Center for Quantitative Sciences, Vanderbilt University, Nashville, Tennessee, United States of America.

Investigators have discovered numerous protein-coding and non-coding tumor suppressor genes (TSGs) and oncogenes (OCGs) that control cell proliferation during cancer development. Additionally, TSGs and OCGs may act as modulators of transcription factors (TFs) to influence gene regulation. To better understand the post-translational modulation of TSGs and OCGs to TF gene regulation, we developed a novel computational framework for identifying target genes of TSGs and OCGs. This framework uses TFs as bridges through the integration of protein-protein interactions and gene expression data. We applied this pipeline to ovarian cancer and constructed a three-layer regulatory network. In the network, the top layer consisted of modulators (TSGs and OCGs), the middle layer included TFs, and the bottom layer contained target genes. Based on regulatory relationships in the network, we compiled TSG and OCG profiles and performed clustering analyses. Interestingly, we found TSGs and OCGs formed two distinct branches. The genes in the TSG branch were significantly enriched in DNA damage and repair, macromolecule metabolism regulation, cell cycle and apoptosis, while the genes in the OCG branch were significantly enriched in intracellular signaling, such as in the ErbB signaling pathway. Remarkably, their specific targets showed a reversed functional enrichment in terms of apoptosis and the ErbB signaling pathway: the target genes regulated by OCGs were only enriched in anti-apoptosis and the target genes regulated by TSGs were only enriched in the ErbB signaling pathway. Our study suggested a competitive regulatory mechanism acting upon apoptosis and the ErbB signaling pathway through specific target genes. We further collected 710 human TSGs, 628 Mouse TSGs and 567 Rat TSGs from thousands of literature sources and provide a web-based tumor suppressor gene database (TSGene DB, http://bioinfo.mc.vanderbilt.edu/TSGene/) with extensive annotations on cellular function, gene expression, mutation, methylation, transcription factors, post-translational modification, and protein-protein interaction.

1102T

Establishment of Breast, Leiomyosarcoma and Thyroid Cancer Stem Cells derived from carriers of p.R337H TP53 germline mutation within one Li-Fraumeni Syndrome Family. *L.I. Mambelli[2,3], I. Kerkis[3], F. Fortes[2], P. Hainaut[4], M.I.W. Achatz[1,2].* 1) Departamento de Oncogenética, Hospital A.C. Camargo, São Paulo, São Paulo, Brazil; 2) Centro Internacional de Pesquisa, Hospital A.C. Camargo, São Paulo, Brasil; 3) Laboratório de Genética, Instituto Butantan/Universidade de São Paulo, São Paulo, Brasil; 4) International Prevention Research Institute, Lyon, France.

Li-Fraumeni Syndrome (LFS), an inherited cancer predisposition syndrome, is associated with germ line mutations in TP53 and it is characterized by a high risk of multiple early cancers. In Brazil, a variant form of this syndrome is exceedingly frequent due to a widespread founder TP53 mutation, p.R337H, detected in about 0.3% of the general population in Southern Brazil. This mutation occurs in p53 oligomerization domain and disrupts p53 oligomerization and DNA-binding activity in pH-dependent manner. In this study we have isolated and characterized Cancer Stem Cells (CSC) from tumors (breast, leiomyosarcoma and thyroid) of p.R337H mutation carriers, who received clinical diagnosis of LFS. Clinical characterization and molecular profiling, through direct sequencing of TP53 gene in 91 member of one large LFS family was performed. This analysis confirmed the presence of p.R337H mutation in 38 members . All carriers received screening according to NCCN protocol and in a period of 4 year- follow up three tumors were diagnosed in three carriers within this family. Tumors were collected during surgery and primary cultures were initiated. Adherent cells and spheroids were derived from these different tumor types. Spheroids derived from a breast cancer (BC) were further analyzed to demonstrate positive immunolabeling for CD44, Oct4, Sox2, PCNA and Ki-67 antibodies. Time-lapse video microscopy showed rapid growth, frequent asymmetric division and absence of senescent phenotypes of these cells for least 17 passages. Floating and adherent spheroids contained self-renewing stem cells capable of long-term proliferation without losing their primary characteristics. These properties are similar to those of stem cells with inactivated p53 protein. p.R337H CSC may provide an useful model for studying the underlying defects related to inherited cancer in p.R337H carriers.

1103F

Functional validation using eQTL needs cell-specific approaches. *S. Kim[1], H.-H. Won[2], J.-W. Kim[3], D.H. Kim[4], S.-T. Lee[3], S. Kim[2], S.-H. Kim[3], C.W. Jung[5].* 1) Dept. of Health Sciences and Technology, Graduate School, Samsung Advanced Institute for Health Science and Technology, Sungkyunkwan University, Korea; 2) Samsung Biomedical Research Institute, Samsung Medical Center, Korea; 3) Dept. of Laboratory Medicine and Genetics, Sungkyunkwan University School of Medicine, Samsung Medical Center, Korea; 4) Chronic Myelogenous Leukemia Group, Department of Hematology/Medical Oncology, Princess Margaret Hospital, University Health Network, University of Toronto, Toronto, Canada; 5) Department of Hematology/Oncology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea.

Genome-wide association study (GWAS) has been used to find loci related to diseases but it needs functional validation. Finding expression quantitative trait loci (eQTLs) is an important tool for revealing the functional relationships between variations and the corresponding gene. eQTL studies have focused on the gene expression of lymphoblastoid cell lines (LCL), and subsequently, most eQTL databases contain the results based on LCL. Chronic myeloid leukemia (CML) is a clonal bone marrow stem cell disorder in which proliferation of mature granulocytes and their precursors is a major characteristic. We previously identified novel locus 6q25.1 associated with susceptibility to CML (Kim *et al.*, Blood, 2011, 117(25):6906-11). Therefore, we conducted experiments to confirm whether SNPs in this locus affect expression of CML candidate genes in two types of white blood cells, mononuclear cells and granulocytes. Peripheral blood samples from normal subjects were obtained. We isolated granulocytes and mononuclear cells from the blood to compare gene expression levels and purified DNA for genotyping. We then extracted total RNA and equal amounts of RNA were used to synthesize cDNA. Quantitative PCR was performed using Real-Time PCR 7900 HT (Applied Biosystems, USA). Relative expression of a candidate gene, *RMND1*, was calculated using the threshold cycle number (Ct) and normalized according to *GAPDH*. We selected SNPs which were known as eQTLs and DNase I hypersensitive QTLs (dsQTLs) in the locus 6q25.1 using the eQTL database (http://eqtl.uchicago.edu/) and genotyped eight SNPs. Statistical significance was evaluated by Kruskal-Wallis test using the R software. We observed that three SNPs (rs3900024, rs7765741, and rs78970048) in this locus were marginally associated with gene expression of mRNA in mononuclear cells (p-value: 0.06, 0.06 and 0.05 respectively) but not in granulocytes (p-value: 0.53, 0.53 and 0.46 respectively). These SNPs were reported as eQTLs and dsQTLs in the eQTL database. This study revealed that gene expression patterns in each cell according to genotypes were substantially different, consistent with a previous report that several eQTLs had cell type-dependent directional effects in B cells and monocytes (Fairpax *et al.*, Nat Genet, 2012, 25;44(5)). Therefore, we suggest that functional validation using eQTL needs to be conducted in cell lines that are relevant to the disease.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1104W

Functional characterization of HLA allelotypes associated with follicular lymphoma risk. *F.C.M. Sillé, L. Conde, C.F. Skibola.* School of Public Health, Division of Environmental Health Sciences, University of California Berkeley, 237A Hildebrand Hall, Berkeley, CA 94720-7360, USA.

Background: Family and epidemiological studies suggest an important genetic role in the etiology of lymphomas. In recent genome-wide association studies of follicular lymphoma (FL), several genetic susceptibility loci have been identified in the human leukocyte antigen (HLA) Class II region on chromosome 6p21.32 including rs10484561 that is associated with a two-fold increased risk of FL (OR = 1.95, $P = 1.12\times10^{-29}$). Moreover, high-resolution HLA typing using next-generation sequencing identified a number of HLA alleles associated with FL including the extended haplotype *DRB1*01:01—DQA1*01:01—DQB1*05:01* that is in complete linkage disequilibrium with rs10484561 ($r^2 = 0.99$). We hypothesize that these FL risk HLA alleles possibly influence the immune response against tumor development by altering the presentation and expression of specific antigens. To gain further understanding on the mechanism by which these HLA alleles might act, we explored the effect of HLA alleles in gene expression at both transcriptional and protein levels.

Methods: We selected 40 lymphoblastoid cell lines from the Coriell Institute with available genotype and HLA typing. Transcription was tested using RT-qPCR, whereas protein expression was tested using flow cytometry.

Results: Using RT-qPCR, we did not find significant changes in transcript levels of HLA-DRB1, HLA-DQB1 or HLA-DQA1 upon the presence of rs10484561 risk allele or the *HLA-DRB1*0101—DQA1*0101—DQB1*0501* haplotype. Flow cytometry with HLA-specific antibodies and pan-HLA antibodies (against DR/DQ/DP) also did not reveal significant differences in surface expression of HLA-DRB1, HLA-DQB1 or HLA-DQA1.

Conclusion: The extended haplotype *HLA-DRB1*0101—DQA1*0101—DQB1*0501*—rs10484561 was previously identified and validated as a risk haplotype associated with FL. However, we did not detect significant differences in transcript levels or surface protein expression upon the presence of these risk alleles. These findings suggest that the increased risk of FL associated with the *HLA-DRB1*0101—DQA1*0101—DQB1*0501* haplotype and the linked rs10484561 SNP are not due to changes in HLA expression. We will validate these findings by Western blotting and ELISAs and test antigen presentation using T-cell activation assays.

1105T

MicroRNA expression for classification of histiocytic sarcomas in dogs. *J. Aguirre Hernandez, J.M. Dobson, T.M. Hoather, D.R. Sargan, F. Constantino-Casas.* Veterinary Medicine, University of Cambridge, Cambridge CB3 0ES, United Kingdom.

Soft tissue sarcomas are rare in humans but common in dogs (up to 40 percent of malignancies in some breeds). This makes it possible to rely on the dog to better understand these tumours. Flat-coated retrievers have a particularly high incidence of histiocytic sarcomas of which two forms can be recognised: limb and visceral. To aid in their diagnosis, classification and prognosis, the expression of 22 microRNAs (miRNA) was studied in formalin fixed paraffin embedded tumours of this breed. Samples comprised visceral histiocytic sarcomas, limb histiocytic sarcomas, peripheral nerve sheath tumours, histiocytomas and a high grade soft tissue sarcoma. MiRNA expression profiling correctly distinguished all of them. However, contrary to what was expected, the expression profiles did not classify histiocytic sarcomas according to the site of presentation. Instead, 2 different histiocytic sarcoma groups were distinguished, each including both limb and visceral forms. In addition to this, a single limb histocityc sarcoma was not grouped with any of the two histiocytic sarcoma clusters, or with any other tumour, appearing to be an outlier in terms of its miRNA expression profile. This clustering of the histiocytic sarcomas would be consistent with the recognition of several dendritic cell lineages.

1106F

Molecular urine assay of PCA3,MMP2,MMP9 and KLK3 in men with prostate cancer and benign prostatic hyperplasia. *H. Akin[1], A. Simsir[2], F. Hazan[1], T.R. Ozdemir[1], I. Cureklibatir[2], G. Itirli[1], C. Gunduz[3], C. Ozkinay[1], F. Ozkinay[1].* 1) Department of Medical Genetics, Ege University, Medical Faculty, IZMIR, Turkey; 2) Department of Urology, Ege University, Medical Faculty, IZMIR, Turkey; 3) Department of Medical Biology, Ege University, Medical Faculty, IZMIR, Turkey.

Men with elevated serum prostate-specific antigen (PSA) levels and negative prostate biopsy findings present a complicated status in definitive prostate cancer approach. In this study our aim was the evaluating the diagnostic and prognostic utility of the investigational prostate cancer gene 3 (PCA3), matrix metalloproteinase 2(MMP2) and 9 (MMP9) and KLK3 mRNA expression profiling in prostate cancer(PCA) and benign prostatic hyperplasia (BPH). Seventy four PCA patients, 60 BPH patients and 41 controls were included in the study. The average ages of the groups PCA, BPH and controls were 61.20±11.05, 65.25±11.90 and 44.12±15.26 respectively. Urine was collected after digital rectal examination (three strokes per lobe) from the patients and controls. The specimens were collected from April 2009 to November 2011. The mRNA expression levels of PCA3, MMP2, MMP9 and KLK3 genes were detected using a non-invasive molecular urine assay method, RT PCR. PCA3 expression levels were found to be significantly higher (p=0.001) in PCA patients group comparing to the BPH patients and controls using student t test. Receiver operating characteristic curve analysis yielded an area under the curve of 0.659 for the PCA3 with p:0.002 significance. On the other hand, MMP9 expression level showed significant correlation with Gleason score in patients with PCA. PCA3 and MMP9 expression profiling in the urine of PCA patients may be useful in the diagnosis and management of PCA.

1107W

The effect of copy number variation (CNV) in the phase II detoxification genes, UGT2B17 and UGT2B28, on colorectal cancer risk. *A.Y. Angstadt[1], A.S. Berg[1], J. Zhu[1], P. Miller[2], T.J. Hartman[3], S.M. Lesko[4], J.E. Muscat[1], P. Lazarus[1,5], C.J. Gallagher[1,5].* 1) Department of Public Health Sciences, Penn State College of Medicine, Hershey, PA, United States; 2) National Cancer Institute, Rockville, Maryland, United States; 3) Department of Nutritional Sciences, Penn State University, University Park, PA, United States; 4) Northeast Regional Cancer Institute, Scranton, PA, United States; 5) Department of Pharmacology, Penn State College of Medicine, Hershey, PA, United States.

Colorectal cancer (CRC) incidence levels are highest in the developed world as the Western-Style diet is thought to be a contributing factor. Metabolism of dietary compounds is completed by phase II detoxification enzymes like UGT2B17 and UGT2B28. These enzymes are located in commonly deleted regions of the human genome, occurring in Caucasians at 27% and 13.5%, respectively. UGT2B17 copy number variation (CNV) was found associated with lung and prostate cancer risk and it has been shown that loss of UGT2B28 in CRC cells compared to nonneoplastic cells resulted in lower expression. Yet, no genetic association studies of UGT2B17 or UGT2B28 and CRC risk have been previously performed. We hypothesized that CNV of these genes may be causing different CRC risk levels in individuals exposed to particular dietary factors. We genotyped UGT2B17 and UGT2B28 CNVs in 665 Caucasian CRC cases (368 men, 297 women) and 621 Caucasian controls (359 men, 262 women) that had completed an extensive demographic/dietary questionnaire. Unconditional logistic regression was used to test the separate effect of UGT2B17 and UGT2B28 CNV on CRC risk, while polytomous logistic regression was used to analyze associations between genotype and anatomical subsites. All models were adjusted for age, BMI, kcal/day, NSAID use, CRC family history, and physical activity. A significant association between the UGT2B17 deletion allele (0/0) and decreased CRC risk was found in our entire population (0.815 OR, 0.668–0.994 95% CI). In addition, a significant association between UGT2B17 (0/0) and decreased CRC risk was found in rectal cancer patients (0.670 OR, 0.502–0.894 95% CI) but not colon cancer patients. Sex stratification yielded a decreased risk in UGT2B17 (0/0) men (0.724 OR, 0.552–0.950 95% CI), but no association was found in women. No significant association was found between UGT2B28 CNV and CRC risk before and after stratification. UGT2B17 has been shown to have higher expression and activity in men, which may be responsible for the sex-specific association we observed. In addition to its role in androgen metabolism, UGT2B17 is known to metabolize NSAIDs and flavonoids (antioxidants) so individuals with two copies of this gene may excrete these protective dietary compounds at a faster rate than individuals with one or zero copies. Therefore, decreasing the antioxidant effect on carcinogen detoxification, resulting in the observed decreased CRC risk in UGT2B17 (0/0) individuals.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1108T

Genetic sequence variant in microRNA genes and survival in non-small cell lung cancer. *A.K. Azad[1], X. Qiu[3], T. Popper[1], K. Boyd[1], Q. Kuang[1], Y.B. Berhane[3], H. Henrique[1], N. Perera[1], P. Prakruthi[1], D. Patel[1], S. Momin[1], M. Nakhla[1], E. Marjan[1], Z. Chen[1], D. Cheng[1], R. Feld[2], N.B. Leighl[2], F.A. Shepherd[2], M.-S. Tsao[4], W. Xu[3], G. Liu[1,2,5], S. Cuffe[2].* 1) Ontario Cancer Institute, Princess Margaret Hospital, University Health Network, Toronto, Ontario, Canada; 2) Division of Medical Oncology and Hematology, Princess Margaret Hospital, University Health Network, Toronto, Ontario, Canada; 3) Department of Biostatistics, Princess Margaret Hospital, University Health Network, Toronto, Ontario, Canada; 4) Department of Pathology, Princess Margaret Hospital, University Health Network, Toronto, Ontario, Canada; 5) Department of Environmental Health, Harvard School of Public Health, Boston, Massachusetts, USA.

Background: Lung cancer is the leading cause of cancer-related mortality worldwide and approximately 80% of the patients have NSCLC. Recent studies indicate that small non-coding RNA molecules known as miRNA can function as tumor suppressors and oncogenes. GSVs in miRNA genes, misexpression and altered miRNA processing have been implicated in carcinogenesis and cancer outcome. miRNA expression profiles and specific miRNA have been shown to be correlated with survival of lung adenocarcinoma (ADC). The aim of this study was to evaluate whether miRNA-related GSVs associated with cancer risk are also associated with overall survival (OS) and progression free survival (PFS) in stage-IV NSCLC patients. **Methods:** A total of 41 candidate GSVs in 33 genes in miRNA biogenesis pathways previously reported to be associated with cancer risk/outcome were genotyped in 325 stage-IV NSCLC patients. Patients were recruited for prospective study of the molecular epidemiology of thoracic lesions in PMH from 2006–2010. Illumina Custom GoldenGate Genotyping Panel was used for the genotyping assay platform. Multivariate Cox proportional hazard models adjusted for potential clinical prognostic factors were generated for OS and PFS. **Results:** Median age was 64y (range: 31–88); 71% had ADC. Median OS time was 1.10y (95% CI: 0.90–1.23); Median follow-up time was 1.9yr. Median PFS time was 6m (95% CI: 0.49–0.60). The variant allele of *C14orf101*:rs4901706:c.*170G>A and *NEIL1*:rs5745925:c.719-113delT (hsa-mir-631) found to be associated with OS and PFS when compared with the wild type allele. For OS, rs4901706 [aHR 1.42 (95% CI: 1.04–1.93), *p* 0.03] and rs5745925 [1.77 (95% CI: 1.16–2.70), *p* 0.01]; and for PFS, rs4901706 [1.46 (95% CI: 1.11–1.93), *p* 0.01] and rs5745925 [1.62 (95% CI: 1.07–2.44), *p* 0.02] showed significant association. **Conclusions:** The miRNA-related GSVs are not only a risk factor for lung cancer but also are associated with survival in patients with stage-IV NSCLC. The data demonstrated the GSVs are marginally associated with poorer survival in stage-IV NSCLC patients. Variant alleles of rs4901706 and rs5745925 have been previously associated with cancer risk (PMID: 17584784, 19138993). However, their functional relevance in cancer outcome has not been investigated. The validation of these findings in multiple cohorts and testing of their functionality at cellular level is also required. *GL and SC are senior co-authors.

1109F

The PROFILE Study: Genetic prostate cancer risk stratification for targeted screening. *E.K. Bancroft[1,2], E.C. Castro[2], N. Taylor[1,2], E. Page[2], E. Saunders[2], T. Dadaev[2], A. Lee[3], A. Antoniou[3], Z. Kote-Jarai[2], R. Eeles[2].* 1) Royal Marsden Hospital, Sutton, United Kingdom; 2) Institute of Cancer Research, Sutton, United Kingdom; 3) The University of Cambridge, UK.

A family history (FH) of prostate cancer (PC) is one of the main risk factors for the disease suggesting the importance of genetics in PC development. GWAS have identified a number of common single nucleotide polymorphisms (SNPs) that confer small but cumulatively substantial risks of PC. These findings open the possibility for exploring the use of SNPs in PC risk stratification for targeted screening. Objectives: PROFILE has been developed as a pilot study. The primary aim is to determine the feasibility of targeted PC screening using prostatic biopsy (PB) and its association with specific genetic profiles in men with FH. The secondary aims are to (1)determine PC risk related SNP signatures in men with FH; (2)study the correlation between those SNP signatures and PB outcomes; (3)estimate the sensitivity and specificity of PSA screening and optimal PSA threshold for PB in men with FH; (4)to assess the value of Diffusion Weighted Magnetic Resonance Imaging (DW-MRI) as a PC screening tool; (5)evaluate the psychosocial impact of undergoing genetic profiling in this context. Method: Eligible men were aged 40–69 years with a FH of PC. After informed consent patients provided blood samples to measure PSA level and for DNA extraction. All participants were asked to undergo PB +/− DW-MRI regardless of baseline PSA result. DNA from each participant was analysed for a panel of 39 SNPs associated with increased risk of PC using iPLEX Sequenom technology. The results were fed back to participants and an associated psychosocial study has collected data on cancer worry, anxiety, depression and illness perceptions. Results: 92 men were recruited and 77 PB performed by 1st May 2012. 33 men had a predicted relative risk (RR)<1 (median age 55 yrs; median PSA 1.5). In this subgroup 6 men (18.2%) were diagnosed with PC (median age 66.5 yrs, PSA 2.1). Amongst the 34 men with a RR>1 (median age 50.5 yrs; median PSA 1.3) 9 PCs (26.5%) have been identified (median age 58 yrs, PSA 5.4). DW-MRI had 33% false positives and 10% false negatives. 56 men completed baseline & follow-up questionnaires and no measurable adverse psychological outcomes were recorded. Conclusion: The results indicate that PB as a means of PC screening is feasible and acceptable in men with a FH of PC. Men had a good understanding of the limitations of genetic profiling. Our findings support the use of SNPs in PC risk stratification for targeted screening in a larger study.

1110W

Lgr5 expression and its relatedness with other tumor stem cell markers in human gastric adenocarcinoma. *W. Fan, Y. W. Guo, Y. Y. Xie, Y.S. Li, W. Chen.* Molecular Genetics, Hebei University, Baoding, Hebei, China.

Purpose: To investigate if the Lgr5 serves as a tumor stem cell marker in the human gastric adenocarcinoma as it does in mouse, and to understand its relatedness with other stem cells markers (Bmi1 and Villin) and putative tumor stem cell markers (CD44, ALDH1, CD26 and CD133) as reported . Methods: Immunohistochemical detection of Lgr5, Bmi1, Villin, CD44, ALDH1, CD26 and CD133 in dissected human gastric adenocarcinoma samples and the adjacent normal gastric tissues. RT-PCR is also applied to detect Lgr5 expression in transcription level of several gastric cell lines. Results: Lgr5 is consistently expressed at gland bottom of normal gastric mucosa in human, and is evenly distributed in pyloris, cardia, lesser curvature, and great curvature, but not in the low-differentiated gastric adenocarcinoma. Lgr5 is co-expressed with Bmi1 and Villin, but Villin is also weakly expressed in biopsy tissue samples of gastric adenocarcinoma. Lgr5 expression is overlapped with putative gastrointestinal tract tumor stem cells markers such as CD44, ALDH1, CD26 and CD133. The results from immunofluorescence double labeling show that Lgr5 is co-expressed with CD44 and ALDH1 even in the same cells. In addition, CD26 and CD133 are expressed in low differentiated gastric adenocarcinoma respectively, and CD133 demonstrates a very strong positive marker. RT-PCR results show that Lgr5 is weakly expressed in the gastric tumor cell line cultures in vitro though it can not be detected by immunohistochemistry and immunoflourescence. Conclusion: Lgr5 is expressed at the bottom of the glands in the human gastric mucosa and its expression is overlapped with other stem cell markers, indicating Lgr5 may serves as the putative stem cells marker in the human stomach as it is in mouse, and it may share the similar pathway in cell development and differentiation with Bmi1, Villin, CD44, ALDH1, CD26 and CD133.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1111T

The TP53 16 bp duplication polymorphism is associated with breast cancer. *M. Gallegos, Jr.[1], A. Gutiérrez[1,5], A. Ramos[1,6], E. Salas[2], J. M. Castro[3], A. M. Puebla[7], L. E. Figuera[4], G. M. Zúñiga[1]*. 1) División de Médicina Molecular, CIBO; 2) Jefatura de Oncología, UMAE, Hospital de Gineco-Obtetricia; 3) Jefatura de Onco-Cirugía, UMAE, Hospital de Especialidades; 4) División de Genética, CIBO, CMNO, IMSS; 5) Licenciatura en Químico Farmacobiologo, CUCEI, UdeG; 6) Doctorado de Farmacología, UdeG; 7) Laboratorio de Inmunofarmacología, CUCEI, UdeG,.Guadalajara, Jalisco. México.

The TP53 tumor suppressor gene encodes the nuclear phosphoprotein p53, which plays an important role in cell cycle, apoptosis, DNA reparations and angiogenesis. The TP53 contains common genetics polymorphisms that influence gene activity. In the present study we describe to frequency of 16 bp duplications polymorphism of TP53 gene in 304 breast cancer patients and 204 woman health controls from Mexico. The frequency observed was of 4% (13/304) and 1% (1/204) of 16 bp duplication genotype; 20% (59/304) and 11% (24/204) of heterozygous (16 bp duplication/16 bp no duplication) genotype; 76% (232/304) and 88% (179/204) of 16 bp no duplication genotype in cases and controls respectively. We found that odds ratio (OR) = 9.0 95% confidence intervals 1.33–387, p = 0.001. These results suggest that 16 bp duplication polymorphism of TP53 gene is associated with Mexican breast cancer patients.

1112F

Statistics of Cellular Evolution in Leukemia: Allelic Variations in Patient Trajectories Based on Immune Repertoire Sequencing. *H. Gao[1], C. Wang[1], AC. Logan[2], CD. Bustamante[3], J. Seok[1], DB. Miklos[2], RW. Davis[1], MW. Feldman[4], W. Xiao[1,5]*. 1) Stanford Genome Tech. Center and Biochemistry Dept, Stanford Univ, Stanford, CA; 2) Division of Blood and Marrow Transplantation, Medicine Dept, Stanford Univ, Stanford, CA; 3) Genetics Dept, Stanford Univ, Stanford, CA; 4) Biology Dept, Stanford Univ, Stanford, CA; 5) Department of Surgery, Shriners Burn Center and Massachusetts General Hospital, Harvard Medical School, Boston, MA.

The evolution of a cancer system consisting of cancer clones and normal cells is a complex and dynamic process with multiple interacting factors including clonal expansion, somatic mutation, and sequential selection. As a typical example, in patients with chronic lymphocytic leukemia (CLL), a monoclonal population of transformed B cells expands to dominate the B cell population in the peripheral blood and bone marrow. This expansion of transformed B cells suggests that they might evolve through processes distinct from those of normal B cells. Recent advances in next generation sequencing enable the high-throughput identification and tracking of individual B cell clones through sequencing of the V-D-J junction segments of the immunoglobulin heavy chain (IGH). Here, we developed a statistical approach to modeling cellular evolution of the immune repertoire. Adapting the infinitely many alleles model from population genetics, we studied abnormalities occurring in the immune repertoire of patients as substantial deviations from the null model. The Ewens sampling test (EST) distinguished the immune repertoires of CLL patients with imminent relapse from healthy controls and patients in sustained remission. Extensive simulations based on sequencing data showed that EST is sensitive in detecting cancer-related derangements of the IGH repertoire. In addition, we suggest two potentially useful parameters: the rate at which donor's B cell clones enter the circulation and the average time to regenerate a transplanted immune repertoire, both of which help to classify relapsing CLL patients from those in sustained remission and provide additional information about the dynamics of immune reconstitution in these patients. We anticipate that our models and statistics will be useful in diagnosis and prognosis of leukemia, and may be adapted for application to other diseases related to adaptive immunity.

1113W

Genetic polymorphisms and colorectal cancer risk: A systematic review of meta-analyses. *B. Haerian[1], M. Haerian[2]*. 1) University of Malaya, Kuala Lumpur, Malaysia; 2) Shaheed Beheshti of Medical Sciences, Tehran, Iran.

Background: Colorectal cancer (CRC) is the second leading cause of morbidity and mortality in United States. A series of genetic and environmental factors drive the development of CRC. Accumulated evidence have identified involvement of numerous low penetrance susceptibility genes loci in CRC. Alterations in these genes could change protein structure and function associated with the hallmarks of cancer-related biology. Objectives: The aim of this study is to highlight an array of genes that reflect the cellular processes and regulate tumour activities and progression in CRC using the significant loci from evaluated meta-analyses in the literature. Methods: In this study, we have searched MEDLINE, Embase, and the Cochrane Database of Systematic Reviews for meta-analyses and pooled analyses involving the associations between various genes polymorphisms and CRC risk and then we have systematically evaluated them. Results: A total of 133 published meta-analyses and pooled analyses articles evaluating the association of low penetrance susceptibility variants with CRC risk have been investigated. Of these reports, 118 have been included in our study. The data from 118 polymorphisms of 65 genes, one microRNA, and six chromosomal regions are analyzed. Finally, based on the six hallmark trait, the role of the significant loci in CRC is classified and discussed. Conclusion: The data of this comprehensive study shows the considerable role of some susceptible genes variants in cancer hallmarks during multi-steps CRC development.

1114T

NRF2 and sulfiredoxin genetic polymorphisms and protein expression predict outcome in breast cancer. *J.M. Hartikainen[1,2,3], M. Tengström[4,5], V.-M. Kosma[1,2,3], V. Kinnula[6], A. Mannermaa[1,2,3], Y. Soini[1,2,3]*. 1) Institute of Clinical Medicine, Pathology and Forensic Medicine, University of Eastern Finland, Kuopio, Finland; 2) Biocenter Kuopio and Cancer Center of Eastern Finland, University of Eastern Finland, Kuopio, Finland; 3) Imaging Center, Clinical Pathology, Kuopio University Hospital, Kuopio, Finland; 4) Institute of Clinical Medicine, Oncology, University of Eastern Finland, Kuopio, Finland; 5) Cancer Center, Kuopio University Hospital, Kuopio, Finland; 6) Department of Medicine, Pulmonary Division, University of Helsinki and Helsinki University Central Hospital, Helsinki, Finland.

Nuclear factor erythroid 2-related factor 2 (NRF2) is a transcriptional factor which senses oxidative and xenobiotic stress in the cells. As a response to oxidative and xenobiotic stress NRF2 activates several protective genes, such as sulfiredoxin. Defects in the NRF2 pathway may increase cancer susceptibility. In tumor cells the activation of NRF2 may lead to chemo- and radioresistance and thus affect patient outcome. Nine single nucleotide polymorphisms on *NRF2* gene and eight on *SRXN1* (sulfiredoxin) were genotyped in a material of 452 breast cancer patients and 370 controls. Protein expression of NRF2 and sulfiredoxin was studied in 373 breast carcinomas by immunohistochemistry. Statistical significance of the associations between genotypes, protein expression, clinicopathological variables and survival was assessed. High extent (>25%) cytoplasmic NRF2 positivity was observed in 237/361 (66%) and sulfiredoxin positivity in 82/363 (23%) of the cases. *NRF2* rs6721961 genotype TT associated with increased risk of breast cancer (P=0.008, OR=4.656, CI=1.350–16.063) and T allele with low extent NRF2 protein expression (P=0.0003, OR=2.420, CI=1.491–3.926) and negative sulfiredoxin expression (P=0.047, OR=1.867, CI=1.002–3.478). *NRF2* rs2886162 allele A associated with low extent NRF2 expression (P=0.011, OR=1.988, CI=1.162–3.400) and genotype AA with poorer survival (P=0.032, HR=1.687, CI=1.047–2.748). *NRF2* rs1962142 allele T associated with low extent cytoplasmic NRF2 (P=0.036) and negative sulfiredoxin expression (P=0.042). *NRF2* rs2706110 genotype AA associated with increased risk of breast cancer and *SRXN1* rs6053666 allele C was protective (P values 0.011 and 0.017). Four other *SRXN1* SNPs associated with breast cancer survival. Our results demonstrate that in the NRF2 pathway there are genetic polymorphisms that affect both the susceptibility for breast cancer and the outcome of the breast cancer patients, thus underlining the complex effect of NRF2 in cancer progression. The results implicate that mechanisms associated with ROS and NRF2 pathway are involved in breast cancer initiation and progression also at the level of genetic predisposition. ROS associated mechanisms appear to play a role in the behavior and treatment of breast cancer to the extent of being reflected in the survival of the patients. Studying antioxidative mechanisms may thus pave the way for new treatment modalities based on inhibition of such mechanisms in breast cancer cells.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1115F

Complex tumor genomes inferred from plasma-DNA and circulating tumor cells of patients with cancer. *E. Heitzer[1], M. Auer[1], E.M. Hoffmann[1], C. Beneken[2], M. Pichler[3], P. Ulz[1], S. Lax[4], J. Waldispuehl-Geigl[1], O. Mauermann[2], G. Pristauz[5], C. Lackner[6], G. Höfler[6], F. Eisner[3], E. Petru[5], H. Sill[7], H. Samonigg[3], K. Pantel[2], S. Riethdorf[2], T. Bauernhofer[3,8], J.B. Geigl[1], M.R. Speicher[1].* 1) Institute of Human Genetics, Medical University of Graz, Harrachgasse 21/8, A-8010 Graz, Austria; 2) Institute of Tumor Biology, University Medical Center Hamburg Eppendorf, Martinistr. 52, D-20246 Hamburg, Germany; 3) Division of Oncology, Medical University of Graz, Auenbruggerplatz 15, A-8036 Graz, Austria; 4) Department of Pathology, General Hospital Graz West, Goestingerstrasse 22, A-8020 Graz, Austria; 5) Department of Obstetrics and Gynecology, Medical University of Graz, Auenbruggerplatz 14, A-8036 Graz, Austria; 6) Institute of Pathology, Medical University of Graz, Auenbruggerplatz 25, A-8036 Graz, Austria; 7) Division of Hematology, Medical University of Graz, Auenbruggerplatz 38, A-8036 Graz, Austria; 8) LKH Leoben, Department for Hematology and Oncology, Vordernberger Straβe 42; A-8700 Leoben, Austria.

As the number of available predictive biomarkers is steadily rising, the serial accurate monitoring of tumor genotypes, which are prone to changes, becomes increasingly important in the clinical management of human cancer. In this study, we investigated whether complex tumor genomes can be inferred from peripheral blood, i.e. plasma DNA and/or CTCs from patients with cancer. Therefore, we analyzed peripheral blood of 32 patients with advanced-stage colorectal cancer (CRC) for mutant DNA-fragments in the plasma and circulating tumor cells (CTCs) using array-CGH and next generation sequencing technology. We found that a subset of patients simultaneously releases tumor DNA from two different cell populations, i.e. apoptotic and non-apoptotic cells, into the circulation. This phenomenon co-occurs with increased numbers of CTCs and elevated plasma-DNA concentrations and has consequences on diagnostic tests, as tumor specific copy number changes can be established directly from plasma-DNA. Our CTC analyses included massively parallel sequencing of a panel of 68 CRC associated genes and a comparison with the mutation status in the primary tumors and associated metastases. Ultra-deep sequencing of “private mutations” found in only one CTCs revealed allele frequencies ranging from 0.02 to 0.45 in corresponding samples including plasma, tumor, and metastasis. We demonstrate that CTCs have an enormous heterogeneity, which is characterized by different copy number changes and mutations in cancer driver genes. The fact that these results were confirmed in 35 patients with breast cancer suggests that our approach is applicable to a variety of tumor entities. Our analyses provide novel insights into tumor evolution and may expand current non-invasive approaches for disease monitoring.

1116W

Novel integrative genomics approach for associating GWAS information with triple negative breast cancer. *C. Hicks[1,2], K. Backus[1], A. Pannuti[1], L. Miele[1].* 1) Cancer Institute Univ Mississippi Med Center, 2500 N State Street, Jackson, MS 39216; 2) Dept Medicine, Univ. Mississippi Med. Center, 2500 N State Street, Jackson, MS 39216.

Introduction: Genome-wide association studies (GWAS) have achieved great success in identifying genetic variants associated with increased risk for breast cancer. However, the association between genetic variants and the most aggressive subset of breast cancer, the triple negative breast cancer (TNBC) remains a central puzzle. TNBC (Tumors that do not express estrogen receptors, progesterone receptors, or HER2) constitute 15%–20% of all breast cancers diagnosed. TNBC is significantly more aggressive than other subtypes of breast cancer and disproportionately affects younger premenopausal women, with a higher mortality rate among African-American women. Currently, chemotherapy remains the only effective therapeutic modality. The objective of this study was to determine whether genes containing SNPs associated with increased risk for breast cancer are associated with and could stratify TNBC. Methods: Here we present a novel integrative genomics approach that combines GWAS information with gene expression data from breast cancer patients classified as TNBC at diagnosis. Specifically, we first identified 600 SNPs (genetic variants) mapped to 200 genes identified from 50 GWAS involving over 300,000 cases and over 300,000 controls. We then investigated the association of these variants and genes with TNBC and identified biological pathways using gene expression as the intermediate phenotype. Gene expression data was derived from normal-like (N=20), basal-like (N=29), and basal (N=75) breast cancer patients and cancer-free controls (N=142 subjects). We performed unsupervised and supervised analysis, and controlled for multiple hypothesis testing using the false discovery rate (FDR). Pathway prediction and network visualization was performed using Ingenuity (IPA). Results: We identified 159 genes contain SNPs, that were significantly (P<10-5) associated with and stratified TNBC. 40 of these genes contained SNPs replicated in multiple independent GWAS studies. We identified networks made of multiple pathways including DNA repair, DNA mismatch repair, cell cycle, apoptosis, P53, metabolism, cellular growth and proliferation. The integration of GWAS information with gene expression data provides insights about the association of genetic variants with TNBC. In conclusion, our results demonstrate the potential applicability of integrating GWAS information and gene expression data to identify genetic variants and genes associated with TNBC.

1117T

Germ line variation in TP53 regulatory network genes associates with breast cancer survival and treatment outcome. *M. Jamshidi[1], M. K. Schmidt[2, 3], T. Dörk[4], M. Garcia-Closas[5, 6], T. Heikkinen[1], S. Cornelissen[2], S. van den Broek[3], P. Schürmann[4], A. Meyer[7], T. W. Park-Simon[4], J. Figueroa[8], M. Sherman[9], J. Lissowska[10], G. T. Keong[11], A. Irwanto[11], M. Laakso[12], S. Hautaniemi[12], K. Aittomäki[13], C. Blomqvist[14], J. Liu[11], H. Nevalinna[1].* 1) Department of Obstetrics and Gynecology, Biomedicum Helsinki, University of Helsinki and Helsinki University Central Hospital, P.O. Box 700, 00029 Helsinki, Finland; 2) Department of Molecular Pathology, Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands; 3) Department of Psychosocial Research and Epidemiology, Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands; 4) Clinics of Gynecology and Obstetrics, Hannover Medical School, Hannover, Germany; 5) Division of Cancer Epidemiology and Genetics, National Cancer Institute, Department of Health and Human Services, Bethesda, Maryland; 6) Division of Genetics and Epidemiology, Institute of Cancer Research and Breakthrough Breast Cancer Research Centre, London, UK; 7) Clinics of Radiation Oncology, Hannover Medical School, Hannover, Germany; 8) Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD, 20892; 9) Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, MD, USA; 10) Department of Cancer Epidemiology and Prevention, M. Sklodowska-Curie Memorial Cancer Center and Institute of Oncology, Warsaw, Poland; 11) Human Genetics Division, Genome Institute of Singapore, Singapore; 12) Computational Systems Biology Laboratory, Genome-Scale Biology Research Program, Institute of Biomedicine, University of Helsinki, Finland; 13) Department of Clinical Genetics, Helsinki University Central Hospital, Helsinki, Finland; 14) Department of Oncology, Helsinki University Central Hospital, Helsinki, Finland.

Germ line variation in the TP53 network genes PRKAG2, PPP2R2B, CCNG1, PIAS1 and YWHAQ was previously suggested to have an impact on drug response in vitro. Here, we investigated the effect on breast cancer survival of germ line variation in these genes in 925 Finnish breast cancer patients and further analyzed 5 SNPs in PRKAG2 (rs1029946, rs4726050, rs6464153, rs7789699) and PPP2R2B (rs10477313) for 10-year survival in breast cancer patients, interaction with TP53 R72P and MDM2-SNP309, outcome after specific adjuvant therapy, and correlation to tumor characteristics in 4701 invasive cases from four data sets. We found evidence for carriers of PRKAG2-rs1029946 and PRKAG2-rs4726050 having improved survival in the pooled data (HR 0.53, 95% CI 0.3–0.9; P = 0.023 for homozygous carriers of the rare G-allele and HR 0.85, 95% CI 0.7–0.9; P = 0.049 for carriers of the rare G allele, respectively). PRKAG2-rs4726050 showed a significant interaction with MDM2-SNP309, with PRKAG2-rs4726050 rare G-allele having a dose-dependent effect for better breast cancer survival confined only to MDM2 SNP309 rare G-allele carriers (HR 0.45, 95% CI 0.2–0.7; P = 0.001). This interaction also emerged as an independent predictor of better survival (P = 0.047). PPP2R2B-rs10477313 rare A-allele was found to predict better survival (HR 0.82, 95% CI 0.6–0.9; P = 0.018), especially after hormonal therapy (HR 0.66, 95% CI 0.5–0.9; P = 0.048). These findings warrant further studies and suggest that genetic markers in TP53 network genes such as PRKAG2 and PPP2R2B might affect prognosis and treatment outcome in breast cancer patients.

1118F

Detection of FLT3 Internal Tandem Duplication and D835 Mutations in Iranian patients with acute myeloid leukemia. *L. Kokabee[1], A. Ahmadzadeh[2], M. Kokabee[2], M. Karimipoor[1].* 1) Molecular Medicine Group, Biotechnology Department, Pasteur Institute of Iran, Tehran, Tehran, Iran; 2) Shariati Hospital, Tehran university of medical science, Tehran, Iran.

Acute myeloid leukemia (AML) is a disease with marked heterogeneity in both response to therapy and survival. The advent of molecular diagnostics has heralded an explosion in new prognostic factors. The FLT3 gene encodes the tyrosine kinase receptor critical for normal hematopoiesis. There are two types of FLT3 mutations: an internal tandem duplication (ITD) and a point mutation (D835), which both have constitutively active tyrosine kinase, thereby promoting proliferation and inhibiting apoptosis in leukemic cells. In this study, mutations in 31 untreated cytogenetic normal AML patients (The mean age of patients was 39 year) were analyzed. For this purpose, genomic DNA was extracted from peripheral blood (referred to Shariati hospital, Tehran, Iran) by standard methods. PCR amplification was performed for FLT3 mutation status (ITD and D835). After amplification, PCR product of FLT3/D835 mutation was digested with EcoRV. FLT3/ITD was found in 2 patients (6.45%) and D835 mutation in 4 patients (12.90%). These mutations also appear to be activating and to portend a worse prognosis. Identification of FLT3 mutations is important because it provides prognostic information and may play a pivotal role in determining appropriate treatment options.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1119W

Association of Gastric Cancer With Amerindian Ancestry, Socioeconomic and Nutritional Factors and a Gene Candidate Study in a Latin American Admixed Population. *L. Lacorte[1], R. Zamudio[1], G. Soares-Souza[1], P. Herrera[2], L. Cabrera[2], C. Hooper[3], J. Cok[4], J. Combe[5], G. Vargas[6], W. Prado[7], S. Schneider[8], M. Rodrigues[1], S. Chanock[9], D. Berg[10], R. Gilman[2,3,11], E. Tarazona-Santos[1,2], F. Kehdy[1].* 1) Departamento de Biologia Geral, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil; 2) Asociación Benéfica PRISMA and Universidad Peruana Cayetano Heredia, Lima, Peru; 3) 3 Laboratorios de Investigacion y Desarrollo, Facultad de Ciencias, Universidad Peruana Cayetano Heredia, Lima, Peru; 4) Departamento de Patología, Hospital Nacional Cayetano Heredia, Lima, Peru; 5) Departamento de Gastroenterologia, Instituto Nacional de Enfermedades Neopláscas, Lima, Peru; 6) Servicio de Gastroenterologia, Hospital Nacional Arzobispo Loayza; 7) Servicio de Gastroenterologia, Hospital Dos de Mayo, Lima, Peru; 8) Departamento de Estatística, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil; 9) Laboratory of Translational Genomics of the Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Grovemont Circle, Advanced Technology Center, Gaithersburg, USA; 10) Department of Molecular Microbiology, Washington University Medical School, St Louis, MO, USA; 11) Department of International Health, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, USA.

Gastric cancer is one of the most lethal types of cancer and its incidence varies worldwide, with the Andean region of South America showing high incidence rates. We evaluated the genetic structure of the population from Lima (Peru) and performed a case-control genetic association study to test the contribution of African, European, or Native American ancestry to risk for gastric cancer, controlling for the effect of non-genetic factors. A wide set of socioeconomic, dietary, and clinic information was collected for each participant in the study and ancestry was estimated based on 103 ancestry informative markers. Although the urban population from Lima is usually considered as mestizo (i.e., admixed from Africans, Europeans, and Native Americans), we observed a high fraction of Native American ancestry (78.4% for the cases and 74.6% for the controls) and a very low African ancestry (<5%). We determined that higher Native American individual ancestry is associated with gastric cancer, but socioeconomic factors associated both with gastric cancer and Native American ethnicity account for this association. Therefore, the high incidence of gastric cancer in Peru does not seem to be related to susceptibility alleles common in this population. Instead, our result suggests a predominant role for ethnic-associated socioeconomic factors and disparities in access to health services. In addition to ancestry analysis, we used a case-control design to assess if the SNPs in promoter of candidate genes IL8 (rs4073), IL8RB (rs4674258), PTGS2 (rs689465 and rs689466) and SNPs in other two genes (MSMB(rs10993994), FGFR2 (rs1219648)) associated with other types of cancer are associated with gastric cancer in the Peruvian population. Although our results have shown no association with any analyzed SNP, we are performing functional studies to test differential IL8RB gene expression caused by polymorphisms which shows high different allele frequencies between Native Americans and other populations. Since Native Americans are a neglected group in genomic studies, we suggest that the population from Lima and other large cities from Western South America with high Native American ancestry background may be convenient targets for epidemiological studies focused on this ethnic group.

1120T

Detection of P16 Expression and Human Papillomavirus in Oropharyngeal Squamous Cell Carcinoma. *S. Lai[1,2], V. Sandulache[3,4], J. Zevallos[3,4].* 1) Department of Pathology and Immunology, Baylor College of Medicine, Houston, TX; 2) Department of Pathology, Michael E. DeBakey VA Medical Center, Houston, TX; 3) Department of Otolaryngology, Baylor College of Medicine, Houston, TX; 4) Department of Otolaryngology, Michael E. DeBakey VA Medical Center, Houston, TX.

The association between human papillomavirus (HPV) and head and neck squamous cell carcinoma has been recognized in the last decade. HPV is thought to cause tumor suppressor p16 overexpression by E7-induced inactivation of the retinoblastoma protein. P16 immunoreactivity is often considered as an excellent surrogate marker for HPV infection in head and neck squamous cell carcinoma. However, the incidence of HPV infection in oropharyngeal squamous cell carcinoma of veteran male patients has not been extensively studied. Tumors in oropharyngeal area are difficult to be resected completely, often unresectable due to complexity of location. Study of the pathogenic mechanism of HPV infection, P16 overexpression and patient's clinical outcomes plays an important role for therapeutic implication. We studied P16 expression by immunohistochemistry in 51 oropharyngeal squamous cell carcinomas of veteran male patients. The mean ages at diagnosis are 60.7 year-old. Tumor sites include oropharynx (13 cases), base of tongue (20 cases) and palate (18 cases). Most of the patients underwent chemoradiation therapy without radical resection. P16 immunoreactivity was present in 33 of 51 tumors (64.7%). P16 is diffusely and strongly expressed in 14 of 20 cases (70%) in base of tongue tumor, compared with 2 of 13 cases (15.4%) of oropharynx, and 4 of 18 cases (22.2%) of palate tumor with similar level of expression. Mean follow-up time (months) in P16 negativity, focally expression and diffusely expression are: orophargnx-5, 7, 29 (range from 1 to 36); base of tongue-24, 20, 35 (range from 4 to 73) and palate-14, 21 and 36 months (range from 1 to 73), respectively. HPV DNA will also be detected by polymerase chain reaction and subsequently confirmed by sequencing. The result will be correlated with clinical pathologic data and P16 expression. Our preliminary data suggests squamous cell carcinoma of base of tongue shows frequent diffuse and strong P16 expression. High P16 expression level in oropharyngeal carcinoma predicts a better pathological response to chemoradiation therapy with longer overall survival than do patients with low level. P16 immunoreactivity as well as its associated HPV infection can be used as a prognostic marker for chemoradiation adjuvant therapy.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1121F

BORIS expresion in squamous intraepithelial lesions and cervical cancer. B. *Lazalde[1], N. Velazquez-Hernandez[2], M. Barragán-Hernandez[3], M.A. Reyes-Romero[4].* 1) Department of Genetics, Faculty of Medicine and Nutrition, UJED, Durango, Durango, Mexico; 2) Scientific Research Institute, UJED, Durango, Dgo., Mexico; 3) Cancer Center, General Hospital of the Ministry of Health of Durango State, Durango, Dgo., Mexico; 4) Department of Molecular Medicine, Faculty of Medicine and Nutrition, UJED, Durango, Dgo., México.

The infection by high-risk human papillomavirus is the first step in the development of cervical cancer; however, most infected women do not develop cervical cancer. Different studies supports that histological benign lesions already contain genetic and epigenetic alterations that make them more susceptible to neoplastic transformation. CCCTC-binding factor (CTCF) and its paralogue Brother of the regulator of imprinted sites (BORIS) are involved in epigenetic reprogramming events. BORIS is considered to be a new oncogene, since is abnormally activated in a wide range of human cancer. *BORIS* has been studied, in endometrial and uterine mixed mesodermal tumors and its expression has been observed in a high percentage; however the expression of *BORIS* has not been studied in primary cervical cancer or cervical premalignant lesions. The aim of this study was to analyze the mRNA expression levels of *BORIS* and *CTCF* in cervical smears from women with normal cytology, squamous intraepithelial lesions and cervical cancer. The study was approved by the institutional bioethical committee; informed consent was obtained from each patient before sample collection. Cervical smears were taken and the cyto-brushes were immediately immersed and stored in Trizol Reagent for total RNA isolation. The cDNA synthesis was done with the Quantitect cDNA Kit (Qiagen). Specific *BORIS* and *CTCF* expression was quantitated using a real-time reverse transcriptase PCR (RT_PCR) assay with RT-PCR Kit, and QuantiTect Primer Assay primers (Qiagen Gmbh, Hilden). *GAPDH* was used as a normalizer gene. RT-PCR experiments were performed on Eco thermocycler (Illumina), and the data was processed with the Qiagen REST Software. A total of 55 samples were collected. *BORIS* expression was observed in 1/18 normal samples, 1/12 Low-grade squamous intraepithelial lesions, 1/17 high grade squamous intraepitelial lesion and 3/8 invasive cancer samples. Although, *BORIS* expression was not only observed in cancer, its expression level relative to *CTCF* was higher in the invasive cancer smear samples compared with smears of benign lesions and normal cytology. The *CTCF* expression levels relative to *GAPDH* were significantly downregulated in cancer smear samples compared with smears of benign lesions and normal cytology. In conclusion, these findings suggest that *BORIS* expression could be linked to incipient molecular changes leading to cervical cancer progression, although further studies in this issue are needed.

1122W

An informative clinical applicable serum miRNAs 371–3 and 302/367 test for (germ cell) cancer patients. *L.HJ. Looijenga[1], A.J.M. Gillis[1], M.A. Rijlaarsdam[1], R. Eini[1], L.C.J. Dorssers[1], Y. van der Zwan[1], C.M. de Boer[2], S.J. White[2].* 1) Pathology, Erasmus MC, Rotterdam, Zuid-Holland, Netherlands; 2) Monash Institute, Melbourne, Australia.

Expression of microRNAs (miR) can be specific for malignant cells, as for human germ cell cancers (GCC), seminomas (SE) and nonseminomas (NS). GCC are most in Caucasian young males, show high expression of the miRNA 371–3 and 302/367 clusters, related to the stem cell component. No recurrent DNA copy number changes of these miR and mutation in DICER was found. Current clinically used serum markers for GCC are AFP and hCG (and LDH1), which are informative in a selection of patients (yolk sac tumor and/or choriocarcinoma). Although the fore-mentioned miR have been suggested as GCC serum markers, no robust protocol is available so far. A novel, sensitive and specific pipeline for miRNA profiling is developed based on a magnetic anti-probe bead purification step. This method shows stable recovery based on multiple synthetic non-human spike-ins, allowing calibration. Out of a set of 9 potential serum-specific controls, two were identified by Normfinder and geNorm as most stable (miR-20a and 93), used for normalization. This protocol was applied to serum samples of 12 controls (no cancer), 16 SE and 16 NS patients, taken at time of primary diagnosis. Undetermined values were replaced by Ct maximum per assay +2. Histology, stage of the disease, as well as serum marker measurements, of each patient was known. Patients with a GCC showed significantly higher serum levels of miR-302a/367/371/372/373 ($p<0.006$, Mann-Whitney test, two sided, Bonferroni corrected (MWB)). Pooled analysis, irrespective of histology, showed that miR-302a/367/373 were significantly higher in patients with metastatic - compared to localized disease ($p<0.006$, MWB). Histology specific analysis (SE or EC), demonstrated that miR-367 was significantly higher in patients with metastasized - versus localized SE ($p=0.003$, MWB). The findings indicate that serum miRNA profiling based on a limited set of embryonic miRNAs is informative to diagnose patients with a GCC, better than known serum markers. The method is suitable to define patients with a stage I and higher (metastasized) disease. Defined steps within the protocol for calibration, normalization, and quantification result in a robust method, suitable for clinical implementation. The protocol can be applied to diagnosis and follow up of patients with different types of cancer, or diseases in general, for which specific miRNA profiles are reported, to be detected in various body fluids.

1123T

Gene Variants in Chemokine and Chemokine Receptor with Risk of Prostate Cancer in North Indian Cohort. *R.D. Mittal, R.K Mandal.* Urology & Renal Transplantation, Sanjay Gandhi PGI, Lucknow, Uttar Pradesh, India.

Introduction: The wide and differential distribution of chemokines and their receptors account for the pleiotropic actions of chemokines in prostate cancer (PCa). Allelic variants of the genes involved in inflammatory pathways are logical candidates as genetic determinants for PCa risk. In a case control study we analyzed the frequencies of wild, heterozygous and variant genotypes of the CCL2 Ins/Del, -2518 (A>G) and CCR2 V64I (G>A) in PCa from North Indian cohort. Methodology: The hospital-based case-control study constituted of histologically confirmed 195 PCa patients and 250 healthy controls that were age matched and were of similar ethnicity. Genotypes were determined by PCR-RFLP (Polymerase Chain Reaction Restriction Fragment Length Polymorphism) method Results: Heterozygous genotype I/D (OR=1.71; p=0.010) and variant allele carrier (ID+DD) (OR=1.67; p=0.010) of CCL2 Ins/Del polymorphism demonstrated increased risk of PCa. However, -2518 (A>G) and CCR2 V64I (G>A) were not significantly associated with PCa. We observed significantly enhanced risk for PCa due to interaction between CCL2 I/D, -2518 (A>G) and CCR2 (G>A) genotypes. High risk to PCa was also observed with respect to diplotypes, I-G (OR= 1.83; Bonferroni corrected p value (Pc)=0.004) and D-A (OR=2.11; Pc= 0.004) of CCL2 I/D and -2518 (A>G). Association of genotypes with clinical characteristics, homozygous D/D (OR=7.40; Pc=0.042) and variant allele carrier ID+DD (OR=2.42; Pc=0.036) genotype of CCL2 Ins/Del polymorphism conferred risk of PCa in high Gleason grade tumor. However, smoking did not further modulate the risk for PCa. Conclusion: CCL2 Ins/Del gene polymorphism demonstrated increased risk of PCa. Further studies evaluating the role of these polymorphisms in ethnically diverse populations and a larger cohort may help in understanding the etiopathogenesis of PCa Acknowledgment: UPCST New Delhi, India.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1124F

The regulatory BCL2 promoter polymorphism (-938C,A) is associated in Iranian women Breast Cancer Patients. *B. Motahari[1,4], M. Ghaffarpour[2,3,4], GH. Javadi[1], M. Houshmand[2,4]*. 1) Department of Biology,Science & Research Branch,Islamic Azad University, Tehran, Iran; 2) Medical Genetics Dep,National Genetics Engeeniering and Biotechnology(NIGEB),-Tehran, Iran; 3) Iranian Research Organization for Science and Technology, Tehran, Iran; 4) Medical Genetics Dep, Special Medical centers, Tehran, Iran.

Introduction: Despite recent advancements in the treatment and management of cancers, Breast cancer still remains the most common malignancy and second leading cause of cancer-related deaths among women in Iran. Apoptosis and cellular proliferation play an important role during normal mammary development and carcinogenesis of the mammary gland. In normal tissues, the homeostasis between apoptosis and cell proliferation is regulated by Bcl-2 family proteins. Tumor cells tend to interfere with the mechanism of Apoptosis by activating anti apoptotic genes such as bcl-2. BCL-2 gene, located at 18q21.3 and consists of three exons and two promoters (P1 and P2), which have different functions. The second promoter, P2, is located 1,400 bp upstream of the translation initiation site and functions as a negative regulatory element to the P1 promoter.Bcl-2 is the one of the most important anti apoptotic genes.Bcl2 gene has been demonstrated with breast cancer development and a single nucleotide polymorphism (SNP-938C >A) has been identified recently and can provide a susceptible factor for trigger breast cancer. The aim of this study was evaluation of the frequency of the regulatory BCL2 promoter polymorphism (-938C>A) in Iranian women patients with Breast cancer Materials and Methods: Patients; 27 paired fresh tumor and adjacent normal samples were obtained from consecutive patients with BC who underwent surgery for mama mastroctomy from IRANIAN National Tumor Bank, National Cancer Institute, Imam Khomeini Hospital Complexes, Medical Tehran University, Tehran, Iran.Histopathological examinations were performed, and all tumors were confirmed as adenocarcinoma. Mutational analysis: fresh tumors and their adjacent were extracted for genomic DNA using the QIAamp Mini Kit according to the manufacturer's instructions. We searched DNA samples of Iranian patients for identification of SNP; -938C>A in the inhibitory P2 promoter of the BCL2 gene by means of PCR sequencing. Results and conclusions: we found heterozygosity in this location (-938 C, A) in 16/27 (59.25%) cases with breast cancer and 5 of 27(18.51%) SNP-938 C >A. According to the findings, the point mutation in -938C allele displayed a positive relation with increasing tumorgenesis in breast cancer. .Our result indicated that the regulatory BCL2 promoter polymorphism (-938C,A) may be useful as a tumor marker for increasing tumorgenesis in women breast cancer.

1125W

Type II transmembrane serine protease gene variants associate with breast cancer. *K.M. Nieminen[1, 2, 3], J.M. Hartikainen[1, 2, 3], V. Kataja[4, 5], V.-M. Kosma[1, 2, 3], A. Mannermaa[1, 2, 3]*. 1) Institute of Clinical Medicine, Pathology and Forensic Medicine, University of Eastern Finland, Kuopio, Kuopio, Finland; 2) Biocenter Kuopio and Cancer Center of Eastern Finland, University of Eastern Finland, Kuopio, Finland; 3) Imaging center, Clinical Pathology, Kuopio University Hospital, Kuopio, Finland; 4) Institute of Clinical Medicine, Oncology, University of Eastern Finland, Kuopio, Finland; 5) Cancer center, Kuopio University Hospital, Kuopio, Finland.

Type II transmembrane serine proteases (TTSP) are related to tumor growth, invasion, and metastasis in cancer. Genetic variants in these genes may alter their function leading to cancer onset and progression, and also the outcome of the patients. In this study 464 invasive breast cancer cases and 370 healthy controls from Eastern Finnish origin in the Kuopio Breast Cancer Project (KBCP) were genotyped for 82 SNPs in five different TTSPs, and in three related genes. Association of the genotypes was estimated against breast cancer risk (Armitage trend test), breast cancer specific survival (COX regression analysis), and clinicopathological variables (logistic regression analysis). Five SNPs, *TMPRSS2* SNPs rs468811 and rs7275220, *TMPRSS3* SNPs rs3814903 and rs11203200, and *HGF* SNP rs5754752, associated significantly with breast cancer risk. In addition, *TMPRSS1* SNPs rs12151195 and rs12461158, *TMPRSS2* SNP rs2276205, *TMPRSS3* SNP rs3814903, and *HGF* SNP rs2399403, associated with prognosis (P-values 0.009–0.050). We assessed a new variant by summing the number of risk alleles from significantly associated SNPs, and estimated the additive effect of the risk alleles. As a result, the risk of breast cancer was significantly higher when having five to seven risk alleles (P=0.0020; OR, 2.632; 95% CI, 1.426–4.859) compared to 0–2 alleles. The additive effect of the associating variants was also seen in survival analysis. Women with six to eight survival associating alleles had 3.3 times higher risk of dying of breast cancer compared to women with one to three alleles (P=0.0012; RR, 3.365; 95% CI, 1.612–7.024). Our results suggest this gene family to be strongly associated with breast cancer risk and survival. More importantly, the number of carried risk alleles seems to be important when estimating the breast cancer risk and the prognosis of the patient.

1126T

Prevalence and prognostic impact of Ikaros deletions in childhood acute lymphoblastic leukemia treated according to NOPHO protocols. *I. Ofverholm[1], A.N. Tran[1], G. Barbany[1], A. Nordgren[1], M. Heyman[2], M. Nordenskjold[1]*. 1) Department of Molecular Medicine and Surgery, CMM, Karolinska Institute, Stockholm, Sweden; 2) Department of Women's and Children's Health, Karolinska Institute, Stockholm, Sweden.

Survival rates for childhood acute lymphoblastic leukemia (ALL) have increased dramatically over the last two decades; still 20–25% of the children relapse and 15% will not survive. Deletions of the *IKZF1* locus, coding for the Ikaros transcription factor, have been described as characteristic of Philadelphia-positive ALL, but have also been proposed to predict relapse in other types of B-cell progenitor ALL (BCP ALL). Gene-expression studies have revealed that heterozygous deletion of *IKZF1* results in haploinsufficiency while deletion of a subset of exons (4–7) results in a dominant-negative protein. The aim of our study is to assess the frequency and prognostic impact of *IKZF1* mutations in a cohort of 87 children diagnosed with Philadelphia-negative BCP ALL from 2001 to 2008 at the Karolinska Hospital and treated according to the NOPHO 2000 protocol. We have used Multiplex Ligation-dependent Probe Amplification (MLPA) to investigate the presence of *IKZF1* deletions in DNA extracted from bone marrow samples taken at diagnosis. In order to confirm and further characterize the deletions we also performed array CGH in cases with deletions detected by MLPA. Deletions of *IKZF1* where present in 10% (9/87) of cases, out of which 4/9 had deletions of a subset of exons (4–7) and the remaining cases had heterozygous deletions of the whole *IKZF1* gene. Array CGH revealed large deletions of chromosome arm 7p encompassing several genes in a majority of these latter cases. The frequency of *IKZF1* deletions in our material was consistent with previous reports of unselected BCP ALL cohorts, although we detected a lower rate of cases with intragenic deletions resulting in a dominant-negative isoform of Ikaros. There was a good correlation between MLPA and array CGH results, but while MLPA can identify the presence of a deletion in a single exon it is unable to determine the extent of the deletion beyond the specific gene of interest. MLPA also has a limited ability to distinguish the presence of mixed cell populations, which is often the case in hematological malignancies. We are currently evaluating the clinical and prognostic relevance of *IKZF1* deletions, preliminary data reveals that 22% of cases with *IKZF1* deletions suffered from relapse, a high frequency compared to the non-deleted cases in our study. To our knowledge, this is the first study to examine the presence of *IKZF1* deletions in a cohort of pediatric BCP ALL treated according to the NOPHO protocol.

1127F

Genotype proline/proline of TP53 codon 72 polymorphism is enriched in breast cancer. *A. M. Puebla[1], A. Gutiérrez[2,6], A. Ramos[3,6], E. Salas[4], J. M. Castro[5], R. Ramírez[7], M. P. Gallegos[6]*. 1) Laboartorio de Inmunofarmacologia, UdeG, Guadalajara, jalisco, Mexico; 2) Licenciatura en Químico Farmacobiologo, CUCEI, UdeG; 3) Doctorado de Farmacología, UdeG; 4) Jefatura de Oncología, UMAE, Hospital de Gineco-Obtetricia, CMNO, IMSS, Guadalajara; 5) Jefatura de Onco-Cirugía, UMAE, Hospital de Especialidades, CMNO, IMSS; 6) División de Médicina Molecular, CIBO, IMSS, Guadalajara, Jalisco, México; 7) Doctorado en Genética Humana, UdeG.

The nuclear phosphoprotein p53 encodes to TP53 tumor suppressor gene, which plays an important role in cell cycle, apoptosis, DNA reparations and angiogenesis. The TP53 contains common genetics polymorphisms that influence gene activity. In the present study we describe to frequency of codon 72 polymorphism of TP53 gene in 253 breast cancer patients and 232 woman health controls from Mexico. The frequency observed was of 34% (87/258) and 62% (142/232) of arginine/arginine genotype; 49% (124/253) and 30% (71/232) of heterozygous (arginine/proline) genotype; 17% (42/253) and 8% (19/232) of proline/proline genotype in cases and controls respectively. The statistically different ($p<0.05$) was observed in the proline/proline genotype with an OR of 2.2 (95%CI 1.2–4.0). These results suggest that the proline/proline genotype of codon 72 polymorphism of TP53 is associated with Mexican breast cancer patients.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1128W

3R/3R genotype of TYMS gene is associated with non response of chemotherapy in breast cancer patients. *A. Ramos[1,6], E. Salas[2], J. M. Castro[3], L. E. Figuera[4], A. M. Puebla[5], M. P. Gallegos[6].* 1) Doctorado en Farmacología, UdeG, Guadalajara, Mexico; 2) Jefatura de Oncología, UMAE, Hospital de Gineco-Obtetricia, CMNO, IMSS, Guadalajara; 3) Jefatura de Onco-Cirugía, UMAE, Hospital de Especialidades, CMNO, IMSS; 4) División de Genética, CIBO, IMSS; 5) Laboratorio de Inmunofarmacologia, CUCEI, UdeG; 6) División de Médicina Molecular, CIBO, IMSS, Guadalajara, Jalisco, México.

Thymidylate synthase (TYMS) catalyzes the reductive methylation of dUMP to form dTMP, which is essential for cell proliferation. Has been show that patients high levels of TYMS are not respond to 5-fluorouracil. The polymorphisms 2R/3R located in 5'untranslated region (5'UTR) of TYMS gene, contains 28 bp repeated in the sequence. The alleles with two repeat (2R), have low translational efficiency of TYMS mRNA than three (3R) repeats with greater translational efficiency. In the present study we describe association of genotype 3R/3R polymorphism of TYMS gene with non response in breast cancer patients from Mexico. In 175 woman with breast cancer was observed a frequency of 61% (19/31) and 39% (70/251) of 2R/2R genotype; 63% (61/97) and 49% (36/97) of 2R/3R genotype, 38% (18/47) and 62% (29/47) of 3R/3R genotype in breast cancer responds and non response of chemotherapy respectively. The statistically different ($p<0.05$) was observed in the 3R/3R genotype with an OR of 2.2 (95%CI 1.2–4.0). These results suggest that the 3R/3R genotype of TYMS polymorphism is associated with non response to chemotherapy in Mexican breast cancer patients.

1129T

Characterization of breast cancer intrinsic subtypes in Mexican women using integrative multiplatform analysis. *C. Rangel-Escareño[1], I. Imaz-Rosshandler[1], S. Muñoz-Montero[1,2], JE. Castillo-Fernandez[1,2], A. Hidalgo-Miranda[3].* 1) Computational Genomics, National Institute of Genomic Medicine, Periferico Sur 4809 Arenal Tepepan, Tlalpan C.P. 14610; 2) Instituto Tecnológico y de Estudios Superiores de Monterrey ITESM Campus Ciudad de México, Biotechnology Program. Calle del Puente #222 Col. Ejidos de Huipulco, Tlalpan C.P. 14380, México D.F; 3) Cancer Genomics Lab, National Institute of Genomic Medicine, Periferico Sur 4809 Arenal Tepepan, Tlalpan C.P. 14610.

Breast cancer tumors can be classified according to gene-expression subtypes: Luminal A, Luminal B, Her2-enriched, and basal-like. These intrinsic subtypes differ in genomic complexity, key genetic alterations, and clinical prognosis. For instance, luminal subtypes are associated with expression of estrogen (ER+) and progesterone (PR+) receptors and differentiated luminal epithelial cell markers. Molecular subtype classification not only provide information about prognosis but also may become useful in planning treatment and developing of new therapies. Previous studies have shown a good performance classifying tumors based on gene expression profiles such as MammaPrint, OncotypeDX, 3-gene model by Haibe-Kains(2011) and PAM50. Nevertheless, these models do not always agree and have been extensively tested in mostly Caucasian populations. We tested these four approaches on a set of 106 tumors from Mexican women and found some inconsistencies. The molecular classifiers do not perform as good as they do in non-Mexican samples. To explain why this might be happening we included other sources of biological information in the form of data generated using high-throughput genomic technologies on the same set of patients. In this work we present our findings when we integrate information about gene expression transcripts on human whole-transcript microarrays (Human Gene ST 1.0, Affymetrix, Santa Clara CA) of 141 Mexican samples included 35 normal and 106 tumors, copy number variation using Affymetrix SNP6.0, and somatic mutations of 100X sequencing depth of whole-exome sequencing of 56 tumors on the Illumina sequencing platform. All samples were obtained under institutional IRB approval and with documented informed consent.

1130F

International collaborative analysis of the InSiGHT database quantifies unique reports and redefines the V.U.S. challenge for DNA mismatch repair genes (MMR). *R.P. RAVAL[1,2], L. BAEZ-CABRERA[1,2], JP. PLAZZER[3], S. OKOCHI[1,2], F. MACRAE[3], M. GENUARDI[4], TK. WEBER[1,2].* 1) DEPARTMENT OF SURGERY, VA NEW YORK HARBOR MEDICAL CENTER, BROOKLYN, NY; 2) SUNY DOWNSTATE MEDICAL CENTER, BROOKLYN, NY, USA; 3) DEPARTMENT OF COLORECTAL MEDICINE AND GENETICS, ROYAL MELBOURNE HOSPITAL, MELBOURNE, AUSTRALIA; 4) UNIVERSITY OF FLORENCE, ITALY.

Background: The International Society for Hereditary Gastrointestinal Tumors (InSiGHT) (www.insight-group.org) has collected MMR gene sequence variant reports from clinicians and scientists on its database (DB) since 1994. In 2009 InSiGHT, in collaboration with the Human Variome Project significantly expanded its MMR gene DB using the Leiden Open Variation Database platform (www.lovd.org). However, literature reviews, including rigorous clarification of sequence variant “pathogenicity” has been limited. We also provide a systematic analysis of the heterogeneous literature cited for many alterations and the limitations of current search and analysis tools. Methods:We performed a detailed analysis of the 12,673 DNA MMR variants reported on the InSiGHT DB as of 2012 including frequency of alteration histograms organized by gene, exon, codon and specific nucleotide change. Our report focused on analysis of unique alterations and reviewed the conclusions of all literature cited for the 2,240 unique MSH2 and MLH1 alterations listed. Variant pathogenicity assignments were based on the LOVD legend. Results:Our analysis demonstrates a dramatic increase in the total number of MMR variant reports from 884 in 2008 to 12,673 as of May 2012. For MLH 1, 1150 unique alterations (UAs) were cited of which 168 were classified as “probably pathogenic” based on +/? literature reports. For MSH2, there were 1090 UAs cited of which 108 were classified as “probably pathogenic” having only +/? literature reports. Of the 10,601 MSH2 and MLH1 reports cited only 3 were reported as +/+ and interpreted as “pathogenic” Importantly, 50% (996) of the MSH2 and MLH1 UAs were listed as “variants of uncertain significance” (?/? reports only). 34% of the combined MLH1 and MSH2 UAs listed were associated with multiple reports with differing conclusions (heterogeneous literature) regarding pathogenicity for the exact same variant. Conclusions:Total reports to the InSiGHT LSDB have increased 14 fold since 2009. Detailed analysis of literature cited for each of the 2,240 unique MSH2 and MLH1 alterations reported indicates the clinical significance of 84% of these variants remains uncertain. These results suggest a quantitative analysis/index of this heterogeneous literature would assist their characterization and facilitate prioritization of further analysis. Our collective results expand the dimensions of the V.U.S. problem and underscore the clinical genetic relevance of the HVP mission.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1131W

Frequency of non-DNA-binding Ikaros isoforms in childhood acute lymphoblastic leukemia and relapse or death risk. *A. Reyes-León[1], R. Juárez-Velázquez[1], A. Medrano-Hernández[1], T. Cuenca-Roldán[1], C. Salas-Labadía[1], P. Navarrete-Meneses[1], R. Paredes-Aguilera[2], R. Rivera-Luna[3], R. Cárdenas-Cardos[3], G. López-Hernández[4], R. Bernaldez[5], P. Pérez-Vera[1].* 1) Laboratorio de Cultivo de Tejidos, Instituto Nacional de Pediatria, Mexico city, Mexico; 2) Servicio de Hematología, Instituto Nacional de Pediatría, Mexico city, Mexico; 3) Servicio de Oncología, Instituto Nacional de Pediatría, Mexico city, Mexico; 4) Hospital del Niño Poblano, Puebla, Mexico; 5) Servicio de Hematología, Hospital de Pediatría, Centro Medico Nacional siglo-XXI, IMSS, Mexico city, Mexico.

Introduction: Expression of non-DNA-binding Ikaros isoforms (Ik-6 and Ik-8) is strongly associated with the development of childhood acute lymphoblastic leukemia (ALL). Nowadays, the presence of the non-functional isoforms of Ikaros, Ik-6 and Ik-8, in ALL patients is related to -high risk of relapse and death. The purpose of this study is to determine the Ik-6 and Ik-8 Ikaros isoforms in ALL children, in order to associate them with the presence of adverse events. Methods: The sample included 76 children with newly diagnosed ALL. Informed consent was obtained from parents or tutors. Mononuclear cells were isolated from bone marrow for total RNA extraction. Ikaros isoforms were analyzed by nested RT-PCR and gel electrophoresis. Bands were analyzed based on their size and classified according to reported Ikaros isoforms (Ik-1, Ik-3A, Ik-2, Ik-3, Ik-4, Ik-2A, Ik-5/7, Ik-4A, Ik-8 and Ik-6). Sequencing analysis was performed in non-expected bands. Results: Sixty-nine patients were diagnosed with B-ALL and 21 relapsed or dead. The non-DNA binding isoforms Ik-6, Ik-8 and a deleted-variant of Ik-8, were detected in 18/69 (26.1%) cases. Fourteen B-ALL patients expressed Ik-6, 3 presented Ik-8 and 1 showed Ik-8 variant, 7 of them developed adverse events (3 relapses and 4 deaths); it is remarkable that only 5 of the 7 presented standard risk by conventional prognostic factors at diagnosis. The remaining 7 patients were diagnosed with T-ALL, 3 of them developed adverse events. Ik-8 was the only non-functional isoform detected in this group and it was expressed in 3 (42.8%) patients; 2 of them relapsed. Two B-ALL patients were BCR-ABL1(+), however they did not present non-functional isoforms. Conclusion: The Ik-6 and/or Ik-8 were found in relapsed/death ALL-children in the analyzed group; however the increase of the studied sample and a long-term survival analysis is required to confirm this observation. In addition, unexpected bands must be analyzed by sequencing looking for non-functional isoforms, which could be relevant in the patient's prognosis.

1132T

GENE EXPRESSION PROFILE OF HUMAN TELOMERE AND TELOMERASE COMPLEX IN GASTRIC CANCER. *L.C. SANTOS[1], F. WISNIESKI[1], D.Q. CALCAGNO[1], M.F. LEAL[1], T.B. PONTES[1], C.O. GIGEK[1], E.S. CHEN[1], S. DEMACHKI[2], P.P. ASSUMPÇÃO[2], R.R. BURBANO[3], M.A.C. SMITH[1].* 1) MORFOLOGY AND GENETIC, UNIFESP, SÃO PAULO, BRAZIL; 2) Serviço de Cirurgia, Hospital Universitário João de Barros Barreto, ufpa, Belém, Brasil; 3) Laboratório de Citogenética Humana, Instituto de Ciências Biológicas, UFPA, Belém, Brasil.

Gastric cancer is the fourth most prevalent and the second cause of cancer death worldwide. Changes in gene expression of telomere and telomerase complex have been observed in carcinogenic processes. The investigation of these genetic mechanisms may help to determine more accurate diagnosis, prognosis and the establishment of a therapeutic approach. In the present study, we aimed to identify differentially expressed genes of telomere and telomerase complex in gastric tumors compared to matched noncancerous gastric samples. The expression of 84 key genes involved in telomere replication and maintenance was evaluated using the Human Telomeres & Telomerase RT^2 Profiler™ PCR Array (SABiosciences, PAHS-010). We analyzed six paired gastric cancer and adjacent noncancerous gastric tissues from individuals of Northern Brazil. We observed that ACD, BCL2 and SNRPN expression were downregulated in neoplasic samples compared to matched noncancerous (0.34, 0.50 and 0.47 -fold change; and p = 0.005, p = 0.04, p = 0.03; respectively). ACD gene encodes a protein that is involved in telomere maintenance. This protein is one of six core proteins of shelterin telomeric complex. The BCL2 gene encodes a protein that has a role in the control of apoptosis and inhibition of the telomerase action. SNRPN encodes a protein that acts in pre-mRNA processing and in telomerase regulation. These differentially expressed genes highlight the importance of telomere and telomerase complex in gastric carcinogenesis. These results contribute to the understanding of gastric cancer biology with possible future clinical implications to be tested in a subsequent study in a larger number of gastric cancer patients.

1133F

Chemotherapy response in ovarian cancer is modulated by the DNA encoded microRNA family miR-17–92a. *I. Shapira[1,2], A. Lee[1,3], M. Oswald[1,3], J. Lovecchio[1], A. Menzin[1], J. Whyte[1], L. Dos Santos[1], K. Sultan[1], V. John[1], K. Cheng[1,2], S. Liang[1], T. Bradley[1,2], D. Budman[1,2].* 1) Hofstra North Shore LIJ School of Medicine, Hempstead NY, NY; 2) Monter Cancer Center Lake Success NY, NY; 3) Feinstein Institute for Medical Research Manhasset, NY.

Background Over 75% of women are diagnosed with ovarian cancer are diagnosed at stage III due to inadequate screening techniques. One fifth are cured of their disease with surgery followed by chemotherapy. Patterns of circulating miRs may explain the role of surgery and chemotherapy in ovarian cancer cure. MiRs-17-92 have essential roles in embryonal and immune system development. MiR-17-92 cluster is deleted in 20% of ovarian cancers. Aims We set out to identify changes microRNA patterns in plasma obtained before surgery for ovarian cancer and during chemotherapy that correlated with the patients' long-term outcomes. Methods Between 2004 and 2011 we investigated patterns of plasma miRNAs collected before, after surgery, during and after chemotherapy in 50 patients presenting for surgery for ovarian cancer and 10 normal controls. We also collected blood at ovarian cancer relapse and tumor and benign ovary for miRNA analysis. 2-sample t-test was used for all 2-sample comparison and ANOVA followed by Benjamin-Hochberg method for multiple testing to limit the false discovery rate (FDR) at the 5% level. All tests were 2-tailed and results with a $p<0.05$ were considered statistically significant. Results Fifty patients were operated for EOC mean age at surgery 65 (range 51–78), 11 race matched women consented as controls - mean age of 58 (range 25–67). Four patients were cured, 21 patients died within 36 month of their diagnosis and 21 patients who survived long term not cured but continue to receive chemotherapy. MiR-19a was high in plasma obtained before surgery in patients with short over all survival ($p< 0.003$). During chemotherapy miR-19a decreased 38-fold compared to the pre-surgical levels ($p<0.0002$). MiR-19b was 83 fold higher in the pre-surgical plasma and it decreased 38 fold during chemotherapy ($p<0.003$). MiR-92 was 12 fold higher in the pre-surgical plasma and decreased 17-fold during chemotherapy ($p<0.003$). MiR-25 was 71-fold higher in the pre-surgical plasma and decreased 47-fold during chemotherapy ($p<0.002$). Conclusions During chemotherapy, levels of miR-17-92 change dramatically compared to the pre-surgical levels. Shifts in microRNAs from the cluster 17-92 observed in the nadir phase of adjuvant chemotherapy for ovarian cancer patients correlate with long term outcomes. Patients with normal miR-17-92 levels pre-surgical plasma levels tended to have better long-term survival.

1134W

Single-nucleotide polymorphisms in genes encoding toll-like receptor -2, -3, -4, and -9 in case-control study with bladder cancer susceptibility in North Indian Population. *V. Singh[1], N. Srivastava[2], R. Kapoor[1], R.D Mittal[1].* 1) Department of Urology and Renal Transplantation Sanjay Gandhi Post Graduate Institute of Medical Science Raebareli Road, Lucknow, Uttar Pradesh, India; 2) Department of Physiology, Chhatrapati Shahuji Maharaj Medical University, Lucknow, Uttar Pradesh, India.

Background: Impairment of the immune system could contribute to bladder cancer risk as Toll like receptors are important for innate immunity. Aim of the study: we examined the association between candidate disease-susceptibility polymorphisms in the single nucleotide polymorphism (SNPs) like TLR2 (-196 to-174del), TLR3 (C1377T), TLR4 (Thr399Ile) and TLR9 (G2848A) genes in bladder cancer patients of North Indian population. Methods: The SNPs comprised of TLR 2 (-196 to -174 Del), TLR 3(C1377T), TLR4 (Thr399Ile) and TLR9 (G2848A) genes. The allelic and genotypic frequencies of these TLR SNPs from histopathologically confirmed patients of bladder cancer (n=200) and age and sex matched unrelated, healthy controls of similar ethnicity (n=200) were genotyped by Polymerase Chain Reaction and Restriction Fragment Length Polymorphism (PCR-RFLP) analysis. Result: In TLR2 I/D gene polymorphism, combination of ID+DD showed significant 3 fold increased risk P=0.001. TLR2 with combination of heterozygous and variant genotype (ID+DD) also showed 5 fold risks with tumor stage/grade of BC patients. The other genotypes of TLR 3, 4 and 9 did not exhibit any significant association with bladder cancer risk. Conclusion: Our results suggested TLR2 (-196 to-174 del) to be involved in BC susceptibility, however, TLR 3, 4 and 9 genes were not associated with BC risk implicating that polymorphisms in these tested TLR genes are not likely to be associated with increased risk for developing bladder cancer. Functional studies in ethnically diverse populations may provide a more comprehensive involvement of innate immunity in identifying the disease-associated variants for BC aetiology.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1135T

Clinical impact of MMP-3 and TIMP-3 gene polymorphisms in prostate cancer. *P. Srivastava, R. Kapoor, R.D. Mittal.* Urology and Renal Transplantation, Sanjay Gandhi Postgraduate Institute of Medical Sciences, Lucknow, Uttar Pradesh, India.

Objective: Matrix metalloproteinases (MMPs) have been implicated in progression and metastases of different tumors. The balance between the MMPs and their natural inhibitors (tissue inhibitors of matrix metalloproteinases; TIMP) seems to be an important factor related to this role. The purpose of this study was to evaluate polymorphisms in the MMP-3 and TIMP-3 genes for their associations with prostate cancer (PCa) risk in North Indians. Materials and Methods: Genotypes were determined by PCR-RFLP (Polymerase Chain Reaction Restriction Fragment Length Polymorphism) method in 150 PCa patients and 200 age-ethnicity matched controls. Results: We found significant association in the MMP-3(1171)5A/6A and TIMP-3 (1298) C/T polymorphism with PCa risk. Variant genotype (5A/5A) of MMP-3(1171)5A/6A polymorphism had a high risk of developing PCa (p= 0.037, OR=3.52, 95%CI=1.08–11.5). Individuals with TIMP-3 (1298) CT genotype as well as T allele showed reduced risk of PCa (p<0.001; OR= 0.31; 95%CI= 0.18-0.52, and p=0.001; OR=0.49; 95%CI=0.32–0.75). This effect was even more evident in case of T allele carrier (CT+TT) (p<0.001; OR=0.36; 95%CI=0.22–0.59). Overall no significant association was observed statistically in MMP-3 and TIMP-3 with any of the grading stages and smoking habits in PCa. Haplotype analysis of MMP-3 showed that A-5A-A was associated with three folds (OR=3.06; 95%CI=1.71–5.47; p<0.001) increased risk in PCa patients. Conclusion: This is the first reported association of polymorphisms in the MMP-3 and TIMP-3 gene and PCa risk and supports the hypothesis that the protease/antiprotease balance has an important role in this disease. We propose that analysis of these genes polymorphism may help in identifying patient subgroups at high risk and poor disease outcome. Keywords: Matrix metalloproteinases, Tissue Inhibitors of metalloproteinases, Polymorphism, Prostate cancer, PCR-RFLP, Haplotype.

1136F

Role of SOCS3 Promoter Methylation in The Pathogenesis of Myeloproliferative Neoplasms and Secondary/Reactive Erythrocytosis/Thrombocythemia. *D. TORUN[1], O. NEVRUZ[2], M. AKYOL[3], S. KOZAN[1], M. BAHCE[1], S. GURAN[4], C. BEYAN[2].* 1) Dept. of Medical Genetics, Gulhane Military Medical Faculty, Ankara, Turkey; 2) Dept. of Haematology, Gulhane Military Medical Faculty, Ankara, Turkey; 3) Dept. of Biostatistics, Gulhane Military Medical Faculty, Ankara, Turkey; 4) Dept. of Medical Biology, Gulhane Military Medical Faculty, Ankara, Turkey.

Myeloproliferative neoplasms (MPN) like essential thrombocythemia (ET), polyctyhemia vera (PV) and primary myelofibrosis (PMF) are acquired clonal hematopoietic stem cell disorders and originate from a multipotent hematopoietic stem cell. MPNs demonstrate phenotypic/genotypic similarity with each other. The diagnosis of MPNs is made by step by step elimination of the other clinical disorders. The lack of specific molecular marker makes difficult the differential diagnosis. Hitherto, a few genetic aberrations have been shown which are related with MPNs. Single nucleotide polymorphisms at various loci and somatic mutations, such as those in JAK2 V617F, MPL W515L/K, exon12 JAK2 may contribute to the pathogenesis of MPNs. JAK2 V617F mutation is responsible for most of the PV, ET and PMF patients. Despite the high prevalence of JAK2 V617F mutation, several question still remains. For example, how do some patients develop MPNs in tha absence of JAK2 V617F mutation? SOCS1 and SOCS3 genes are negative regulators of JAK/STAT signal pathway. Promoter methylation of these genes may deteriorate the inhibitor effect and may cause the development of MPNs. For this purpose, it was aimed to investigate the promoter methylation of SOCS1 and SOCS3 genes with methylation specific polymerase chain reaction (MS-PCR). No disease-specific CpG island methylation of SOCS1 was observed. Hypermethylation of the SOCS3 promoter was identified in 5 of 19 (26,3%) PV, 2 of 21 (9,5%) ET, 1 of 5 (20%) PMF and 9 of 42 (21,4%) secondary/reactive erythrocytosis/thrombositosis. As a result of this study, promoter methylation of SOCS3 gene suggesting a possible role for SOCS3 methylation in the pathogenesis of MPNs and secondary/reactive erythrocytosis/thrombositosis.

1137W

Association of common polymorphisms in TNF-α pathway genes and apoptotic genes with risk and prognosis of Esophageal cancer. *M. Umar[1], R. Upadhyay[1], S. Kumar[2], U.C. Ghoshal[3], B. Mittal[1].* 1) Department of Genetics, Sanjay Gandhi Post-Graduate Institute of Medical Sciences, Lucknow, India; 2) Department of Radiotherapy, Sanjay Gandhi Post-Graduate Institute of Medical Sciences, Lucknow, India; 3) Department of Gastroenterology, Sanjay Gandhi Post-Graduate Institute of Medical Sciences, Lucknow, India.

Introduction: Esophageal cancer (EC) is highly aggressive malignancy, involves excessive cellular proliferation and low apoptotic activity. As, TNF-α mediated activation of NF kappa beta (NFKB) triggers cellular proliferation, while p53/its homologue p73 initiate apoptosis in response to irreparable DNA damage, so variation in TNF-α pathway genes and apoptotic genes may affect EC risk and prognosis. Thus, the present study was designed to investigate association of common variants of TNF-α pathway genes and apoptotic genes with susceptibility to EC or its clinical phenotypes and their role in survival outcome of EC patients. Methodology: We genotyped 250 EC cases (including 143 followed up cases) and 250 healthy controls to analyze the polymorphisms in TNF-α pathway genes (TNF-α-308 G>A, NFKB1-94ins/del and NFKBIA-881A>G polymorphisms) and in apoptotic genes (p73 G4C14>A4T14, p53 intron3 16bp ins/del, CASP8 -678 del and CASP8 IVS12-19G>A) using PCR RFLP and TaqMan assay. Results: Individuals with TNF-α-308 GA genotype were found to be at 1.9 fold higher risk of EC (OR=1.88, 95% CI=1.17–3.05, P=0.010). In case of apoptotic gene polymorphisms, p73 AT/AT genotype in dominant model and CASP8 IVS12-19 AA genotype in recessive model were found to confer increased risk of EC (OR=1.82, 95% CI=1.21–2.73, P=0.004 and OR = 4.08, 95% CI=1.13–14.76, P=0.032 respectively). Stratification of subjects based on clinical phenotypes showed that TNF-α GA, p73 AT carrier and CASP8 AA genotypes specifically modulated risk of squamous cell carcinoma tumor histopathology and lymph node metastasis. Tumor location specific risk was observed with association of TNF-α-308 GA genotype with upper-third, p73 AT carrier genotype with middle-third and CASP8 IVS12-19 GA genotype with lower-third esophageal tumors. Other variants like NFKB1-94ins/del and NFKBIA-881 A>G, p53 intron3 ins/del and CASP8 -678 del polymorphisms did not seem to play any role in susceptibility to EC. Also, no interaction with environmental risk factors was observed. Kaplan Meier and COX regression analysis showed CASP8 IVS12-19 AA and p53 DD genotypes associated with lower median survival and adverse outcome in EC patients treated with radiotherapy with/without chemotherapy. Conclusion: TNF-α-308 G>A, p73 G4C14>A4T14 and CASP8 IVS12-19 polymorphisms modulate risk of EC, while p53 intron3 and CASP8 IVS12-19 polymorphisms affect the survival outcome of EC patients in north Indian population.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1138T

Colon cancer in a 9-year-old female due to combined inherited EPCAM deletion and hMSH2 mis-sense mutation leading to tissue-specific bi-allelic loss of expression of hMSH2. *P. J. Ainsworth[1,2,4], H. H. Li Chang[2], H. Levin[3], D. K. Driman[2], V. M. Siu[3,4], A. E. L. Cairney[3], N. L. Scanlan[5], K. Buckley[6].* 1) Molec Diagnostic Lab, London Hth Sci, London, ON, Canada; 2) Department of Pathology, London Hth Sci/Western Univ, London, ON, Canada; 3) Department of Pediatrics, London Hth Sci/Western Univ, London, ON, Canada; 4) Department of Biochemistry, Western Univ, London, ON, Canada; 5) Cancer Genetics, London Hth Sci, London, ON, Canada; 6) Cancer Genetics, Grand River Hospital, Kitchener, ON, Canada.

Monoallelic mutation of hMSH2 is associated with Lynch Syndrome, and leads to a risk of colonic, as well as extracolonic cancers, which are typically not seen before 30 years of age. Bi-allelic mutations in hMSH2 are associated with a much earlier onset of colon cancer, and can be associated with astrocytoma and features of neurofibromatosis[1]. It has also been reported[2] that heterozygous germline deletions of the last exons of EPCAM can lead to colon cancer associated with aberrant methylation, and consequent inactivation, of the normal downstream *in cis* hMSH2 gene. A 9 year old female (the proband), who presented with severe anemia from gastrointestinal bleeding, demonstrated adenocarcinoma after colorectal biopsy, and underwent total colectomy. Pathologic examination documented 2 mucinous adenocarcinomas, 70–80 adenomas, and lymph node metastases. Immunohistochemistry showed loss of expression of hMSH2 and hMSH6, both within the normal-appearing colonic mucosa as well as in the adenocarcinomas; however, in the tumour-infiltrating lymphocytes, muscle, peripheral nerve, and fibroblast tissues, both of these MMR proteins were intact. Family medical history was notable for multiple relatives with malignancies on both sides of the family, although neither parent nor their siblings had a personal history of cancer. An hMSH2 mis-sense mutation (hMSH2:c.2075G>T, p.Gly692-Val), previously identified in the proband's maternal grandfather's half-sister's daughter, demonstrated immunohistochemical loss of hMSH2 protein in her endometrial tumour tissue. The same mis-sense mutation was also identified in leukocyte-derived genomic DNA from the proband, as well as her mother. In addition the proband's DNA, analyzed by MLPA, showed an EPCAM deletion mutation (EPCAM: EX:8_9del), which was also shown to be present in genomic DNA obtained from her father. This case demonstrates that very early onset colon cancer may also result from a germline mutation in EPCAM, in concert with a germline mutation in hMSH2. Tissue-specific expression of EPCAM can explain the immunohistochemical loss of hMSH2 in the normal-appearing colonic epithelium and adenocarcinoma, but its presence in other cell types. As extracolonic cancers appear to be rare in EPCAM deletion carriers[3], our proband may have a more favourable prognosis compared to individuals with biallelic MMR mutations[1]. 1. Familial Cancer (2009) **8:**187–194 2. Nat Genet (2009) **41(1):**112-7 3. Am J Gastoenterol (2011) **106:**1829-36.

1139F

NF1 Second Hit Leads to the Up-Regulation of HLA Class II Genes and Recruitment of Tregs in Neurofibromas. *E.M. Jouhilahti[1,2], S. Peltonen[3], E.P. Jokinen[1], T. Callens[4], H. Aho[5], E. Legius[6], O. Lassila[2,7], L. Messiaen[4], J. Peltonen[1].* 1) Department of Cell Biology and Anatomy, Institute of Biomedicine, University of Turku, Turku, Finland; 2) Turku Doctoral Programme of Biomedical Sciences, University of Turku, Turku, Finland; 3) Department of Dermatology, University of Turku and Turku University Hospital, Turku, Finland; 4) Department of Genetics, Medical Genomics Laboratory, University of Alabama at Birmingham, Birmingham, Alabama, USA; 5) Department of Pathology, Turku University Hospital, Turku, Finland; 6) Department of Human Genetics, KU Leuven, Leuven, Belgium; 7) Department of Medical Microbiology and Immunology, University of Turku, Turku, Finland.

Background: A germline mutation in the *NF1* tumor suppressor gene invariably manifests as a neurocutaneous-skeletal syndrome neurofibromatosis 1 (NF1). Cutaneous neurofibromas are the hallmark lesions of the syndrome. They are composed of an *NF1* heterozygous tumor microenvironment and a subpopulation of Schwann cells carrying an *NF1* null ($NF1^{-/-}$) genotype, which is a prerequisite for the development of neurofibromas. **Purpose:** The current study investigated the transcriptome of neurofibroma-derived Schwann cell cultures enriched towards $NF1^{+/-}$ or $NF1^{-/-}$ genotype. The aim was to answer to the question "How does the NF1 second hit mutation affect the reading of the whole genome?" **Methods:** Illumina Sentrix® HumanHT-12 v3 Expression BeadChips were used for gene expression profiling of neurofibroma-derived Schwann cell cultures enriched towards $NF1^{+/-}$ or $NF1^{-/-}$ genotype. For microarray analysis, cell cultures were established from a total of nine tumors from three different NF1 patients. The *NF1* mutations were characterized from each established cell line. The microarray results were validated using qPCR, western blots and immunolabelings. Furthermore, *NF1* siRNAs were applied on normal human keratinocytes and keratinocytes from a patient with NF1, $NF1^{+/-}$ Schwann cells, and neurofibroma-derived precursors (NFPs). **Results:** Expression profiling of NF1+/− and NF1−/− cells derived from the same tumor highlighted the processes of inflammatory response, antigen presentation and processing, and myelination. More specifically, the results showed that HLA class II expression was associated with $NF1^{-/-}$ genotype in Schwann cells. When $NF1^{-/-}$ and $NF1^{+/-}$ Schwann cell culture pairs from each tumor were analyzed by western blots, the results showed that HLA-DR protein levels were consistently higher in the $NF1^{-/-}$ Schwann cells. In neurofibroma tissues, HLA-DR was expressed by a subpopulation of cells potentially rendering them as non-professional antigen presenting cells. Our results also demonstrated that the tumors hosted scattered CD4/ FoxP3-positive cells, consistent with Tregs which are known to contribute to the immunologic tolerance through interaction with antigen presenting cells. Thus, the results pinpoint a novel mode of action how a second hit in a tumor suppressor gene helps the mutated cells hide from the immune system. Targeting Tregs in NF1 patients could provide a novel therapeutic approach to interfere with the development of neurofibromas.

1140W

Genetic screening of colorectal cancer patients for the Lynch syndrome detection in Estonian population. *M. Kask[1], E. Oitmaa[1], K. Toome[1], K. Raime[1], K. Vaidla[1], P. Laidre[2], J. Jaal[2], J. Soplepmann[2], V. Afanasjev[3], T. Erm[2], H. Roomere[1].* 1) Asper Biotech, Tartu, Tartumaa, Estonia; 2) Tartu University Hospital, Tartu, Tartumaa, Estonia; 3) The North Estonia Medical Centre, Tallinn, Estonia.

About 700 new colorectal cancer cases are diagnosed in Estonia every year, which brings the incidence of colorectal cancer to the third place of all the cancer cases. Each year about 400 people die of colorectal cancer in Estonia. About 4% of all colorectal cancer cases are associated with dominantly inherited Lynch syndrome caused by germline mutations in MLH1, MSH2, MSH6 or PMS2 genes. Mutation carriers have a lifetime risk of up to 80% for colorectal cancer, 20–60% risk of endometrial cancer, as well as other tumors. Aim of the study was to detect Lynch syndrome in patients diagnosed with colorectal cancer to improve the screening strategy in Estonia and provide genetic testing and counceling to the family members of the patients as a cancer preventive measure. The study involved a systematic analysis of 180 patients diagnosed with colorectal cancer. Microsatellite instability (MSI) testing, BRAF V600E mutation detection, immunohistochemistry (IHC) and mutation analysis of MLH1, MSH2, MSH6, PMS2 genes were performed. MSI analysis was performed on tumors from all the patients using Bethesda panel (BAT-25, BAT-26, D2S123, D5S346, D17S250) and an additional BAT-40 marker. Also, BRAF V600E mutation analysis was performed on tumors from all the patients. Based on revised Bethesda guidelines and the results of MSI (MSI-H or MSI-L) and BRAF (V600E negative) analyses, samples were chosen for the IHC analysis and DNA sequencing of MLH1, MSH2, MSH6 and PMS2. The average age of colorectal patients was 72 (36 to 88) years. MSI-H and BRAF mutation were observed in 30 and 28 out of all cases, respectively. Several polymorphisms in MLH1 (13); MSH2 (11); MSH6 (10) and PMS2 (15) genes, and a few previously not described variants of unknown significance were found.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1141T

CASP8AP2 and H2AFZ expression as a prognosis factor in childhood acute lymphoblastic leukemia. *R. JUAREZ-VELZQUEZ[1], A. REYES-LEON[1], C. SALAS[!], R. PAREDES[2], R. CARDENAS[2], G. LOPEZ-HERNANDEZ[3], A. LOPEZ[4], A. CARNEVALE[5], R. ORTIZ[6], R. BERNALDEZ[7], P. PEREZ-VERA[1].* 1) Human Genetics Dept, Instituto Nacional de Pediatria, Mexico, D.F; 2) Hemato-Oncology Dept, Instituto Nacional de Pediatria, Mexico, D.F; 3) Hematology Dept, Hospital del Niño Poblano, Puebla; 4) Gastroenterology Dept, INCMNSZ, Mexico, D.F; 5) INMEGEN, Mexico, D.F; 6) Universidad Autonoma Metropolitana-Iztapalapa, Mexico, D.F; 7) Hematology Dept, Hospital de Pediatria, CMNSXXI, Mexico, D.F.

Cancer is the second global cause of infant mortality, among them, acute lymphoblastic leukemia (ALL) is the most common type. Up to date, most patients with ALL achieve a cure through the use of adequate treatments according to the risk classification of each patient. Recently, new potential genetic markers have been identified to predict outcome of patients with ALL. The low gene expression levels of CASP8AP2 and H2AFZ had been considered an independent predictor of adverse outcome. In order to know if these genes predict outcome in our population, we studied the association between adverse events (relapse or die) and the expression levels of these genes, in a group of ALL children. Bone marrow samples were collected from 86 ALL pediatric patients diagnosed at Instituto Nacional de Pediatría, Centro Médico Nacional SXXI and Hospital del Niño Poblano in Mexico, between 2008 and 2010. Quantitative RT-PCR was used for evaluating expression levels. Seventy-four patients were diagnosed with B-ALL and 12 with T-ALL. Twenty-eight patients presented adverse events, of them, 15 (54%) and 14 (50%) showed low expression for CASP8AP2 and H2AFZ, respectively. Survival analysis (Kaplan-Meier) revealed an event-free survival (EFS) significantly lower between low expression of CASP8AP2 (log-rank p=0.004) and H2AFZ (log-rank p=0.028). Patients with low expression of both CASP8PA2 and H2AFZ genes presented the lowest EFS (log-rank p<001). These results suggest that expression of CASP8AP2 and H2AFZ genes could be useful as prognostic markers to identify ALL children at risk of relapse.

1142F

Circulating miRNAs as non-invasive blood based diagnostic biomarkers. *R. Duttgupta[1], R. Jiang[1], J. Gollub[1], T. Stamato[3], B.C. Getts[2], K.W. Jones[1].* 1) Assay and Product Developement, Affymetrix, Santa Clara, CA; 2) Research and Development, Genisphere LLC, 2801 Sterling Drive, Hatfield, PA, 19440; 3) Lankenau Institute for Medical Research, Wynnewood, PA 19096.

Effective diagnosis and surveillance of complex multi-factorial disorders such as cancer can be improved by screening of easily accessible biomarkers. Highly stable cell free Circulating Nucleic Acids (CNA) present as both RNA and DNA species have been discovered in the blood and plasma of humans indicating a functional role in relevant biological pathways. Correlation between the deregulation of these pathways and tumor-associated genomic/epigenetic/transcriptional alterations in CNA levels are strong predictors of the utility of this class as promising clinical indicators. Towards this goal microRNAs (miRNAs) representing a class of naturally occurring small non-coding RNAs of 19–25nt in length are important members that can associate specific expression profiles with cancer development. In this work, we discuss some of the pre-analytic considerations for isolating plasma fractions for the study of miRNA biomarkers. The relative proportions of cell-derived and cell-free miRNAs in hematopoietic fractions are outlined in a whole genome manner. Cellular miRNA signatures in cohorts of normal individuals are catalogued and the abundance and gender specific expression of bona fide circulating markers explored after calibrating the signal for this interfering class. A map of differentially expressed profiles is presented and the intrinsic variability of circulating miRNA species investigated in subsets of healthy males and females. Additionally, as proof-of-principle the diagnostic potential of circulating miRNA signatures for non-invasive prediction of Ulcerative Colitis (UC) - a subcategory of Inflammatory Bowel Disease (IBD) is investigated. We identify a panel of 31 platelet-derived miRNAs biomarkers with performance measurements of 92.8% accuracy, 96.2% specificity and 89.5% sensitivity in distinguishing UC patients from normal individuals. The mRNA targets/gene pathways and physical linkage between known IBD susceptibility loci and the qualified biomarker panel is investigated. Prospective cohorts for sub-stratifying IBD spectrum disorders are discussed.

1143W

Comparison of PCR method and RUT test for detection of *Helicobacter pylori* infection in gastric tissue. *N. Bagheri[1], L. Salimzadeh[2], F. Azadegan[3], Gh. Rahimian[4], A. Taghikhani[4], M. Hashemzadeh[5], H. Shirzad[5].* 1) MSc of Immunology, shahrekord, 1., Iran; 2) Department of Immunology, Molecular Research Center, Shahrekord University of Medical Sciences; 3) MSc Genetic, Cellular and Molecular Reserch Center, Shahrekord University of Medical Sciences; 4) Molecular Research Center, Shahrekord University of Medical Sciences; 5) Cell and Molecular Research Center, Shahrekord University of Medical Sciences.

Background and Aim: Culture has been for long the method of choice to detect infectious agents. However Helicobacter pylori (H.pylori) is a fastidious bacteria and growing slowly that its diagnosis by this method may take several days. This study was done to comparatively evaluate PCR and the rapid urease test in order to determine which of target housekeeping genes (glmM and 16SrRNA gene) was more appropriate in the diagnosis of Helicobacter pylori (H pylori) infection. Methods: The specimens used for this study were gastric biopsy samples which were collected from 438 patients undergoing upper gastroduodenal endoscopy for various dyspeptic symptoms. CLO test was compared with the PCR method for detection of selected housekeeping genes of H.Pylori, glmM and 16srRNA. Results: Totally 33.78% of the biopsies were positive for H pylori using both CLO test and PCR method for glmM and 16srRNA reference genes. 43.1% of the biopsies were negative for H pylori in all three mentioned methods. 193 of the 438 study subjects (43.6%) were found positive for CLO test. glmM and 16srRNA were positive in 48.4% and 38.3% of all specimens respectively. In positive CLO test specimens, only 156 and 148 subjects were positive for glmM and 16srRNA gene. Of the 245 negative CLO test subjects, 56 (22%) were positive with glmM gene and 20 (16%) were positive for 16srRNA gene. CLO test gave 7.9% false positive (glmM and 16srRNA Negative) and 4.5% false negative (glmM and 16srRNA Positive) results. Conclusion: Our results indicated that the CLO test gave various false positive and false negative results and gained poorer performance than PCR for detection of H.pylori infection in biopsy of dyspeptic Patients. This shows that in addition to the CLO test, PCR method using specific primers sets of both glmM and 16srRNA is useful for correctly diagnosis of H.pylori infections in laboratory.

1144T

Lowered PTEN protein dosage predicts for underlying germline *PTEN* mutations amongst patients presenting with thyroid cancer and Cowden-like phenotypes. *J. Ngeow[1,2], X. He[1,2], J. Mester[1,2,3,4], J. Lei[1,2], T. Romigh[1,2], M. Milas[3,4], M. Orloff[1,2,3], C. Eng[1,2,3,4,5,6].* 1) Genomic Medicine Institute, Cleveland Clinic, Cleveland, Ohio 44195, USA; 2) Lerner Research Institute, Cleveland Clinic, Cleveland, Ohio 44195, USA; 3) Taussig Cancer Institute, Cleveland Clinic, Cleveland, Ohio 44195, USA; 4) Thyroid Cancer Center, Endocrinology and Metabolism Institute, Cleveland Clinic, Cleveland, Ohio 44195, USA; 5) Stanley Shalom Zielony Institute of Nursing Excellence, Cleveland Clinic, Cleveland, Ohio 44195, USA; 6) Department of Genetics and Genome Science, and CASE Comprehensive Cancer Center, Case Western Reserve University, Cleveland, Ohio 44106, USA.

Statement of Purpose: Thyroid cancer is a major component of Cowden Syndrome (CS), a heritable syndrome also characterized by breast, uterine and kidney cancers. CS patients with an underlying *PTEN* mutation (*PTEN*$^{mut+}$) have a 70-fold increased risk of developing epithelial thyroid cancer. In contrast, <1% of sporadic epithelial thyroid cancer patients carry a germline *PTEN* mutation. Thus, cost-efficient markers capable of shortlisting thyroid cancers for CS genetic testing would be clinically useful. Here, we analyzed the utility of patient blood-PTEN protein levels in predicting germline *PTEN* mutations. **Patients and Methods:** We conducted a 5-year, multicenter prospective study of 2792 CS and CS-like patients, all of whom had comprehensive *PTEN* analysis. Analysis of PTEN and downstream proteins by immunoblotting was performed on total protein lysates from patient-derived lymphoblast lines. We compared PTEN protein expression levels between *PTEN*$^{mut+}$ patients and those with variants of unknown significance (VUS) or wildtype *PTEN* (*PTEN*$^{wt/vus}$). **Results:** Of 2792 CS/CS-like patients, 722 had thyroid cancer. *PTEN* germline pathogenic mutations were present in 36 (5%) patients. *PTEN* frameshift and nonsense mutations accounted for 61% of mutations. Ninety-three percent (26/28) of *PTEN*$^{mut+}$ patients had blood-PTEN protein levels in the lowest quartile as compared to 25% (139/555) of *PTEN*$^{wt/vus}$ patients (p<0.001). Low blood-PTEN levels predicted for *PTEN*$^{mut+}$ cases with a 99.5% NPV (95%CI 98.3- 99.3) and a positive test likelihood ratio of 3.7 (95%CI 3.11–4.42). **Conclusions:** Our study shows that PTEN protein dosage could serve as a screening molecular correlate predicting for germline *PTEN* mutation in CS and CS-like presentations of thyroid cancer.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1145F

Cellular uterine leiomyomata with chromosome 1p deletions represent a distinct entity. *J.C. Hodge[1], A.T. Florin[2,3], K.E. Pearce[1], A.C. Clayton[1], E.A. Stewart[2].* 1) Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN; 2) Department of Obstetrics and Gynecology and Surgery, Mayo Clinic, Rochester, MN; 3) Department of Obstetrics and Gynecology, University of Tübingen, Germany.

Introduction: Cellular leiomyomata (CL) represent a subgroup of uterine leiomyomata (UL) that are distinguished by hypercellularity. Developing evidence suggests that CL may have malignant potential. Our recent case-control analysis of CL versus typical UL demonstrated that CL have a distinct clinical phenotype with characteristics of leiomyosarcomas (LMS). Another study showed a subset of women with CL or atypical leiomyomata have disease-related deaths occurring >6 years after diagnosis. In addition, a limited genetic analysis found a portion of CL have chromosome 1p deletions, and the expression profiles of two such tumors segregated with LMS rather than myometrium/typical UL in a hierarchical cluster analysis. **Methods**: The current study included the same cohort as our case-control analysis of women who had surgery at Mayo Clinic between Jan 1989 and Dec 2008 that were diagnosed with a CL (n=99). Each case with available tissue that was confirmed to meet criteria for CL when reviewed by a single pathologist underwent determination of 1p deletion status using interphase FISH that targeted 1p21.1 plus control probes at 1q21.3 and 1q25. Univariate analysis of the successful cases (n=69) differentiated the clinical characteristics of women with CL that did or did not have 1p deletion. Mean values of continuous variables were compared with t-tests and percentages of categorical variables were contrasted using chi square analysis. **Results**: Twenty-two percent of our CL cohort had a 1p deletion. Women with a CL that had 1p deletion were significantly older (52.4±15.2 vs 43.7±10.5 yrs; P=0.47) and more likely postmenopausal (33.3 vs 9.8 %; P=0.040) with postmenopausal bleeding (20.0 vs 3.92 %; P=0.039). The uteri from these women tended to be significantly larger (1148.3±1494.7 vs 563.4±599.8 g; P=0.026) with a larger dominant fibroid (11.0±5.2 vs 8.1±4.5 %; P=0.027) containing hyaline necrosis (46.7 vs 17.7 %; P=0.037). **Conclusions**: CL with 1p deletion are common and have a distinct clinical phenotype with some characteristics in common with LMS. Such a striking difference in clinical behavior depending on the presence of 1p deletion suggests this molecular signature may help classify delayed malignant potential. Our findings are of particular importance considering the increasing preference for uterine-sparing treatment options, suggesting 1p deletion status in cellular UL may be an important factor in clinical surveillance decisions.

1146W

Cancer Biomarker Research using castPCR™ Technology. *T. Hartshorne, Y. Bao, B. Ching, M. Mouanoutoua, Y. Wang, D. Keys, S. Desai, J. Stevens.* Life Technologies, 850 Lincoln Centre Drive, Foster City, CA, USA.

Cancer biomarkers have applications in the diagnosis, staging, prognosis and monitoring of disease progression, as well as in the predication and monitoring of drug response. Profiling and validation research tools are needed that exhibit the combined features of high sensitivity and high specificity for cancers. However, the sensitivity of molecular methods such as DNA sequencing and conventional genotyping in tumor samples is limited, typically ranging from 5–20%. We have recently developed TaqMan® Mutation Detection Assays using our competitive allele specific TaqMan® PCR (castPCR™) technology for cancer biomarker research. TaqMan® Mutation Detection Assays were tested with >300 tumor research samples (either fresh/frozen or formalin-fixed, paraffin-embedded samples) and cell lines to assess mutation status at multiple independent laboratories. The results showed that castPCR™ technology can robustly detect mutations as low as 0.1% and has >99% concordance to other technologies including PCR-based technology and sequencing. In this study, a large panel of castPCR™ assays for AKT1, BRAF, CTNNB1, HRAS, KRAS, NRAS, PIK3CA, PTEN and TP53 genes were used for investigating somatic mutations in breast tumor research samples. Initially, 4 model FFPE cell lines were used to validate the assays. Mutant gDNAs were titrated into wild type gDNAs from 50% to 0.1%. Mutations were identified down to 0.1% titration with high reproducibility. No false positives were found in non-tumor samples. The results obtained by TaqMan® Mutation Detection Assays for 20 breast tumor samples (FFPE/fresh frozen) were concordant to those reported by other methods. Our data showed that castPCR™ technology provides an excellent tool for identifying cancer biomarkers or confirming potential cancer markers such as those obtained by next-generation sequencing and other technologies.

1147T

Naturally Missing Teeth and Ovarian Cancer: A Potential Genetic Link. *L. Morford[1], A. Vu[2], K. Kirk[1], M. Gilbey[1], G. Falcao-Alencar[1], M. Sakamoto[2], D. Fardo[3], J. Hartsfield, Jr.[1,2].* 1) College of Dentistry, Center for Oral Health Research, Univ Kentucky, Lexington, KY; 2) College of Dentistry, Division of Orthodontics, Univ Kentucky, Lexington, KY; 3) College of Public Health, Department of Biostatistics, Univ Kentucky, Lexington, KY.

Objective: Women diagnosed with Epithelial Ovarian Cancer (EOC) have an 8.1 times increased odds of having 1 to 6 naturally missing teeth (NMT; also termed hypodontia) when compared to controls. This is not trivial, as 20% of the women diagnosed with EOC have hypodontia, compared to only 3% of control subjects. Of the EOC patients with hypodontia, 60% had a family history of hypodontia, while only 30% of the same patients had a family history of ovarian cancer. By comparison, control subjects with hypodontia had no family history of either hypodontia or EOC. Hence, NMT may be good pre-malignancy marker of EOC, especially since hypodontia can be diagnosed at young ages. The aim of this nested case-control study was to determine the association between SNP rs10088218 within an ovarian cancer susceptibility locus on chromosome 8q24 and NMT in a population of young, cancer-free orthodontic patients. Design and Patients: Genomic DNA was isolated from 110 orthodontic patients (30 subjects with NMT and 80 controls). Subjects with NMT were sub-classified according to the number of agenic teeth as having hypodontia (1 to 6 NMT) or oligodontia (7 or more NMT). Underdeveloped peg-shaped teeth were also recorded. Variation at rs10088218 was assayed using Taqman®-based SNP genotyping. Hardy-Weinberg Equilibrium (HWE) testing was employed to assess genotyping quality, and logistic regression was used for association testing. Results: Within this orthodontic patient population, more females had NMT than males (3.3 to 1 ratio); with maxillary lateral incisors being most commonly affected by agenesis and/or peg formation, followed by mandibular second premolars and maxillary secondary premolars. No departures from HWE were detected. After an adjustment to remove any gender bias, rs10088218 was shown to be significantly associated with the presence of hypodontia (p=0.019) and also with the presence of generalized-NMT (hypodontia + oligodontia; p=0.021) under an additive mode of inheritance. Conclusions: Following adjustment for gender, each copy of the G allele at rs10088218 conferred an 11.51-times greater odds of having hypodontia, and a 4.37 greater odds of having generalized-NMT. Future studies will need to examine the potential dual association of rs10088218 with both NMT and ovarian cancer in a single population. It remains to be seen if this particular marker is a better predictor for the development of EOC than the NMT status alone.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1148F

A meta-analysis of genome-wide association studies to identify prostate cancer susceptibility loci associated with aggressive and non-aggressive disease. *A. Amin Al Olama[1], Z. Kote-Jarai[2], F.R. Schumacher[3], F. Wiklund[4], S.I. Berndt[5,6], S. Benlloch[1], G.G. Giles[7,8], G. Severi[7,8], D.E. Neal[9], F.C. Hamdy[10], J.L. Donovan[11], D.J. Hunter[12], B.E. Henderson[3], S. Chanock[6], H. Gronberg[4], C.A. Haiman[3], P. Kraft[12], D.F. Easton[1], R.A. Eeles[2,13], The PRACTICAL Consortium.* 1) Centre for Cancer Genetic Epidemiology, Strangeways Laboratory, Worts Causeway, Cambridge, CB1 8RN, UK; 2) The Institute of Cancer Research, 15 Cotswold Road, Sutton, Surrey, SM2 5NG, UK; 3) Department of Preventive Medicine, Keck School of Medicine, University of Southern California/Norris Comprehensive Cancer Centre, Los Angeles, California, USA; 4) Department of Medical Epidemiology and Biostatistics, Karolinska Institute, SE- 171 77 Stockholm, Sweden; 5) Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, Bethesda, Maryland, 20892; 6) Core Genotyping Facility, SAIC-Frederick, Inc., National Cancer Institute, NIH, Gaithersburg, MD, USA; 7) Cancer Epidemiology Centre, Cancer Council Victoria, 1 Rathdowne street, Carlton Victoria 3053, Australia; 8) Centre for Molecular, Environmental, Genetic and Analytic Epidemiology, The University of Melbourne, 723 Swanston street, Carlton, Victoria 3053, Australia; 9) Surgical Oncology (Uro-Oncology: S4), University of Cambridge, Box 279, Addenbrooke's Hospital, Hills Road, Cambridge, UK and Cancer Research UK Cambridge Research Institute, Li Ka Shing Centre, Cambridge, CB2 2QQ, UK; 10) Nuffield Department of Surgery, University of Oxford, Oxford, UK, Faculty of Medical Science, University of Oxford, John Radcliffe Hospital, Oxford, OX3 9DU, UK; 11) School of Social and Community Medicine, University of Bristol, Canynge Hall, 39 Whatley Road, Bristol, BS8 2PS, UK; 12) Program in Molecular and Genetic Epidemiology, Department of Epidemiology, Harvard School of Pubic Health, Boston, MA 02115, USA; 13) Royal Marsden NHS Foundation Trust, Fulham and Sutton, London and Surrey, UK.

Genome-wide association studies have identified multiple common genetic variants associated with an increased risk of prostate cancer, but these explain less than one third of the heritability. To identify further susceptibility alleles, we conducted a meta-analysis of four genome-wide association studies including 5,953 cases of aggressive prostate cancer and 11,463 controls. We computed association tests for ~2.6M SNPs and followed up the most significant SNPs by genotyping 49,121 samples in 29 studies through the international PRACTICAL and BPC3 consortia. We confirmed the association of a prostate cancer susceptibility locus, rs11672691 on chromosome 19, and demonstrated that the SNP was associated with aggressive prostate cancer(OR=1.12, 95% CI 1.03–1.21; *P*-value= 1.4×10^{-8}), and more weakly with non-aggressive prostate cancer(OR = 1.08, 95%CI 1.05–1.12; *P*-value=8.2×10^{-7}) however the difference was not statistically significant. This work was supported by Cancer Research UK and the NIH GAME-ON Initiative.

1149W

Identifying genetic modifiers in two p16 melanoma pedigrees. *M.H. Bailey[1], C. Teerlink[2], J.M. Farnham[2], L. Cannon-Albright[2].* 1) Life Sciences, Brigham Young University, Provo, UT; 2) Genetic Epidemiology, University of Utah, Salt Lake City, UT.

Introduction/Background: The CDKN2A gene, also known as p16, is a well-documented tumor suppressor gene predisposing to melanoma, and less often to pancreatic cancer. It has not been determined to what extent additional modifier genes/variants in p16 mutation carriers affect risk. Methods: We performed genotyping (500,000 SNPs) in 25 melanoma cases in two-linked high-risk melanoma pedigree with confirmed segregation of p16 mutations. Using genomically matched population controls from the Illumina iControl data set, we performed case/control genome wide association analysis (GWAS) to identify variants modifying melanoma risk in p16 carriers. Most GWAS approaches are not applicable to related individuals, thus we used PLINK solely as a naïve approach to identify possible associations to disease. We then used EMMAX (which adjusts for relatedness by way of a pairwise relationship matrix estimated from genotype data) to identify SNPs associated with melanoma risk in the presence of p16. The intersection of significantly associated SNPs from both analyses were then analyzed with PedGenie, an association analysis software that appropriately adjusts for know relationships among cases using 100,000,000 Monte Carlo inheritance simulations. Results: 7 SNPs were identified to have genome wide significance with p-values less than 1×10-8 from these two families. These SNPs are proposed as candidate melanoma risk modifying variants and must be validated in other p16-linked cases and/or pedigrees. Discussion: Due to the difficulty of determining association in highly related individuals we propose this method as a possible solution to identifying associated SNPs in high-risk pedigrees. Further investigation of these variants will to include SNP × SNP interaction models and additional p16 confirmed individuals.

1150T

Genome-wide association studies identify four novel ER-negative specific breast cancer risk loci. *F.J. Couch[1], M. Garcia-Closas[2, 3], S. Lindstrom[4], K. Michailidou[5], M.K. Schmidt[6], M. Brook[2], N. Orr[3], S. Slager[1], D.J. Hunter[4], J. Simard[7], J. Benitez[8], A. Dunning[5], M.E. Sherman[9], G. Chenevix-Trench[10], S.J. Chanock[9], P. Hall[11], P. Pharoah[5], C. Vachon[1], D.F. Easton[5], C.A. Haiman[12], P. Kraft[4] for BPC3, TNBCC, and BCAC.* 1) Departments of Laboratory Medicine and Pathology and Health Sciences Research, Mayo Clinic, Rochester, MN; 2) Division of Genetics and Epidemiology, Institute of Cancer Research, Sutton, United Kingdom; 3) Breakthrough Breast Cancer Research Centre and Division of Breast Cancer Research, The Institute of Cancer Research, London, United Kingdom; 4) Program in Molecular and Genetic Epidemiology, Harvard School of Public Health, Boston, Massachusetts, USA; 5) Departments of Public Health and Primary Care and Oncology, University of Cambridge, Cambridge, United Kingdom; 6) Netherlands Cancer Institute, Antoni van Leeuwenhoek hospital, Amsterdam, The Netherlands; 7) Centre Hospitalier Universitaire de Québec and Laval University, Quebec, Canada; 8) Spanish National Cancer Research Centre [CNIO], Madrid, Spain; 9) Epidemiology and Biostatistics Program, National Cancer Institute, Bethesda, MD, USA; 10) Department of Genetics, Queensland Institute of Medical Research, Brisbane, Australia; 11) Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden; 12) Department of Preventive Medicine, Keck School of Medicine and Norris Comprehensive Cancer Center, University of Southern California, Los Angeles, CA, USA.

Estrogen receptor negative (ER-negative) tumors represent 20–30% of all breast cancers, with a higher proportion in younger women and women of African ancestry. Compared with tumors that express ER (ER-positive), ER-negative tumors are associated with a worse short-term prognosis and have weaker associations with reproductive risk factors. There are also important differences in genetic susceptibility to these two types of tumors. BRCA1 mutations predispose primarily to ER-negative disease and most common susceptibility loci for breast cancer, known to date, show stronger associations for ER-positive than ER-negative tumors. Exceptions are variants in the chromosome 5p15 (TERT-CLPTM1L) and 19p13 (MERIT40) loci that predispose primarily to ER-negative tumors. To identify additional ER-negative specific loci, three genome-wide association studies of 4,193 ER-negative breast cancer cases and 35,194 controls from the National Cancer Institute Breast and Prostate Cancer Cohort Consortium (BPC3), the Triple Negative Breast Cancer Consortium (TNBCC) and the Combined Breast Cancer Association Consortium (BCAC) ER-negative GWAS (C-BCAC) were combined with a replication series of 40 studies from BCAC (6,514 cases and 41,455 controls) in a meta-analysis. Four novel loci at 1q32.1 ($P=3.9\times10^{-13}$) and ($P=4.9\times10^{-9}$), 2p24.1 ($P=1.5\times10^{-8}$) and 16q12 ($P=2.0\times10^{-8}$) were associated with ER-negative but not ER-positive breast cancer (P>0.05). These findings provide further evidence for distinct etiologic pathways for invasive ER-positive and ER-negative breast cancer.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1151F

cAMP and cGMP signaling may play an important role in the development of prostate cancer. *R. de Alexandre[1,2], A. Horvath[1], A. Manning[1], D. Carraro[3], F. Soares[3], M. Nesterova[1], S. Constantine[1], F. Faucz[1,2].* 1) Section on Endocrinology & Genetics, PDEGEN, NICHD, NIH, Bethesda, MD, USA; 2) Laboratory of Molecular Genetics, NIMA, PPGCS, Pontificia Universidade Catolica do Parana, Curitiba, Brazil; 3) Laboratory of Genomics and Molecular Biology - A.C. Camargo Hospital, São Paulo, Brazil.

The second messengers cAMP and cGMP are responsible for a variety of cellular functions. Cyclic AMP and cGMP intracellular levels depend on the balance between their synthesis and degradation, i.e. on the activity of the adenylate and guanylate synthases and, the highly polymorphic family of phosphodiesterases (PDEs). Variants in PDEs have been associated with different disorders. The first PDE to be implicated in predisposition to tumors is *PDE11A*, whose inactivating mutations have been reported to be frequent in patients with adrenal, testicular and prostate tumors. Further, higher levels of cAMP have been measured in prostate cancer (PCa) compared to the normal tissue. In addition, cAMP and cGMP signaling is involved in PCa cell growth and modifies androgen receptor effects. We sequenced 16 different PCa tumor DNAs on SOLiD4 platform, after targeted enrichment for the coding parts of 196 genes, including the family of the phosphodiesterases, adenylate and guanylate synthases, all genes involved in cAMP and cGMP pathways and other genes related to endocrine tumors. After alignment and variation calling, 79 novel variants in 20 genes, and 120 previously reported SNPs in 21 of a selected list of 23 genes related to the cAMP and cGMP signaling were called in the 16 samples. The first 3 genes sequenced by Sanger for confirmation were *PDE11A*, *PDE5A* and *PRKAR1A*. While no mutations in *PRKAR1A* were confirmed, 2 novel synonymous SNPs were identified in *PDE5A* that are currently undergoing functional analysis. Analysis for *PDE11A* coding variants revealed results concordant with our previous findings: missense and nonsense variations were seen in the PC patients (allelic frequency 0.16 vs 0.043 in the controls, $\chi 2=18.23$, $p<0.0001$, OR=4.18, 95%CI 2.14–8.16). We also selected twenty-nine confidently called novel mutations in 15 genes from the same list of 23 genes of cAMP signaling for further confirmation and analysis. As a conclusion - while our results support the previously reported involvement of *PDE11A* in the predisposition to PCa, the role of *PDE5A* is yet to be determined. Genetic variants in *PRKAR1A* are not likely to be involved in the PCa predisposition.

1152W

Genome wide association for Multiple Myeloma reveals overlap between variants associated with susceptibility and survival. *E. Dean[1], D. Hu[1], P. Bracci[2], V. Krepkiy[1], T. Martin[1,3], J. Wolf[1,3], E. Ziv[1,3,4].* 1) Medicine, UCSF, San Franc, CA; 2) Epidemiology and Biostatistics, UCSF, San Francisco, CA; 3) Helen Diller Family Comprehensive Cancer Center, UCSF, San Francisco, CA; 4) Institute for Human Genetics, UCSF, San Francisco, CA.

Multiple Myeloma (MM) is an incurable hematological malignancy characterized by a clonal population of immunoglobulin secreting plasma cells. MM is known to have a significant familial component. One published genome wide association study (GWAS) has identified 2 susceptibility loci. MM is also known to be a heterogeneous disease with more and less aggressive biological subtypes that can be discerned by cytogenetic abnormalities and gene expression studies. We hypothesized that different subsets of disease would have different susceptibility genes and that these would lead to different survival rates among MM patients. We used a GWAS of MM to investigate the possibility that there is an overlap between SNPs that predict survival and susceptibility. We used a dataset of 300 MM patients seen at UCSF and 300 clinic based controls matched by race/ethnicity and gender. We genotyped the cases and controls using Illumina Omni5 arrays. We identified and adjusted for population stratification using principal components analysis (PCA), selecting the top 10 PC's. We performed case-control analysis using logistic regression models to identify potential susceptibility variants, and survival analysis using Cox models to identify variants affecting survival. We analyzed the overlap of markers by dropping SNPs in LD and then performing chi-squared tests to examine whether SNPs associated with survival are more likely to be associated with susceptibility. After dropping SNPs and samples with poor quality, we were left with ~2.5M variants typed on 285 cases and 300 controls. Using a relatively liberal cutoff of nominal significance ($p<0.05$), we evaluated whether there was overlap between the variants associated with survival and susceptibility. Based on chance alone, we would have expected 6934 SNPs nominally associated with susceptibility and survival in our dataset, but we saw 7528, a difference that is highly significant ($p<0.0001$). One locus on 8p11, was significantly associated with survival (p=2×10e-8) and nominally associated with susceptibility (p=0.01). Another locus on 18q12 had suggestive associations with both survival (p=2×10e-5) and susceptibility (p=7×10e-4). Our analysis demonstrates substantial overlap between germline variants that determine susceptibility to and survival from MM.

1153T

Genetic risk factors for malignant pleural mesothelioma (MPM): a genome-wide association study. *I. Dianzani[3], S. Guarrera[1], M. Betti[3], G. Fiorito[1], D. Ferrante[4], F. Voglino[1], G. Cadby[5], A. Russo[1], C. Di Gaetano[1,2], F. Rosa[1], E. Casalone[3], M. Padoan[4], M. Giordano[6], A. Aspesi[3], C. Casadio[7], F. Ardissone[8], E. Ruffini[9], P.G. Betta[10], R. Libener[10], R. Guaschino[11], E. Piccolini[12], L. Palmer[5], M. Neri[13], D. Mirabelli[14], D. Ugolini[15], S. Bonassi[13], C. Magnani[4], G. Matullo[1,2].* 1) Human Genetics Foundation, HuGeF, Turin, Italy; 2) Dept.Genetics, Biology and Biochemistry, Univ. Torino, Italy; 3) Laboratory of Genetic Pathology, Dept.Health Sciences, Univ.Piemonte Orientale, Italy; 4) CPO-Piemonte and Unit of Medical Statistics and Epidemiology, Dept.Translational Medicine, Univ.Piemonte Orientale,Italy; 5) Ontario Institute for Cancer Research (OICR), Toronto, Ontario, Canada; 6) Laboratory of Genetics, Dept.Health Sciences, Univ.Piemonte Orientale, Italy; 7) Thoracic Surgery Unit, Univ.Piemonte Orientale, Italy; 8) Chest Surgery, Dept.Clinical and Biological Sciences, Univ.Torino, Italy; 9) Chest Surgery, Dept.Clinical Physiopathology, Univ.Turin, Italy; 10) Pathology Unit, Hospital, Alessandria, Italy; 11) Transfusion Centre, Hospital, Alessandria, Italy; 12) Pneumology Unit Hospital, Casale Monferrato, Italy; 13) Unit of Clinical and Molecular Epidemiology IRCCS San Raffaele Pisana, Roma, Italy; 14) CPO-Piemonte and Unit of Cancer Epidemiology, Univ.Turin, Italy; 15) Department of Internal Medicine, University of Genoa and IRCSS AUO San Martino-IST-Istituto Nazionale per la Ricerca sul Cancro, Genoa.

Asbestos is the main risk factor for malignant pleural mesothelioma (MPM). Many different mechanisms have been hypothesized for its carcinogenic effect. Among them, ROS induced oxidative stress plays a crucial role, either directly or through activation of inflammatory cells. Only 10% of subjects exposed to high levels of asbestos develop MPM, which may be explained by individual susceptibility, including genetic risk factors. Previous studies suggested a role of SNPs in DNA-repair and redox enzymes. These studies report significant OR=1.4–3.5 for several SNPs and showed that genetic factors could indeed have a role in the pathogenesis of MPM, though much lower than that entailed by asbestos. Conversely, dominant mutations in BAP1 have been reported to cause a new cancer prone syndrome, characterised by susceptibility to MPM, melanoma, and other types of cancer. The aim of our work was the identification of genetic risk factors involved in the development of MPM. We performed a GWAs on 392 cases and 379 controls. A unique asset of our study is the complete clinical definition of asbestos exposures that is available for each patient and control. Cases and controls are from three towns in Northern Italy (Casale Monferrato, Turin, Genoa). Casale is a small town exposed to a high asbestos pollution because of an asbestos cement factory, active in 1907–1986. Besides factory workers, also Casale dwellers are at increased risk as compared with residents in the surrounding area. All subjects signed an informed consent form. Patients and controls were genotyped for more than 370.000 SNPs, using the Illumina HumanCNV370-Quad.v3 BeadChip. Preliminary data show that several SNPs (located in 6q21, 5q35, 19q13, 3q26, 4q32.1) were associated with MPM at a level of p=10-6/10-7 testing a per-allele model adjusted for gender, age, centre, exposure and PCA (Principal Component Analysis) in the overall and in the exposed-only case-control study (low vs higher exposure). Replication is on-going on a panel of 428 patients and 1269 controls from Australia (collaboration with Prof. L.Palmer). Functional and gene expression studies are ongoing on associated variants. Resequencing of the BAP1 gene on 100 patients is also ongoing. The definition of genetic risk factors contributing to the development of MPM may help to identify the pathophysiology of MPM and help in the definition of the actual risk of individuals exposed to asbestos.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1154F

A region of autozygosity at 8q21.3 and lung cancer risk. *M.T. Landi[1], K. Jacobs[1], M. Yeager[1], D. Albanes[1], M. Thun[2], N.E. Caporaso[1], S. Chanock[1], J. Shi[1]*. 1) National Cancer Inst, Bethesda, MD; 2) American Cancer Society, Atlanta, GA.

Background: Autozygosity occurs when two identical chromosomal segments are inherited from a common ancestor. Increased levels of autozygosity, either globally or locally, have been implicated as a risk factor for human complex diseases including cancer. Autozygosity can be inferred from the analysis of runs of homozygosity (ROHs; long stretches of homozygous single-nucleotide polymorphisms [SNPs]) in outbred populations from genome-wide SNP array data. We examined whether the proportion of the autosome estimated to be within an ROH (*Froh*) or any locus-specific ROH was associated with the risk of lung cancer overall or with specific histological subtypes. Materials and Methods: We performed an ROH analysis in a GWAS with 4,225 controls and 5,297 lung cancer cases of European ancestry, including 1,607 adenocarcinomas, 1,414 squamous carcinomas and 688 small cell carcinomas genotyped for 515,922 SNPs. ROH segments ≥500kb with ≥60 SNPs and no heterozygous genotypes were detected using PLINK. The association between *Froh* and lung cancer risks was evaluated using logistic regression, adjusting for age, sex, cigarettes per day, years of cigarette smoking and principal components for population stratification. Furthermore, we split the autosome into intervals of 500kb in length and, or each interval, we tested if the proportion of cases carrying ROHs was significantly higher than that of controls using the Fisher's exact test. The empirical genome-wide significance levels were evaluated by a permutation procedure where case-control status was permuted 10,000 times. Results: Froh was not significantly associated with the risk of lung cancer overall or with any lung cancer subtype, in ever, former or current smokers. One ROH at 8q21.3 was significantly associated with overall lung cancer risk in ever smokers (1.2% in cases and 0.35% in controls, OR=3.4, nominal P= 0.000015 and genome-wide P=0.017). This association was stronger for small cell carcinomas and squamous cell carcinomas. Significance: These findings offer new insights into the development of lung cancer and call for further research into recessive, lung cancer-predisposing loci.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1155W

IL28B allelic protein isoforms differentially activate interferon-stimulated genes. *A. Mumy[1], B. Muchmore[1], W. Tang[1], H. Park[2], F. Sheikh[3], B. Rehermann[2], R. Donnelly[3], L. Prokunina-Olsson[1]*. 1) Laboratory of Translational Genomics, Division of Cancer Epidemiology and Genetics, NCI, NIH, Gaithersburg, MD; 2) Immunology Section, Liver Diseases Branch, NIDDK, NIH, Bethesda, MD 20892; 3) Division of Therapeutic Proteins, Center for Drug Evaluation & Research, FDA, Bethesda, MD 20892.

IL28B (IFN-λ3) is a type-III interferon encoded by the IL28B gene located on chromosome 19q13.13. Recent genome-wide association studies (GWAS) have identified genetic variants upstream of IL28B, associated with spontaneous and treatment-induced clearance of HCV. Here, we sequenced IL28B gene in the HapMap reference set of 270 samples from 3 populations - Europeans, Asians and Africans. We found 3 protein-coding variants, rs8103142 (Lys70Arg), rs138893424 (Leu31Phe) and rs62120527 (Glu171-Lys), which were genotyped by sequencing and/or with TaqMan assays. We induced IL28B expression in primary human hepatocytes treated with Poly(I:C), and performed 5'rapid amplification of cDNA ends (5'-RACE), but found no alternative IL28B transcripts. We cloned IL28B transcripts with allelic variants of the 3 coding SNPs, and generated four recombinant purified IL28B protein isoforms using a baculoviral expression system. Functional analysis of these proteins was performed using a stable HepG2-ISRE-Luc cell line expressing an interferon-stimulated response element (ISRE) coupled with luciferase reporter gene. Our analysis showed that the IL28B protein carrying the beneficial T allele (70Lys) of rs8103142 and the common alleles of the two other variants was significantly more potent at inducing activation of the ISRE-Luc reporter compared to the other IL28B protein isoforms. We are currently testing these protein isoforms using STAT1 phosphorylation assays and evaluating their effects on inhibition of HCV replication and on activation of specific interferon-stimulated genes.

1156T

Three novel common susceptibility loci for ovarian cancer identified by GWAS meta-analysis and replication. *S.J. Ramus[1], P.D.P. Pharoah[2,3], Y.Y. Tsai[4], C.M. Phelan[4], A.N.A. Monteiro[4], SA. Gayther[1], J.M. Schildkraut[5,6], T.A. Sellers[4] on behalf of the Ovarian Cancer Association Consortium.* 1) Dept Preventive Medicine, Univ Southern California, Los Angeles, CA; 2) The Centre for Cancer Genetic Epidemiology, Department of Oncology, University of Cambridge, Cambridge, UK; 3) The Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK; 4) Department of Cancer Epidemiology, Division of Population Sciences, Moffitt Cancer Center, Tampa, FL, USA; 5) Department of Community and Family Medicine, Duke University Medical Center, Durham, NC, USA; 6) Cancer Prevention, Detection and Control Research Program, Duke Cancer Institute, Durham, North Carolina, USA.

Genome wide association studies (GWAS) have identified four susceptibility loci for epithelial ovarian cancer (EOC) with another two loci being close to genome-wide significance. We pooled data from a GWAS conducted in North America with another GWAS from the United Kingdom. The North American study comprised four independent case-control studies that included 1,952 EOC cases and 2,052 controls. The second study was a two-phase multi-center GWAS that included 1,817 EOC cases and 2,354 controls in the first phase and 4,162 EOC cases and 4,810 controls in the second phase. We carried out a fixed effects meta-analysis from the two GWAS for ~2.5 million genotyped or imputed SNPs. We selected the top 24,551 SNPs for follow-up genotyping and after applying quality control filters, we tested 22,252 SNPs for association with risk of all invasive EOC and for serous invasive EOC in 18,174 EOC cases (10,316 serous cases) and 26,134 controls. The primary analyses were based on the subjects of European ancestry (16,283 cases and 23,491 controls) from 43 studies as part of the Collaborative Oncological Gene-environment Study (COGS). The associations of the four previously reported SNPs at 2q31, 8q24, 9p22 and 19p13 were all confirmed. We validated the two loci at 3q25 and 17q21 previously near genome-wide significance and identified three novel loci associated with risk; two loci associated with all EOC subtypes, at 8q21 (rs11782652, odds ratio (OR) = 1.19, 95% CI: 1.12–1.26, $P=5.5\times10^{-9}$) and 10p12 (rs1243180; OR = 1.10, 95% CI:1.06–1.13,$P=1.8\times10^{-8}$), and another locus specific to the serous subtype at 17q12 (rs757210; OR = 1.12, 95% CI:1.08–1.17,$P=8.1\times10^{-10}$). The association of rs757210 at 17q12 with all invasive EOC was much weaker (OR = 1.05, 95% CI 1.02–1.09, $P=9\times10^{-4}$); there was substantial heterogeneity by tumor subtype (P=<0.0001) with the risk allele for serous EOC being associated with a reduced risk of both clear cell and mucinous EOC. An integrated molecular analysis of genes and regulatory elements within a 1Mb region at each locus provided evidence for functional mechanisms underlying susceptibility with specific involvement of the CHMP4C (8q21), MLLT10 (10p12) and HNF1B (17q12) genes in the pathogenesis of ovarian cancer.

1157F

Insights into pancreatic cancer etiology from pathway analysis of genome-wide association study data. *P. Wei[1], H. Tang[2], D. Li[2]*. 1) Human Genetics Center, The University of Texas Health Science Center at Houston, Houston, TX; 2) Department of Gastrointestinal Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX.

Pancreatic cancer is the fourth leading cause of cancer death in the U.S. and the etiology of this highly lethal disease has not been well defined. To identify genetic susceptibility factors for pancreatic cancer, we conducted a comprehensive pathway analysis of genome-wide association study (GWAS) data in 3,141 pancreatic cancer patients and 3,367 controls with European ancestry, pooled from nested case-control studies-based PanScan1 and case-control studies-based PanScan2. Using the gene set ridge regression in association studies (GRASS) method we analyzed 197 pathways identified from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. We used the logistic kernel machine (LKM) test to identify major contributing genes to each pathway. We conducted functional enrichment analysis of the most significant genes (P<0.01) using the Database for Annotation, Visualization, and Integrated Discovery (DAVID). Six pathways were significantly associated with risk of pancreatic cancer after adjusting for multiple comparisons (P <0.00025 based on 5000 permutations): Fc epsilon RI signaling, maturity onset diabetes of the young, neuroactive ligand receptor interaction, long-term depression (P<0.0002), and the olfactory transduction and vascular smooth muscle contraction pathways (P=0.0002). As a replication effort, we performed the GRASS pathway analysis on the data from PanScan1 (1,796 cases and 1,880 controls) and PanScan2 (1,345 cases and 1,487 controls) separately. Two of the six significant pathways, i.e. the olfactory transduction pathway and the neuroactive ligand-receptor interaction pathway showed consistent small P values across the PanScan1 and PanScan2 cohorts, though not both P values were significant after multiple testing corrections likely due to much smaller sample size in each individual cohort. The meta-analysis P values for these two pathways (1×10^-5 and <1×10^-5, respectively) were significant after the Bonferroni correction. LKM test identified four genes that were significantly associated with risk of pancreatic cancer after the Bonferroni correction (P<1×10^-5): ABO, HNF1A, OR13C4, and SHH. Functional enrichment analysis using DAVID consistently found the G protein-coupled receptor signaling pathway to be the most significant pathway for pancreatic cancer risk in this study population. If confirmed, these novel findings provide new perspectives on genetic susceptibility to and molecular mechanisms of pancreatic cancer.

1158W

Pathway-based analysis of genome-wide SNP data reveals new candidate genes for susceptibility to melanoma. *M. Brossard[1,2,3], A. Vaysse[1,2,3], E. Corda[1,2,3], P. Jeannin[1,2,3], H. Mohamdi[1,2,3], V. Chaudru[1,3,4], N. Lavielle[1,2,3], B. Bressac-de Paillerets[1,5], M.F. Avril[6], M. Lathrop[3], F. Demenais[1,2,3]*. 1) INSERM U946, Paris, France; 2) Université Paris Diderot, Paris, France; 3) Fondation Jean Dausset-CEPH, Paris, France; 4) Université d'Evry, Evry, France; 5) Service de Génétique, Institut Gustave Roussy, Villejuif, France; 6) Assistance Publique-Hôpitaux de Paris, Hôpital Cochin, Paris, France.

Genome-wide association studies (GWAS) for melanoma have identified 16 melanoma risk loci. However, these loci explain only a small part of the genetic susceptibility to melanoma. These GWAS were conducted by using single-marker analysis which may be underpowered to detect SNPs with small effects and/or interacting with other SNPs. Multi-marker methods, including pathway-based approaches, have been proposed as an alternative strategy to prioritize new susceptibility genes. Our goal was to use a pathway-based approach implemented in the ALIGATOR software to identify an overrepresentation of biological pathways, indexed by gene-ontology (GO) terms, associated with risk of melanoma. The input data to ALIGATOR consisted of association results of 1.4 millions Hapmap3-imputed SNPs that were obtained from a GWAS conducted in 1186 melanoma cases and 2838 controls from the French MELARISK study. A total of 593,915 SNPs lied within 20kb of 22,648 genes and these genes were assigned to 7,371 GO categories. We found that 12 GO categories achieved a category-specific p-value of 0.001 or lower. This number of GO categories was significantly greater than expected by chance (p=0.001 after correction for multiple testing). The 12 GO classes included 8 of the 16 loci identified by previous GWAS (*MC1R*, *TYR*, *HERC2/OCA2* locus, *SLC45A2*, *TERT*, *MTAP/CDKN2A/CDKN2B* locus, *ATM*, *PARP1*), which confirms the validity of this pathway approach. These loci were grouped into melanin biosynthesis process, DNA metabolic process and telomerase activity pathways. In addition, 5 new candidate genes belonging to DNA replication and replisome pathways were characterized; these genes are located on chromosomes 5q31, 12q13 and 22q11.21. This study outlines the advantage of using a multi-marker approach to reveal new pathways and new candidate loci potentially involved in melanoma susceptibility. Further investigation of these loci in other melanoma datasets and search for gene-gene interactions within and across the pathways characterized by this study are worth pursuing. Funding: Ligue Nationale Contre le Cancer (PRE 09/FD), PHRC (AOM-07–195).

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

† Trainee Award Finalist

1159T

Susceptibility Loci Associated with Specific and Shared Subtypes of Lymphoid Malignancies. *V. Joseph[1,2], T. Kirchhoff[3], J. Brown[5], K.A. Schrader[1,2], A. Dutra-Clarke[1], C. Manschreck[1], N. Hansen[1], R. Rau-Murthy[1], K. Sarrel[1], J. Przybylo[1], S. Shah[1,2], S. Cheguri[1], Z. Stadler[1], L. Zhang[6], O. Paltiel[7], D. Yehuda[7], A. Viale[8], C. Portlock[4], D. Strauss[4], S.L. Lipkin[9], M. Lacher[4], M. Robson[1], R.J. Klein[2], A. Zelenetz[4], K. Offit[1,2,4,9].* 1) Dept. Medicine, Sloan-Kettering Institute for Cancer Research, NEW YORK, NY; 2) Cancer Biology and Genetics Program, Sloan Kettering Institute, New York, New York, USA; 3) NYU Cancer Institute, New York University School of Medicine, New York, New York, USA; 4) Lymphoma Service, Department of Medicine, Memorial Sloan-Kettering Cancer Center, New York, New York, USA; 5) Dana Farber Cancer Institute, Boston, Massachusetts, USA; 6) 6) Diagnostic Molecular Genetics Laboratory, Department of Pathology, Memorial Sloan-Kettering Cancer Center, New York, New York, USA; 7) Department of Hematology, Hadassah-Hebrew University Medical Center, Jerusalem, Israel; 8) Genomics Core, Memorial Sloan-Kettering Cancer Center, New York, New York, USA; 9) Weill Cornell Medical Center, New York, New York, USA.

Using a two-stage GWAS, we tested 530,583 SNPs in 944 cases of lymphoma, including 282 familial cases, and 4059 public shared controls, followed by genotyping of 50 SNPs in 1245 cases and 2596 controls. Six loci achieved genome-wide significance for the subsets of follicular or diffuse large B-cell, and five in joint analysis of all lymphoma subtypes combined (LYM). We discovered a novel locus on 6p23 (rs707824, PLYM=1.22 × 10-9) near JARID2 and another on 11q12.1 (rs12289961, PLYM=3.06 × 10-9; rs948562, PLYM=2.11 × 10-7) near LPXN, both of which confer susceptibility to B-cell lymphoma. The most significantly associated with combined types of non-Hodgkin's lymphoma (NHL) was rs707824 (PNHL=1.67 × 10-10). Candidate genes in these regions have been implicated in B-cell development, hematopoiesis and immune function. Our data also confirms association with a region near the HLA locus at 6p21.32; we demonstrate allelic heterogeneity, with rs2647046 (PFL=1.88 × 10-15), rs4530903 (PFL=6.65 × 10-13), rs2621416 (PFL=2.95 × 10-9) and rs9268853 (PFL=5.53 × 10-11) near HLA-DQA1/2, HLA-DQB1/2, and TAP2, independently associated with follicular lymphoma. Genes implicated by GWAS were also found to be cis-eQTLs in lymphoblastoid cell lines. Eight of the nine SNPs were putative regulatory SNPs as shown by in silico analysis. Thus, these results, showing both locus and allelic heterogeneity, point to the existence of pathways of susceptibility to both shared as well as specific subtypes of lymphoid malignancy.

1160F

Second Generation DCEG Reference Set Improves Performance of Genotype Imputation. *Z. Wang[1,2], K.B. Jacobs[1,2], M. Yeager[1,2], A. Hutchinson[1,2], J. Sampson[2], M. Tucker[2], S.J. Chanock[2].* 1) Core Genotyping Facility, SAIC-F/NCI, Bethesda, MD; 2) Division of Cancer Epidemiology and Genetics, NCI, NIH, Bethesda, MD.

Genotype imputation has become an essential analytic approach for fine-mapping and meta analysis of GWAS data sets genotyped and analyzed on different arrays. The NCI Division of Cancer Epidemiology of Genetics (DCEG) recently released a first version of an imputation reference set through the dbGaP webportal (Wang et al Nature Genetics 44:6-7, 2011 available at http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000396.v1.p1). Earlier, we showed that performance is better for genotyped samples compared to low-pass sequence data. To improve the practical utility of the DCEG Reference Set, we have substantively increased the number of genotyped SNPs in the same set of 1,249 cancer-free individuals, primarily of European ancestry. New content includes SNP microarray data from both Illumina (Omni2.5S, Exome, Immunochip, Metabochip and the iCOGs custom cancer microarrays) and Affymetrix (6.0 and two Axiom microarrays), which will further facilitate larger, international meta-analyses and imputation projects. Ongoing work includes comparison of version 2 to recent builds of the 1000 Genomes Project. Later this year, version 2 of the DCEG Imputation Set will be available on the dbGaP webportal. Funded by NCI Contract No. HHSN261200800001E.

1161W

Are there genetic variants that confer shared genetic susceptibility to pancreatic cancer and melanoma? *L. Wu, K. Rabe, G. Petersen.* Health Sciences Research, Mayo Clinic, Rochester, MN.

Background: Pancreatic cancer researchers (including Mayo Clinic) found that certain subsets of pancreatic cancer patients are at an increased risk for melanoma. Melanoma researchers have similarly found increased risk for pancreatic cancer in their patients. We hypothesized that these two diseases possibly share genetic risk factors. Previous studies found that CDKN2A (p16) can explain some familial clusters of these two cancers, but evidence for a more broad connection between pancreatic cancer and melanoma is largely lacking. Methods: We evaluated the association between 38 single-nucleotide polymorphisms (SNP) within 22 genomic regions which are associated with melanoma risk and the risk of pancreatic cancer using Whole Genome Scan for Pancreatic Cancer Risk in the Pancreatic Cancer Cohort Consortium and Pancreatic Cancer Case-Control Consortium (PanScan) datasets, which include PanScan 1, PanScan 2, combined PanScan, and Mayo Clinic subset data. We also evaluated the association between 9 pancreatic cancer susceptibility variants within 4 genomic regions and the risk of melanoma using MD Anderson Cancer Center melanoma genome-wide association study (GWAS) dataset. Results: In the PanScan datasets, several associations were statistically significant at p=0.05 threshold. For example, OCA2 (rs1800407) (OR=1.152, 95% CI 1.001–1.325, p= 0.049) in PanScan 2 dataset; NCOA6 (rs4911442) (OR=1.323, 95% CI 1.032–1.697, p=0.027), YWHAZP5 (rs17119461) (OR= 2.617, 95% CI 1.079–6.35, p=0.033), and YWHAZP5 (rs17119490) (OR= 2.616, 95% CI 1.078–6.344, p=0.033) in the Mayo Clinic subset data,. In the melanoma GWAS dataset, we detected two pancreatic cancer susceptibility variants to be associated with an increased risk of melanoma: NR5A2 (rs12029406) (OR=1.393, 95% CI 1.011–1.919, p=0.043), and CLPTM1L-TERT (rs401681) (OR=1.16, 95% CI 1.006–1.337, p=0.041). None of the above associations were statistically significant after the Bonferroni correction for multiple comparisons. Conclusion: There is no evidence that GWAS-derived melanoma susceptibility variants play an important role in pancreatic cancer predisposition. Also pancreatic cancer susceptibility variants are not associated with melanoma risk. If there is a shared genetic etiology of these two cancers, it may be detected by genetic analysis other than currently tested GWAS SNPs. Family-based linkage studies or high throughput sequencing studies offer possible approaches.

1162T

Oncogenic rewiring by the t(11;19)(q21;p13) coactivator fusion, CRTC1/MAML2, involves direct activation of the pluripotency factor MYC. *A.L. Amelio[1], F.X. Schaub[1], M. Fallahi-Sichani[2], M.B. Lawani[1], M.R. Southern[3], B.M. Young[4], L. Wu[5], F.J. Kaye[6], J.L. Cleveland[1], M.D. Conkright[1].* 1) Department of Cancer Biology, The Scripps Research Institute, Jupiter, FL; 2) Informatics Department, The Scripps Research Institute, Jupiter, FL; 3) Translational Research Institute and Molecular Therapeutics, The Scripps Research Institute, Jupiter, FL; 4) Genomics Core, The Scripps Research Institute, Jupiter, FL; 5) Department of Molecular Genetics and Microbiology, Shands Cancer Center, University of Florida, Gainesville, FL; 6) Department of Medicine, Shands Cancer Center, University of Florida, Gainesville, FL.

Oncogenesis is a complex, multi-factorial process of cellular transformation that leads to the development of many types of cancers. The factors that contribute to this process reprogram normal cellular functions to allow for uncontrolled cell growth and proliferation. The cAMP Regulated Transcription Coactivator (CRTC) family of CREB coactivators, in conjunction with CBP/p300, cooperate in the regulation of cAMP-inducible genes involved in cell survival, proliferation, glucose metabolism, and adaptive mitochondrial biogenesis. A subset of tumors share a common t(11;19)(q21;p13.1) translocation that forms a chimeric oncogene by fusing CRTC1 to the NOTCH coactivator MAML2. Consequently, the CRTC1/MAML2 (C1/M2) coactivator fusion induces the aberrant expression of genes regulated by CREB and NOTCH and the deregulation of these target genes is believed to cause tumorigenesis. However, using a mammalian genetic screen, we identified a C1/M2 gain-of-function activity that promotes interactions with the MYC:MAX network. RNA-seq analysis confirmed that a significant proportion of genes involved in key aspects of cell growth, survival, and metabolism within the MYC:MAX and CREB transcription networks are induced by C1/M2. Analysis of cellular transformation by RK3E foci formation assays identified a synergistic effect of MYC expression on C1/M2-induced transformation and this can be blocked by a dominant negative MYC molecule. Furthermore, in-frame deletions of C1/M2 that lack transforming activity are unable to induce the expression of MYC-responsive reporters revealing a critical role for MYC target genes in C1/M2-induced cell growth and transformation. Collectively, these studies suggest that ectopic interactions between C1/M2 and MYC contribute to the initial stages of tumorigenesis or in maintaining tumor stem cells given that CRTC1 activity has been implicated in epithelial stem cell proliferation and MYC is important for cell `stemness' and pluripotency. Acknowledgments: This study is supported by a Ruth L Kirschstein National Research Scholar Award from the National Cancer Institute (CA134121), PGA National Cancer Award, and by the Margret Q. Landenberger Foundation.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1163F

Meta-analysis identifies microRNA expression signature in non-small-cell lung cancer. *T. Annilo[1], U. Võsa[1], T. Vooder[2], R. Kolde[3], J. Vilo[3], A. Metspalu[1,4].* 1) Department of Biotechnology, Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia; 2) Clinic of Cardiovascular and Thoracic Surgery of Tartu University Hospital, Puusepa 8, 51014 Tartu, Estonia; 3) Institute of Computer Science, University of Tartu, Liivi 2, 50409 Tartu, Estonia; 4) Estonian Genome Center, University of Tartu, Riia 23b, 50410 Tartu, Estonia.

The prognostic and diagnostic value of microRNA (miRNA) expression aberrations in non-small cell lung cancer (NSCLC) has been studied intensely in recent years. However, due to the application of different technological platforms and small sample size, the miRNA expression profiling efforts have led to inconsistent results between the studies. We performed a comprehensive meta-analysis of 17 published miRNA expression datasets in NSCLC, including a total of 501 tumor and 431 non-cancerous control samples. Using a recently published Robust Rank Aggregation method, we identified statistically significant miRNA meta-signature of four up- and seven down-regulated miRNAs. We conducted a gene set enrichment analysis to identify pathways that are most strongly affected by altered expression of these miRNAs. We found that meta-signature miRNAs cooperatively target functionally related and biologically relevant genes in signaling and developmental pathways. We have shown that such meta-analysis approach is suitable and effective solution for identification of statistically significant miRNA meta-signature by combining several miRNA expression studies. This method allows the analysis of data produced by different technological platforms that cannot be otherwise directly compared or in the case when raw data is unavailable.

1164W

The effect of non-viral magnet assisted transfection on chronic myeloid leukemia cell line with down regulated miR-150. *T. Balci[1], C. Biray Avci[1], S. Yilmaz[1], Z.O. Dogan Sigva[1], G. Saydam[2], C. Gunduz[1].* 1) Medical Biology, Ege University Medical Faculty, Izmir, Turkey; 2) Hematology Department, Ege University Medical Faculty, Izmir, Turkey.

Chronic myeloid leukemia (CML) is a clonal myeloproliferative disease resulting from the neoplastic transformation of multipotent stem cells. The average age of CML cases is between 45 and 55 years old. MicroRNAs (miRNAs) are single strand, non-coding RNA molecules with the length of 18–25 nucleotides which have role in the regulation of gene expression. In various tumors, tumor suppressor miRNAs have been described to be deleted. For example; down regulation of miR-150 was proven in chronic myeloid leukemia. In recent years, polymeric nanoparticles are used as drug and gene carriers. Magnet assisted transfection (MATRA) is one of the most effective non-viral transfection methods. We aimed to determine the expression profiles of 96 cell cycle control genes in K-562 cell lines which transfected with down regulated miR-150 by MATRA mediated non-viral method. Total RNA was isolated from the cells exposed to transfection, the expressions of 96 genes each from apoptotic pathway and cell cycle controls were studied by real time online RT-PCR. Results were compared with control cells. As a result the expression profiles of CASP5, Fas, CD27, TRAF4, TNFRSF1A genes showed a significant increase and, CDK5RAP1, ATM, MYC, CDK8, CCND2 genes showed decrease, compared to the control cells. In conclusion, down-regulated miR-150 substitution showed the regulation of miRNA-mediated cell-cycle and apoptosis. The alterations of the genes' expressions have to be elucidated in protein level for new approaches in CML treatment.

1165T

The novel resequencing Diagnostic Microarray: RDMGGA1.0, can detect at a diagnose level mutations in patients with Breast, Ovarian, Colon, Skin and Multiple Cancers. *D. Bercovich[1], Y. Plotsky[2], T. Morad[2], S. Allon-Shalev[3], AM. Lichanska[4], LA. Borsuk[4], C. Tibbetts[4].* 1) Human Molec Gen & Pharm, MIGAL-Galille Tech Ctr, Kiryat-Shmona, Israel; 2) Galil Genetic Analysais (GGA), Kazarin 12900, Israel; 3) The Institute for Genetics, Ha'Emek Medical Center, Afula 18101, Israel; 4) TessArae, Potomac Falls, VA 20165 USA.

Despite the remarkable progress in sequencing methods for genes causing the most common inherited cancers, most current diagnostic methods and algorithms do not analysis mutations at a clinical level of affected patients and relatives. In addition, the large genes size and the lack of highly predominant mutational hotspots for some populations frequently make mutation detection in these cancer genes outstandingly challenging, which is costly and time consuming. To fulfill this technological gap, we developed a new customized resequencing gene chip (RDMGGA1.0) that is focused on 11 genes: BRCA1, BRCA2, APC, MUTYH, MLH1, MSH2, MSH6, TP53, PTEN, P16 and KRAS simultaneously, in a single assay, with high call rate and accuracy. Probes were designed to identify each base for all exons of these genes with 20 bases of intronic sequence bordering exons in the main tiles. Additionally, 3320 known and most frequent mutations were subtiled as well. The array uses the Affymetrix resequencing platform. Novel software was developed by TessArae for the data analysis. Amplicons were hybridized to the chip, and nucleotide detection was validated by standard capillary sequencing methods. Hybridization of amplicons with the chip produced high nucleotide sequence readout for all 11 genes in a single assay, with an automated call rate of over 98%. In a validation procedure of 150 samples with known mutations, the accuracy of nucleotide calls was 99.99% when compared with capillary sequencing in two separate labs (GGA, Israel; TessArae, USA). The new resequencing chip enables efficient analysis of 11 genes and identifies disease-causing mutations in two working days.

1166F

A New Approach for an Anti-diabetic Drug Metformin in Prostate Cancer. *C. Biray Avci[1], E. Harman[2], S. Yilmaz[1], Y. Dodurga[3], C. Gunduz[1].* 1) Medical Biology, Ege University Medical School, Izmir, Turkey; 2) Endocrinology Dept, Katip Celebi Universtiy Medical Faculty, Izmir, Turkey; 3) Medical Biology, Pamukkale University, Denizli, Turkey.

Prostate cancer is the most commonly diagnosed cancer in the male population. Metformin is the most widely used antidiabetic drug in the world, and there is increasing evidence of a potential efficacy of this agent as an anticancer drug. Metformin inhibits the proliferation of a range of cancer cells including prostate, colon, breast, ovarian, and glioma. MicroRNAs (miRNAs) are a class of small, non-coding, single-stranded RNAs that negatively regulate gene expression. We aimed to evaluate the effects of metformin treatment on the cellular miRNA expression changes in PC-3 cells and possible association of miRNAs and metformin on prostate cancer. Average cell viability and cytotoxic effect of metformin were investigated at 24 hours intervals for three days by using WST-1 assay. Cytotoxic effects of metformin in PC-3 cells were detected in time and dose dependent manner with the IC50 doses. RNA samples analysed using custom multi-species microarrays containing 1209 probes covering 1221 human mature microRNAs present in miRBase 16.0 database. The microarrays were scanned and data was extracted from images. Each array contains probes targeting 11 different synthetic miRNAs. Sensitivity of the microarray hybridization was confirmed by detection of hybridization signal for all 11 spikes well above the detection threshold. The array also contains a set of specificity control probes complementary to three different miRNAs. For statistical analysis, samples were binned in to different treatment groups. The log2-transformed and normalized spot intensities for the detectable human probes were examined for differences using the ANOVA module. The statistical significance was determined using the False Discovery Rate method. Fold changes between the different treatment and control samples were computed and presented along with the log2 transformed spot intensities of the detectable probes. Data for the detectable human probes were clustered. Genes were median centered prior to hierarchical clustering. Among the human miRNAs investigated by the arrays, 10 miRNAs were up-regulated and 12 miRNAs were down-regulated in metformin dose group compared with control cells. In conclusion, expression changes in miRNAs of miR-146a, miR-100, miR-425, miR-193a-3p, miR-106b in the metformin exposed cells may be important. This work may be demonstrated a new role of metformin on the regulation of miRNAs in prostate cancer.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1167W

mRNA-Seq of Single Prostate Cancer Circulating Tumor Cells Reveals Recapitulation of Gene Expression and Pathways Found in Prostate Cancer. *G. Cann[1], Z. Gulzar[2], S. Cooper[1], M. Tat[1], R. Li[1], S. Stuart[1], S. Luo[1], M. Ronaghi[1], J. Brooks[2], A. Talasar[1].* 1) Illumina Inc, Hayward, CA; 2) Department of Urology, Stanford University, Stanford, CA.

Circulating tumor cells (CTC) mediate metastatic spread of many solid tumors types, and enumeration of CTCs is used as a prognostic indicator for metastatic prostate cancer patients. Evidence suggests that information about primary tumors can be derived from expression analysis of CTCs. We assessed the ability of the MagSweeper to isolate CTCs suitable for mRNA-Seq. Using LNCaP cells spiked into normal blood, cell recovery, purity and viability were measured post-isolation with antibody staining against EpCAM, CD45 and staining of membrane-compromised cells with the nuclear dye DAPI. For enumeration comparisons with CellSearch, duplicate blood samples were collected with one sample processed by the MagSweeper and the other collected in a CellSave tube and processed by Quest Diagnostics. mRNA-Seq libraries were constructed with the SMARTer Ultra Low Input RNA Kit for Illumina Sequencing (Clontech) followed by single-read 50bp sequencing on an Illumina Genome Analyzer IIx. mRNA-Seq data was analyzed using CASAVA 1.8, and aligned to human genome build 19 using Tophat v1.3.3. Counts per transcript were calculated in iGenomes using cufflinks v1.1. Gene-by-gene raw counts and FPKM were generated from the cufflinks isoform.fpkm tracking file by identifying all the transcripts with the same gene name and taking the sum of coverage transcript length / read length for each transcript and the sum of the FPKM for each transcript. The MagSweeper isolated CTCs with a capture efficiency matching the CellSearch platform. Unlike CellSearch, the MagSweeper allowed final isolation of live CTCs without contaminating leukocytes. MagSweeper isolation did not have discernable impacts on transcriptional signatures of LNCaP cells isolated from spiked human blood as measured by mRNA-Seq. RNA from patient CTCs showed signs of significant degradation, consistent with reports of short half-lives and apoptosis amongst CTCs. Transcriptional signatures of prostate tissue such as androgen receptor and KLK3 (PSA) expression were detectable in prostate CTC using mRNA-Seq. A gene list of 182 genes up-regulated in prostate CTCs compared to normal prostate was derived. The MagSweeper allows isolation of patient CTCs suitable for mRNAseq. Expression analysis of prostate CTCs confirms expression of prostate specific genes and allowed new gene discovery. These results indicate that MagSweeper isolation coupled with CTC mRNA-Seq has the potential of supplying clinically relevant data.

1168T

Detection of sequences of MMTV-like retroviruses in breast cancer from Mexican women. *A. Cedro-Tanda[1,4], I.A. Cordova-Solis[1], D.J. Arenas-Aranda[1], J. Torres-López[2], F.A. Salamanca-Gomez[1], C. Moctezuma-Meza[3], G. Castelazo-Rodriguez[3], N. Garcia-Hernandez[1].* 1) Unidad de Investigación Médica en Genética Humana, Hospital de Pediatría, Centro Médico Nacional Siglo XXI. Instituto Mexicano del Seguro Social. México, D. F; 2) Unidad de Investigación Médica de Enfermedades Infecciosas, Hospital de Pediatría, Centro Médico Nacional Siglo XXI. Instituto Mexicano del Seguro Social. México, D. F; 3) Centro Médico La Raza. Instituto Mexicano del Seguro Social. México, D. F; 4) Posgrado en Ciencias Biológicas. Universidad Nacional Autónoma de México. México, D. F.

Breast cancer is the leading cause of cancer death in Mexico. In the last two decades has been an increase interest in research on the possible viral etiology of human breast cancer. Around the world, has demonstrated the presence of sequences homologous to the Env gene of Mouse Mammary Tumor Virus (MMTV) in 30–74% of human breast tumors, but not in normal breast tissue. In this work we analyzed the presence of a similar sequence to the MMTV Env gene in biopsies of Mexican women with breast cancer. As a positive control we used the MMTV env gene of CH3 strain. DNA integrity was tested by amplifying GAPDH gene. Was detected a 250pb fragment of MMTV Env gene in 8.8% of tumors analyzed and was not observed in non affected tissue. These data were confirmed by real-time PCR using TaqMan probes. The 250pb amplified fragments were sequenced to reveal it identity and homology among themselves and with MMTV. A total of 300 patients were analyzed. At present time we are working to evaluate the expression of Env and Gag genes through TaqMan probes and to determine insertion sites of the retrovirus into the genome by PCR splinkerette in MMTV-positive tumors. Our results support the hypothesis of a possible viral agent involved in the development and progress of breast carcinogenesis and also support the ongoing studies towards a practical possibility of primary prevention of breast cancer.

1169F

Expression of Stem Cell Gene, Piwil2 and Testis specific genes TSGA10, TEX101 and ODF3 in Breast Cancer. *M. Dianatpour[1], P. Mehdipour[2], M. Miryounesi[2], S. Savad[2], F. Yazarlou[2], M. Mobasheri[2], M.H. Modarressi[2].* 1) Medical Genetics, Shiraz university of medical sciences, Shiraz, Iran; 2) Medical Genetics, Tehran university of medical sciences, Tehran, Iran.

Background: Breast cancer is the most common malignancy in women around the world so finding new biomarkers for early detection and also identification of appropriate target antigens for specific cancer immunotherapy is valuable. Cancer stem cell theory explains the initiation and growth of cancers, like that of breast cancer from a group of cells in the tumor tissues known as cancer stem cells. There are some evidences that cancer stem cells had some similarity to germ line stem cells and express testis specific genes. Cancer testis (CT) genes are a group of genes expressed solely in testis and in a range of human malignancies so they are ideal targets for cancer immunotherapy. Methods: In this study we determined the expression of stem cell gene, Piwil2 and cancer testis genes TSGA10, TEX101 and ODF3 in 50 patients with breast cancer as well as adjacent non cancerous tissues (ANCT) as normal control tissues by RT-PCR. Expression of these genes was also examined in two breast cancer cell lines, MCF 7 and MDA-MB-231. The presence of auto antibody against cancer testis genes in patients' serums in order to show immune response against them in breast cancer patients was carried on by ELISA method. We also used western blot technique to determine the expression of TSGA10 protein in breast cancer cell lines. Results: In our patients 90% showed expression of Piwil2 and 70% showed TSGA10 expression but no one showed expression of TEX101 and ODF3. About12% of patients was positive for anti TSGA10 but all patients were negative for anti TEX101 and anti ODF3. Both of breast cancer cell lines exhibited very strong expression of Piwil2 and TSGA10. Conclusions: We showed a high expression rate of Piwil2 and TSGA10 in breast cancer patients as well as the lack of expression of these antigens in normal breast tissues, suggesting that these two antigens may be of therapeutic value as a vaccine for breast cancer immunotherapy but these antigens should be further studied to check immunoprevention of them in breast cancer patients.

1170W

PMS2 Screening: could massive parallel sequencing facilitate gene analysis? *J. Duclos[1,2], F. Pires[1], G. Legrand[1], C. Colas[1], J. Lefevre[2], M. Eyries[1], C. Wang[3], Y. Parc[2], F. Soubrier[1], F. Coulet[1].* 1) Genetics, GH Pitie-Salpetriere, APHP, PARIS, France; 2) Digestive Surgery, Saint Antoine Hospital, APHP, PARIS, France; 3) Molecular Oncology, Centre Leon Berard, Lyon, France.

Hereditary non polyposis colorectal cancer (HNPCC) syndrome, is known for its familial cancer spectrum directly linked to mismatch repair (MMR) gene mutation. MLH1 and MSH2 genes are most frequently implicated, but PMS2, which forms a heterodimer with MLH1, can also be responsible for HNPCC syndrome. Combination of clinical criteria, positive MSI status, negative MLH1 methylation and loss of MMR protein expression by tumor immunohistochemical (IHC) staining generally lead to targeted MMR gene analysis. Difficulties in PMS2 mutation analysis are inherent to sequence similarity of PMS2 to the PMS2CL pseudogene. To evaluate massive parallel sequencing using 454 pyrosequecing technology after multiplex amplification (HNPCC MASTR®) in PMS2 screening, we report PMS2 analysis in 10 selected patients diagnosed with colorectal cancer. None of the patients fulfilled Amsterdam criteria II but all of them showed positive MSI status, negative MLH1 methylation, and had no germinal mutation of MLH1 (Sanger sequencing tested). Two patients showed specific and complete loss of PMS2 staining, and 4 patients showed both MLH1 and PMS2 loss of expression. One patient had low PMS2 expression level and in 2 patients, PMS2 staining was not interpretable. One of the patients did not have PMS2 IHC analysis as tumor was not available. We also analyzed four control DNAs for whom mutations were already detected on PMS2 by Sanger sequencing. One of the two patients who had not interpretable PMS2 staining showed a frameshift mutation in exon 13 (c.2243_2246delAGAA). In the other patients, we found several reported sequence variants but we failed to identify deleterious mutation. The primers for the last (3') 5 exons (exons 11–15) were not discriminative between PMS2 and the pseudogene PMS2CL because several PMS2CL sequences were obtained. In known-mutated patients, we successfully identified deleterious mutation in exon 2: c.137G>T, in exon 5: c.400C>T, in exon 11: c.1579delA, and in exon 13: c.2249G>A. These findings demonstrated that we were able to detect mutation in PMS2, even in exon 11 and 13 where primers usually amplify the pseudogene sequence, and thus the potential interest of massive parallel sequencing in PMS2 analysis, despite PMS2CL interference. This technique could also be used to test the biallelic mutation of PMS2. However, further larger studies are needed to corroborate our findings.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1171T

FANCD2 immunohistochemical expression is a prognostic marker in breast cancer. *R. Fagerholm[1], K.M. Sprott[2], T. Heikkinen[1], J. Bartkova[3], P. Heikkilä[4], K. Aittomäki[5], J. Bartek[3,6], D.T. Weaver[2], C. Blomqvist[7], H. Nevanlinna[1].* 1) Department of Obstetrics and Gynecology, University of Helsinki and Helsinki University Central Hospital, Helsinki, Finland; 2) On-Q-ity Inc., 610 Lincoln St., Waltham, MA, USA; 3) Institute of Cancer Biology and Centre for Genotoxic Stress Research, Danish Cancer Society, Strandboulevarden 49, DK-2100 Copenhagen, Denmark; 4) Department of Pathology, University of Helsinki and Helsinki University Central Hospital, Helsinki, Finland; 5) Department of Clinical Genetics, University of Helsinki and Helsinki University Central Hospital, Helsinki, Finland; 6) Laboratory of Genome Integrity and Institute of Molecular and Translational Medicine, Palacky University, Olomouc, Czech Republic; 7) Department of Oncology, University of Helsinki and Helsinki University Central Hospital, Helsinki, Finland.

FANCD2 is one of the genes mutated in Fanconi Anemia (FA), a rare genetic disorder characterized by defective DNA repair, with resultant hematological and endocrine abnormalities, developmental disabilities, and increased susceptibility to cancer. FANCD2 functions at a critical juncture in the FA pathway, downstream of the main FA complex, where it facilitates cell cycle control, efficient replication and DNA repair. We have used immunohistochemistry to study FANCD2 protein expression in relation to the clinical, prognostic, and histopathological characteristics of 1238 invasive breast carcinomas from the Finnish population. FANCD2 protein expression, scored as % of positive tumor cells in a given sample, associated with high grade, estrogen and progesterone receptor negativity, HER2 amplification, and positive P53 immunohistochemistry, as well as high Ki67 level ($p < 0.001$ for all variables; Kruskal-Wallis test), but not tumor size or lymph node metastasis. High FANCD2 expression was also associated with the formation of gamma-H2AX nuclear foci ($p = 0.005$), a sensitive marker of DNA damage. Thus, FANCD2 associates with markers of DNA damage and proliferation and characteristics of aggressive tumors. In survival analyses, using distant metastasis or death from breast cancer as endpoint, the highest quartile of FANCD2 expression (approximately > 30% of cells positive) was associated with poor outcome ($p < 0.001$, HR $\approx$ 2). Despite its correlation with several markers of poor prognosis, FANCD2 expression remained an independent prognostic factor in multivariate analysis (Cox's proportional hazards model) adjusted for hormone receptor status, grade, tumor size, nodal metastasis and HER2 status. In the univariate analyses the association with survival was strongest among ER-positive cases ($p < 0.001$, HR $\approx$ 3) and tumors belonging to the luminal A subtype (ER/PR+, HER2-) ($p < 0.001$, HR $\approx$ 3.5), generally a category of favorable prognosis. In contrast, no survival effect was seen among ER-negative or HER2-positive cases. No difference was observed between treatment-based subgroups (anthracycline, methotrexate, endocrine therapy). Based on these findings, it is possible that FANCD2 immunohistochemistry could be used to identify a subset of luminal A breast cancers with a higher risk of metastatic progression and death. Further investigation will be required to elucidate the underlying mechanisms that connect FANCD2 expression to breast cancer progression.

1172F

Effects of Cinnamomum cassia extracts on cytotoxicity, apoptosis and gene expression in HL-60 cells. *C. Gündüz[1], S. Yilmaz Süslüer[1], B. Agrap[2], C. Biray Avci[1], F. Lermioglu[2].* 1) Ege University Medical Faculty Dept. of Medical Biology, Izmir, Turkey; 2) Ege University Faculty of Pharmacy Dept. of Toxicology, Izmir, Turkey.

Cinnamon is widely used in food industry and through volatile oils used in perfumery. It also reduces blood glucose and cholesterol levels. In vitro and in vivo studies suggest that cinnamon barks and leaves have antioxidant, cytotoxic, antidiabetic, antimicrobial and antitumor effects. It was shown that Cinnamomum cassia caused to mitochondrial transmembrane loss and death of Human promyelocytic leukemia cells (HL-60) by increasing caspase-3 activity. We aimed to evaluate the effects of Cinnamomum cassia bark extracts on cytotoxicity, apoptosis and the expression changes of apoptosis pathway and cell cycle genes in HL-60 cell line. Methanolic extract of C. cassia was fractioned with hexane, chloroform, ethylacetate and butanol. Cytotoxic effects of the cinnamon extracts were investigated at 24 hours intervals for three days by using WST-1 assay and the apoptotoic effects determined by using Annexin V and JC-1 assays. As a result of cytotoxicity test, the hexane extract of C. cassia was found to be the most effective extract (IC_{50}: 4.22μg/ml) and this extract was used in apoptosis and gene expression analysis. The hexane extract of cinnamon induced apoptosis and also showed significant cytostatic effect. We found increased expression of genes which are responsible for cell cycle pathway as: CCNG2, CCNG1,CCNH, CCNF, ATR, CDK5R1, CDKN2D, CDK5RAP1, CHEK1, CHEK2 and BAX and increased expression of genes which induces apoptosis were found as: TNFRSF25, TNFRSF1A, CASP4, CASP1, TRADD, CD27, CD40LG, FAS, TRAF2. Hexane extract of cinnamon are considered be effective in the treatment of leukemia, however further studies are needed to elucidate the molecular mechanism of these genes' expressions at protein levels.

1173W

Universal Lynch syndrome screening result notification: Efficacy affected by method. *H. Hampel[1], I. Lattimer[1], W. L. Frankel[2].* 1) Internal Med/Div Hum Gen, Ohio State Univ, Columbus, OH; 2) Pathology, Ohio State Univ, Columbus, OH.

Background: Universal screening of all colorectal cancer (CRC) patients for Lynch syndrome (LS) by immunohistochemistry (IHC) has been ongoing at Ohio State University (OSU) for 6 yrs. Genetic counseling and testing of all patients with abnormal IHC is critical to the success of this program. Methods: OSU has utilized 3 different methods of contact for notification of patients with abnormal IHC results. This study addresses the relative efficacy of these approaches. Results:

	MD Referral	Genetics Phone Contact	Genetics In Person Contact
Time Period	10 mos	7 mos	17 mos
Abnormal IHC	24 (17%)	91 (20%)	36 (22%)
Probable methylated cases	11	16	16
# Needing Contact	13	75	19
Genetics Consult	0	47/ 75 (63.5%)	16/ 19 (84.2%)
Lynch syndrome dx	0	8 (1.8%)	5 (3%)

Conclusion: In person notification has been the most effective notification method in our universal LS screening program and results in the expected number of LS diagnoses (3% observed; 2.8% expected).

1174T

Allele-specific expression imbalances of MCC and reduced susceptibility to colorectal cancer in schizophrenia. *L. He, G. He.* Bio-X Institutes, Shanghai Jiao Tong Univ, Shanghai, China.

To date, evidence has indicated that the incidence of colorectal cancer (CRC) among schizophrenia patients is lower than normal. Genetic research may help to elucidate this longstanding phenomenon. To further investigate this apparent protective effect, we employed an innovative strategy combining an association study with ASE analysis to examine genetic variants of the MCC gene. Four polymorphisms (rs9122, rs2112452, rs2227948 and rs2227947) were selected within the coding/3' UTR region of MCC and were genotyped in 312 CRC patients, 270 schizophrenia patients and 270 healthy controls. As a continuous variable, only ASE values of rs2227948 and rs2227947 presented marginal differences between CRC patients and controls, or between schizophrenia patients and controls when applying the Wilcoxon test. All informative participants were allocated either to an ASE or a non-ASE group on the basis of ASE cut-off points, and higher frequencies of ASE in both MCC and individual polymorphisms were consistently found in CRC patients compared to schizophrenia patients or normal controls. We provide a first indication that imbalances of ASE in MCC might confer a reduced genetic susceptibility to CRC in individuals with schizophrenia. The advances in ASE analysis promise to shed more light on the relationship between schizophrenia and cancer progression.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1175F

RAD51 polymorphisms and breast cancer risk. *M. Hosseini[1], M. Houshmand[2], A. Ebrahimi[3].* 1) Dept Sci Islamshahr Branch, Islamic Azad Univ,Islamshahr Branch, Tehran, Iran; 2) National Institute for Genetic Engineering and Biotechnology (NIGEB), Tehran, Iran; 3) Special Medical Center, Tehran, Iran.

Breast cancer (BC) is one of the most common causes of death among women, and second in Iran. The objectives of this study were to determine the frequency of RAD51 G/C polymorphism in patients with breast cancer. We evaluated these two common polymorphisms and their effects on the folate intake and breast cancer risk association in a Iranian sporadic breast cancer population-based case- control study of 294 breast cancer cases and 315 controls using a PCR-RFLP-based assay. Analyses of affected and controls show that homozygote genotype RAD51 GG has the highest frequency in both groups (33.3 in patients and 41.4 in control group). Genotype RAD51 GG most risk factor were in our population: [CC / GC odds ratio, 0.364 (95% confidence interval; CI, 0.168–0.788) p=0.009, CC / GG odds ratio, 0.828 (95% CI, 0.411–1.668) p=0.596], GG/ GC odds ratio, 2.276 (95% CI, 1.497–3.460) p=0.001]. There was a significant association of breast cancer risk with RAD51 GG and CC polymorphism.

1176W

Design and development of a next-generation sequencing assay for multiplex detection of somatic mutations in FFPE tissues. *E.B. Jaeger, R. Haigis, M. Bauer, J. Yeakley, J. Betley, N. Udar.* Illumina, San Diego, CA.

Current methods for detecting somatic mutations are limited by the detection sensitivity of low-frequency mutations present in highly heterogeneous tumors, or by the number of loci that can be queried in a single assay. Next-generation sequencing provides an opportunity to overcome these challenges. We have developed an assay to survey over 17,500 nucleotides in cancer-associated genes commonly mutated in a variety of solid tumor types. When paired with ultra-deep next-generation sequencing, this assay is capable of detecting mutant alleles at frequencies below 5% of sample input. DNA was extracted from 50 FFPE samples from multiple tissue types using the Qiagen AllPrep kit and the QiaCube extraction system. Genes of interest were targeted using an extension and ligation based amplicon library prep assay, and sequencing libraries were run on the Illumina HiSeq. Average depth of coverage was 6000X, and 89% of assays had at least 1000X coverage. Sequencing reads were aligned to a pseudogenome constructed from targeted sequences in hg19. Background relative to the hg19 reference genome was calculated in 30bp windows across each amplicon. The highest noise from each amplicon was set as the background for that target, and base changes exceeding 5SD of background were considered variant nucleotides. Key variants were confirmed by Sanger sequencing, the Qiagen KRAS PCR Kit, and by complementary library preparation assays. Fifty samples were extracted and processed for sequencing. Forty-five samples yielded sufficient product for sequence analysis. Seven of twenty-two colorectal samples showed mutations in KRAS, as did three of twenty lung samples. Mutant allele frequencies ranged from 1% to 56%, all of which were verified using the KRAS PCR Kit. A subset of these mutations, as well as others identified in BRAF and TP53, could only be confirmed consistently by Sanger sequencing when the mutant allele frequency exceeded approximately 30%. We have developed a targeted extension and ligation based library prep assay for detection of somatic mutations that was applied successfully to 90% of FFPE samples tested. The depth of coverage achieved permits detection of mutant alleles at frequencies as low as 1%. A subset of the mutations could be confirmed by alternate methods, including sensitive allele-specific PCR. With continued development of this assay, we expect it to be capable of similar performance at the >17,500 other nucleotides it surveys.

1177T

GENOTYPIC DISTRIBUTION OF MMP1 (1607, 1G/2G) PRAMOTER IN CHILDHOOD MALIGNANT GLIOMAS. *Priyanka Kawal, Raj Kumar, Anil Chandra, T.N. Dhole T.N, B.K. Ojha. P. Kawal.* neurosurgey, sgppgims, lucknow, u.p, India.

Background: Glioblastoma multiforme & anaplastic Astrocytoma are highly malignant central nervous system tumor, which are extremely refractory to therapy due to their rapid growth & local invasive potential. Several members of the MMP family are thought to contribute to their invasiveness. Objectives: we aimed to investigate the association of MMP-1(-1607, 1G/2G) gene polymorphism with malignant gliomas in pediatrics. Methods: In the present case control study, we enrolled a total of 50 cases (paediatric) of Malignant gliomas confirmed by histopathology and 84 healthy children as control. Polymorphism for MMP1 gene (-1607, 1G/2G) was genotyped by restriction fragment length polymorphism. Results: Frequencies of MMP-1 (-1607, 2G/2G) genotypes & 2G alleles were significantly associated with the cases of malignant Gliomas (22%) in relation to healthy controls (7.1%).[OR= 5.7; p= 0.006; 95%CI= (1.63–19.91)] & [OR=1.7 p=.039; 95% CI= (1.03–3.33)].The Distribution of MMP-1 (-1607 2G/2G) genotype was higher in patients with the nonvegeterian food habit of the Malignant glioma with borderline significance in comparison to vegetarian controls [OR=4.12; p=0.051; 95%CI= (0.99–17.14). However, interaction of MMP1 (-1607, 2G/2G) genotypes with respect of the grades (Anaplastic Astrocytoma group/GBM group = 25/25) of gliomas did not modulate the brain tumor risk significantly [OR=1.6; p= 0.55; 95%CI= (0.28–9.97)]. Conclusions: MMP-1 (-1607, 2G/2G) genotypes & 2G alleles were significantly associated with malignant gliomas in pediatric population. The borderline significance was found with the nonvegeterian food habits of the subjects, but no association of MMP-1 was found with the respect of the grades of gliomas. Keywords: Gliomas, polymorphism, genotypes.

1178F

Antileukemic effect of Paclitaxel in combination with Metformin in HL-60 cell line. *C. Kayabasi[1], C. Biray Avci[1], S. Yilmaz Susluer[1], Z.O. Dogan Sigva[1], T. Balci[1], G. Saydam[2], C. Gunduz[1].* 1) Medical Biology Department, Ege University Medical Faculty, Izmir, Turkey; 2) Hematology Department, Ege University Medical Faculty, Izmir, Turkey.

Acute myeloid leukemia (AML) is the most common myeloid leukemia. Paclitaxel is obtained by extraction of Taxus brevifolia bark. This compound is being produced through acetylation of 10-deacetylbaccatin. It blocks the cell cycle in G2/M phase. Metformin (N,N-dimetilbiguanid) is a drug that used for treatment of overweight patients with type 2 diabetes. Furthermore it is established that, metformin inhibits both in vivo tumorigenesis and in vitro cancer cell proliferation. The aim of the study was to evaluate the synergistic effect of metformin and paclitaxel on the expression of genes including several signal transduction pathways such as; cell cycle, apoptosis, cytokines and costimulator molecules and NF-κB JNK/p38 and MAP/MAPK, PI3K-Akt-mTOR, Toll-like receptors in HL-60 cells. IC_{50} (72nd hours) doses of paclitaxel and metformin were calculated as 10,27 nM and 27,94 mM, respectively. In the cells that treated with combination dose (4μM metformin and 34nM paclitaxel), the percentage of viability was found 44,00% (48th hours), 13,64% (72nd hours) with trypan blue technique and cytotoxicity was found 69,57% (48th hours), 49,47% (72nd hours) by using WST assay. Apoptotic effect of the combination dose was detected 50,00% (48th hours), 77,36% (72nd hours) with Annexin V method and 40,32% (48th hours), 87,25% (72nd hours) with JC-1 method. At 48 hours, expression of genes that are responsible for cell cycle and apoptosis pathway were studied by real time online RT-PCR and were shown that upregulation of TNFRSF25, CASP1, CASP5, CD27 genes were induced apoptosis in AML. Also upregulation of TLR3, IFNG, TNFRSF11A,FKBP1A; downregulation of MAP3K7, NFKB1, NFKBIA and overexpression of ERN1 were significant. In conclusion, if these results are corroborated, the combination of paclitaxel metformin can be effective cure in AML treatment.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1179W

The X-linked Tumor Suppressor TSPX Disrupts the Positive Feedback Loop of the Viral Oncoprotein HBx in HBV-associated Hepatocarcinogenesis. *T. Kido, Y.-F.C. Lau.* Department of Medicine, VA Medical Center-111C5, University of California, San Francisco, San Francisco, CA.

Hepatitis B virus (HBV) infection is a major risk for hepatocellular carcinoma (HCC), the third deadly cancer in the world. Since billions are infected with HBV, this is an important global health problem. HBV encodes the X-factor protein (HBx), a putative oncoprotein. HBx has been shown to promote oncogenesis in multiple pathways in HBV-infected liver cells. HBx promotes viral replication, enhances HBV promoter activity and its own transcription, thereby exacerbating the disease development in a positive feedback manner. Hence, HBx persistence is an important factor in HBV-mediated carcinogenesis. The X-linked tumor suppressor TSPX is a homologue of the Y-linked proto-oncogene TSPY, harboring a unique long Asp/Glu-rich acidic domain on its C-terminus, absent in TSPY. The absence or presence of this acidic domain is important for the respective TSPY and TSPX functions in oncogenesis. In the present study, we demonstrate that TSPX is capable of disrupting the positive feedback loop of the HBx oncogenic actions at two levels. First, TSPX promotes HBx proteosomal degradation via its acidic domain, which interacts with both HBx and a proteasome 19S lid subunit, RPN3. Such interactions abrogate the RPN3 dependent stabilization of HBx, suggesting that TSPX and RPN3 act competitively in regulating HBx stability and proteosomal degradation. Second, a luciferase reporter assay shows that over expression of TSPX abolishes the HBx mediated up-regulation of HBV promoter activity in cultured HEK293 and Huh7 cells, thereby repressing the HBx expression. Since loss-of function mutations and/or epigenetic repression of X-located tumor suppressor gene(s) could significantly predispose males to oncogenesis, our data suggest that TSPX could play key roles in tumor suppression in HBV-mediated initiation and progression of HCC, and inactivation of TSPX could contribute to the male prevalence in HBV-associated liver cancer.

1180T

Comparison of minor groove binding ligands and known carcinogen blastomogenic, recombinogenic and mutagenic activity revealed by SMART in wts/+ hetesozygous flies. *K. Kirsanov, E. Lesovaya, G. Belitsky, M. Yakubovskaya.* Chemical carcinogenesis, Blokhin Cancer Research Center, Moscow, Russian Federation.

BACKGROUND: Loss of heterozigocity (LOH) through homologous recombination, causing manifestation of a recessive mutant allele, has been shown to represent a frequent DNA rearrangement in tumors of patients with inherited retinoblastoma and Li-Fraumeni syndromes. Induction of LOH by different chemicals may be detected easily using the test for somatic mutation and recombination in Drosophila (SMART) which is based on the manifestation of recessive mutant allele after inactivation of wild type one due to somatic point mutation, chromosomal deletion or loss of heterozygocity. In our previous work we revealed for the first time strong blastomogenic effect of several minor groove binding ligands (MGBLs), including bisbenzimidazoles Hoechst 33342, Hoechst 33258 and their derivatives widely used in molecular biology, cytology and fluorescent microscopy. In our experiments all of the bisbenzimidazoles produced significant increase of wtsP4 clone frequency at non-toxic concentrations in SMART.Evaluation of their genotoxicity is complex due to the mode of its action. They neither alkylate DNA nor form interstrand-crosslinks or DNA adducts, but bind DNA via electrostatic, hydrogen and van der Waals bonds and inhibit some enzymes, causing genetic damage in an indirect way. RESULTS: In this study we demonstrated that the bisbenzimidazoles Hoechst 33342, Hoechst 33258 and their derivatives realize their genotoxic effects via HR-mechanism exclusively. Thus, the agents were inactive in SMART when HR was completely blocked by TM3 balancer and showed weaker blastomogenic properties in geterozygotes wts/Mus309 when HR was reduced by inactivation of dmBLM enzyme involved in somatic recombination. We estimated activity of the other AT-specific minor groove binding ligands, Berenil, Pentamidine and DAPI. Surprisingly, all of the MGBLs studied revealed no activity either in Somatic Mutation and Recombination Test on Drosophila or in Ames Test on Salmonella in vivo, and all of them were not able to interrupt the TopoI breakage/reunion cycle in vitro. CONCLUSIONS: Overall, our data support the idea that genotoxicity of the recombinogenic agents may be realized via DNA-dependent enzymes, causing genetic damage in an indirect way. Our investigation of MGBL genotoxic properties demonstrates the necessity of recombinogenic activity screening in short term assays. Drosophila SMART assay using wtsP4/+ heterozygous larvae seems to be promising in this way.

1181F

GR-mediated regulation of gene expression by standard glucocorticoids and selective glucocorticoid receptor agonists (SEGRA) in lymphoma cells. *E. Lesovaya[1], A. Yemelyanov[2], K. Kirsanov[1], V. Gasanova[1], M. Yakubovskaya[1], I. Budunova[2].* 1) Blokhin Cancer Research Center RAMS, Moscow, Russian Federation; 2) Northwestern University, Chicago, IL, USA.

Glucocorticoids provide one of the least toxic mainstay therapies for many types of hematological malignancies; however chronic treatment with steroids results in numerous adverse effects. Glucocorticoid receptor (GR) regulates gene expression via transactivation that requires GR homodimer binding to gene promoters; and transrepression mediated via negative interaction between GR and other transcription factors as well as binding with negative glucocorticoid response elements (nGRE) in genes. GR transactivation is linked to metabolic side effects, while GR transrepression underlies glucocorticoid therapeutic action. Novel GR modulator Compound A (CpdA) prevents GR dimerization and transactivation, specifically activates GR transrepression, and has fewer side effects compared to glucocorticoids. The major goals of this study were to evaluate the anti-lymphoma potential of novel GR modulator CpdA and to test whether BZ will enhance CpdA ligand profile and increase its therapeutic potential in leukemia and lymphoma cells. The study was conducted on human T (CEM) and B (NCEB) lymphoma cells expressing endogenous non-mutated GR, and their counterparts with blocked GR expression (cells were infected with shGR lentivirus). Primary lymphoma and leukemia cells were used for gene expression analysis as well. We found that CpdA strongly inhibited growth and viability of lymphoma cells with endogenous GR expression. In contrast to GCs, CpdA did not induce GR transactivation evaluated by RT-PCR of known GR-target genes such as FKBP51, MKP1 and GILZ. It is known that GR expression is regulated by proteasome. We showed that pretreatment of CEM and NCEB cells with proteasome inhibitor Bortezomib resulted in GR accumulation, and enhanced CpdA ligand properties further shifting GR activity towards transrepression evaluated by inhibiton of NF-kB and AP-1 factors. We also revealed remarkable GR-dependent cooperation between CpdA and BZ in suppressing growth and survival of primary leukemia and lymphoma cells. Overall, our data provide the rationale for novel GR-based therapy for lymphoma chemotherapy using combinations of non-steroidal GR modulators with proteasome inhibitors. The detailed mechanisms of their anti-cancer effects in transformed lymphoma cells are under investigation.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1182W

Sub-cellular localization of Y-Box Protein 1 regulates proliferation, migration and tumorigenicity in astrocytomas. *X. Liu[1], D. Faury[1], C. Sollier[1], N. Gerges[1], B. Meehan[4], Z. Dong[2], P. Siegel[2], A. Korshunov[3], S. Pfister[3], J. Rak[4], N. Jabado[1,4].* 1) Montreal Children's Hospital Research Institute, McGill University, Montreal, Quebec, Canada; 2) Rosalind and Morris Goodman Cancer Research Centre, McGill University, Montreal, Quebec, Canada; 3) Division of Molecular Genetics, German Cancer Research Center, Heidelberg, Germany; 4) Department of Pediatrics, Montreal Children's Hospital, McGill University Health Center, Montreal, Quebec, Canada.

Y-Box-Protein-1 (YB1) is a transcriptional and translational regulator implicated in cancer progression. Previously, we established elevated YB1 levels in pediatric Glioblastoma (GBM), an aggressive high-grade astrocytoma, possibly driving oncogenesis in this cancer. Increased YB1 expression has also been detected and found to correlate with unfavorable outcomes in a wide range of other human cancers. The purpose of this study is to investigate the role of YB1 in astrocytoma genesis and its possible association with poor prognosis. We overexpressed or silenced YB1 protein in pediatric GBM (SF188), adult GBM (U87) and normal human astrocytes cell lines (NHA immortalized with H-Tert). Proliferation, migration, soft agar assay were performed in vitro, accompanied by in vivo mice xenograft to assess the tumorigenic ability of these cells. Meanwhile, tissue microarrays including 150 pediatric GBM and 70 Grade I Pilocytic Astrocytoma were stained immunohistochemically (IHC) with YB-1 and EGFR to assess their expression and correlation. YB-1 silencing using shRNA reduced YB-1 level by 90%, and surprisingly increased cell proliferation. As YB-1 is predominantly cytoplasmic in physiological conditions in cells, and nuclear YB-1 is known to associate with increased cell growth, sub-cellular localization of YB-1 was investigated. In the YB-1 silenced clones, YB-1 was greatly reduced in cytoplasm, while YB-1 in the nucleus and EGFR expression was actually enriched, explaining the increase in cell proliferation. YB1 overexpression was mainly cytoplasmic, and decreased cell proliferation and increased migration. Importantly, in orthotopic mice injections, YB-1 overexpression decreased the tumorigenic ability of GBM cells. IHC on tissue microarrays of patient tumors showed strong YB1 expression in 66% of pediatric GBM samples, but only 8% in Grade I astrocytoma, suggesting a role of YB1 overexpression in a more invasive tumor phenotype. In addition, we showed that nuclear YB1 expression is associated with EGFR overexpression in patient samples. Our results suggest that YB-1 modulates cellular proliferation and migration, based on its sub-cellular localization. Nuclear YB-1 drives cell proliferation, whereas cytoplasmic YB-1 promotes cell migration. Showing the association of nuclear YB-1 with EGFR overexpression and more aggressive tumor phenotype, our data argue for caution in targeting YB1 for therapeutic intervention.

1183T

Integrative analysis of TCGA clinical and genomic data reveals conflicting roles of REST gene expression in recurrent glioblastoma. *J. Madhusoodanan, M. Shekar, J. Su, I. Kupershmidt.* NextBio, Santa Clara, CA.

Background: Increased expression of the NRSF/REST transcriptional regulator is implicated in maintaining tumor cell pluripotency in recurrent glioblastoma multiforme (GBM). Although several studies support this oncogenic role of increased REST expression, little is known about the diagnostic or prognostic significance of this perturbation in GBM patients. We evaluate the clinical significance of increased REST expression in GBM patients and identify novel biomarkers of prolonged recurrence-free survival. Methods: Publicly available genomic and clinical data from TCGA and experimental data sources were curated, normalized and analyzed using the NextBio Clinical platform. Genomatix software suite was used for promoter sequence analysis. GBM cohort of TCGA with increased expression of the REST gene (n =333) were screened for enrichment of known prognostic biomarkers. Results: In the NextBio Clinical module, 97% of GBM patients with increased REST expression were found to have increased methylation of the MGMT gene, a biomarker strongly linked to longer survival and better responses to temozolomide and radiotherapy. In silico analysis of the MGMT promoter sequence reveals conserved RE-1 binding sites suggestive of an association between increased REST expression and MGMT methylation. This cohort was further analyzed for other potential biomarkers predictive of disease recurrence. Comparing patients with increased REST expression who survived less than 3 mos. with those surviving 1 year or longer, we identified VWC2 gene expression as being significantly different between the two groups. Decreased VWC2 expression appears strongly associated with longer recurrence free survival. Only one study so far suggests VWC2 may be a marker of cancer stem cells, and a few others propose a proliferative role in neurogenesis and neuronal development in animal models. Conclusion: We identify an apparently contradictory association between increased REST expression, considered an oncogenic change in GBM, and MGMT promoter methylation, a biomarker strongly linked to improved prognosis in GBM. Within the cohort with increased REST expression, novel biomarkers for prolonged recurrence-free survival were identified, such as decreased expression of VWC2. VWC2 is associated with neuronal proliferation in mice and may serve as a putative biomarker for recurrent GBM.

1184F

Analysis of *MRE11/RAD50/NBN* genes in childhood acute lymphoblastic leukemia. *J. Nowak[1], M. Mosor[1], I. Ziolkowska-Suchanek[1], K. Nowicka[1], D. Januszkiewicz-Lewandowska[1,2].* 1) Inst Human Gen, Polish Academy Sci, Strzeszynska 32, Poznan, Poland; 2) University of Medical Sciences, Fredry 10, Poznań, Poland.

The *MRE11, RAD50* and *NBN* genes encode MRE11-RAD50-NBN (MRN) complex proteins involved in DNA damage cellular response. In our previous study it was showed that germline I171V mutation in *NBN*, may be considered as a risk factor in the development of childhood acute lymphoblastic leukemia (ALL). The aim of the study was to determine whether *MRE11* and *RAD50* alternations have a potential role in childhood ALL. The study was carried out by determining the frequency of constitutional mutations and polymorphisms in selected regions of *MRE11, RAD50*, and *NBN* in the group of 200 children diagnosed with ALL. It was the first time we have carried out simultaneous analysis of *MRE11, RAD50* and *NBN* genes in childhood ALL. Anonymous blood samples collected on Guthrie cards were used as a control group. The analysis was focused on exon 5, 10, 14, 15, 19 of *MRE11*, exon 3, 4, 5 and 7 of *RAD50* and 2, 5, 6, 7, 10, 13 of the *NBN* gene and was performed by specific amplification of region of interest by PCR followed by multi-temperature single-strand conformation polymorphism (PCR-MSSCP) technique. Each sample showing shifts in MSSCP analysis was sequenced. In a sample of the 200 children diagnosed with ALL we did not identify any mutations in the *RAD50* or in the *MRE11*. We identified 2 alternations in 3 out of 506 control group members in the *RAD50* gene - the V127I in exon 4, detected twice, was predicted to be tolerated using the SIFT and PolyPhen analysis. The V315L in exon 7 occurring only once among 506 control group members was predicted to be benign. In addition we were able to detect c.551+19 G>A single nucleotide polymorphisms in the intronic sequence of exon 4 with different frequencies both in ALL patients and controls. The frequency of the AA genotype was higher in leukemia patients, as compared to control group members (p= 0,0032). In the group of ALL patients the higher incidence of the I171V mutation in ALL group (5/200) than among controls (12/2400)(p< 0,0007)was observed. Although variant allele of the intronic SNP in the *RAD50* gene was observed as more frequent in childhood ALL in comparison to control group members, no clear association of mutations in *RAD50* and *MRE11* gene has been seen. Interpretation of these results is limited small sample group number. Nevertheless we confirmed participation of I171V germline mutation of *NBN* gene in ALL.

1185W

Validation of a prostate cancer metastasis signature in a FFPE cohort of primary tumors. *A. Pearlman[1], C. Campbell[1], J. Loke[1], S. Freedland[2], Y. Shao[3], H. Ostrer[1].* 1) Dept of Pathology, Albert Einstein College of Medicine, Bronx, NY; 2) Dept of Pathology, Duke University Medical Ctr, Durham, NC; 3) Division of Biostatistics, NYU School of Medicine, New York, NY.

Genomic copy number alterations (CNAs) resulting in changes in abundance or altered structure of RNA transcripts and proteins (e.g. truncating dominant negative mutations) may impact the fitness of the cell and provide some of the mechanisms necessary for distant site migration, invasion and growth. From multiple CNAs identified in fresh frozen tumors of radical prostatectomies, CNA-based gene signatures were developed into a metastatic potential score (MPS) that suggested the ability to predict metastasis free survival (in press, J. Prob. and Statistics). To assess the validity of our metastasis signature and MPS prediction model, we collected a retrospective cohort of primary prostate cancer tumor and matched normal tissues from thirty men treated with radical prostatectomy at Duke University Medical Center (Duke cohort). The cancer tissue was obtained from archived formalin fixed paraffin embedded (ffpe) blocks. Each block was processed by our pathologist by obtaining a 5 micron H&E stained sections and evaluated for tumor content. The genomic DNA (gDNA) was extracted using a Qiagen ffpe gDNA column extraction kit. The material was sent to Affymetrix Service Center in Santa Clara, CA and run on the OncoscanTM V2 SNP array developed specifically for gDNA samples extracted from ffpe archived tissue. The Duke cohort was made up of primary tumors that metastasized following radical prostatectomy (mPTs, n=12), and, a group of low risk tumors that didn't develop distant metastases (iPTs, n=7). The receiver operating characteristics-area under the curve analysis (ROC-AUC) applied to the Duke cohort mPTs and iPTs resulted in an accuracy of 0.91.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1186T

Molecular Characterization of Small Cell Carcinoma of the Ovary. *P. Ramos[1], M.L. Russell[1], L. Nordstrom[1], S. Smith[1], M.T. Barrett[1], T. Holley[1], E. Lenkiewitz[1], J.G. Pressey[2], J. Farley[3], S.P. Anthony[4], G. Hostetter[5], A.E. McCullough[6], K.A. Furge[5], B.B. Haab[5], J.M. Trent[1], H.E. Cunliffe[1].* 1) The Translational Genomics Research Institute, Phoenix, AZ; 2) Department of Pediatrics, University of Alabama at Birmingham, Birmingham, AL; 3) Department of Obstetrics and Gynecology, St. Joseph's Hospital and Medical Center, Phoenix, AZ; 4) Evergreen Hematology and Oncology, Spokane, WA; 5) Van Andel Research Institute, Grand Rapids, MI; 6) The Mayo Clinic, Scottsdale, AZ.

Background. Small Cell Carcinoma of the Ovary (SCCO) is a rare and highly aggressive malignancy that affects young women and girls. The majority of patients are diagnosed at an advanced stage and do not respond to chemotherapy, or only experience transient responses, quickly exhibiting therapeutic resistance. As a result, more than 75% of patients succumb to their disease within 1–2 years, even when diagnosed at an early stage. The differential diagnosis of SCCO remains a challenge, as the etiology of SCCO is not known. There is no robust molecular evidence to direct therapy for SCCO patients. **Hypothesis.** Comprehensive molecular characterization of multiple SCCO tumors will define the spectrum of SCCO-specific aberrations, leading to new insights for SCCO pathobiology, rationales for targeted treatment approaches, and markers to aid in differential diagnosis. **Methods.** SCCO tumors were obtained from 16 patient cases. Specimens were obtained following informed consent under an IRB-approved protocol (www.tgen.org/scco). We performed Agilent 44Kv2 gene expression profiling on 4 fresh frozen tumors compared to 1) age-matched normal ovary, and 2) 15 individual ovarian tumors of various histology. Agilent 400K array-comparative genomic hybridization (aCGH) was performed on all cases. A tissue microarray (TMA) was constructed for biomarker validation. **Results.** Consistent with previous reports, the average age at diagnosis of our cohort is 23 years. Expression profiling revealed 8,746 genes significantly differentially expressed between SCCO tumors and normal ovary (p-value <0.001). Ontology analyses indicated cell cycle regulation, chromosome organization, and DNA repair. Several candidate therapeutically actionable targets were identified that are currently being validated by Immunohistochemistry. To compare SCCO profiles to the most common histotypes of epithelial ovarian cancer, we performed an unsupervised 2-dimensional hierarchical clustering of 3526 genes most highly deviating across all tumors. SCCO tumors clustered tightly and distinctly from all other tumors and revealed clusters of genes unique to SCCO. We are currently defining biologic processes identified by these clusters to attain insight into the pathobiology and possibly etiology of SCCO tumors. aCGH analyses indicated SCCO genomes are predominantly diploid. **Conclusions.** We anticipate the outcome of this study will provide new insights into the diagnosis and clinical management of SCCO.

1187F

Effects of *Helicobacter pylori* infection on *MGMT* and *MLH1* promoter methylation status and microsatellite instability in paediatric and adult patients. *M.L. Ribeiro[1,2], M.C. Alvarez[1,2], J.C. Santos[1], N.M. Maniezzo[1], M.S. Ladeira[3], I.C.A. Scaletsky[4], J. Pedrazzoli Jr[1].* 1) Sao Francisco University. Clinical Pharmacology and Gastroenterolog, Braganca Paulista, Sao Paulo, Brazil; 2) Programa de Pós Graduação em Genética e Biologia Molecular, UNICAMP, Campinas, SP, Brazil; 3) Departamento de Clínica Médica, UNESP, Botucatu, SP, Brazil; 4) Departamento de Microbiologia, Imunologia e Parasitologia, UNIFESP, SP, Brazil.

It is well established that *H. pylori* infection leads to chronic inflammation in the gastric mucosa, which is associated with DNA methylation. Because this epigenetic alteration plays an important role in the regulation of gene expression and maintenance of DNA integrity and stability, the aim of this study was to analyse the effect of *H. pylori* infection on the methylation status of *MLH1* and *MGMT* and the relationship between this methylation and microsatellite instability (MSI) in children infected or uninfected by the bacterium and in adults with chronic gastritis or gastric cancer, both with and without *H. pylori* infection. Biopsy samples from children and adults with chronic gastritis and adults with gastric cancer, infected or uninfected with *H. pylori*, were included in the study. Methylation was not detected in the promoter regions of *MLH1* and *MGMT* in paediatric samples. In chronic gastritis adult patients, methylation of the *MGMT* promoter was associated with *H. pylori* infection. The levels of *MLH1* and *MGMT* mRNA were significantly reduced in chronic gastritis samples that were also hyper-methylated. We observed methylation of the *MHL1* promoter, as well as an increase in the level of MSI, in samples from gastric cancer patients when comparing them to samples from the adult chronic gastritis patients with *H. pylori* infection. Our study indicates that the methylation of the promoter regions of *MGMT* and *MLH1* might be considered an event that is dependent on the duration of infection because such alterations in methylation were not observed in children. Moreover, in gastric cancer patients, the mRNA levels of *MLH1*, the hyper-methylation of the promoter region of *MLH1* and the increase in the frequency of MSI are all related events.

1188W

Role of novel and GWAS identified PLCE1 genetic variants to gallbladder cancer susceptibility in north Indian population. *K. L. Sharma[1], S. Misra[2], A. Kumar[1], B. Mittal[1].* 1) Sanjay Gandhi Post Graduate Institute of Medical Sciences (SGPGIMS), Lucknow, India; 2) Chatrapati Shahuji Maharaj Medical University (CSMMU), Lucknow, India.

Introduction: Phospholipase C epsilon 1 (PLCE1) gene belonging to the phospholipase family plays crucial roles in carcinogenesis and progression of various cancers. Recent genome wide association studies have reported significant association of genetic variants in PLCE1 as novel susceptibility markers for various cancers like esophageal, gastric and head and neck. However, its association is not known in gallbladder cancer. Therefore, the present study aimed to find out the association of GWAS identified 2274223A>G polymorphism from haplotype block 2 and a novel SNP rs7922612T>C from haplotype block1 of PLCE1 in conferring susceptibility to GBC in north Indian population. Methods: Selected PLCE1 variants were genotyped in 416 GBC cases and 225 controls by PCR-RFLP. Statistical analysis was done by SPSS and SNPstats. In-silco analysis was performed using Bioinformatics tools (FAST-SNP, F-SNP). Results: PLCE1 rs2274223 A>G [AG] and rs7922612T>C, [CC] genotypes were found to be significantly associated with an increased risk of GBC [OR= 1.9, p=0.002, OR= 2.0, p= 0.04 respectively] as compared to healthy controls. PLCE1 haplotype [A-C] showed statistical significant association with GBC [OR= 1.5, p=0.02]. On stratification based on gender, PLCE1 rs2274223 A>G [AG] was significantly associated with GBC n a sex neutral manner [OR=2.3, p=0.04 (males), OR= 1.8, p=0.01(females)] but rs7922612T>C, [CC] genotypes showed increased risk of GBC in males only [OR=4.5, p=0.03]. On dividing GBC patients on the basis of gallstone status, PLCE1 rs2274223 [AG] was risk factor for GBC patients with or without gallstones [OR; 2.1, p=0.006, (with stone/control) OR; 1.8, p=0.012 (without stone/control)] whereas rs7922612 CC genotype imparted higher risk for GBC without gallstones when compared to controls [OR; 2.3, p=0.03]. PLCE1 haplotypes [A-C] and [G-T] [OR= 2.1, p=0.03, OR=2.4, p=0.03] respectively showed statistical significant association with GBC patients with stones. The genetic risk was not further modulated by tobacco consumption or age of onset. In-silico analysis of rs2274223 A>G showed significant change in functional category i.e. protein coding and splicing regulation [SNPeffect: deleterious, ESEfinder: change, ESRSearch: change; (FS score 0.3); FAST-SNP: Low-Medium risk]. Conclusion: The present study found significant association of PLCE1 rs2274223 and rs7922612 polymorphisms with GBC susceptibility.

1189T

Serum metabolomics and prostate cancer survival. *R. Szulkin[1], R. Karlsson[1], A. Heuberger[2], M. Hong[1], C. Broeckling[2], J. Prenni[1], J. Prince[1], F. Wiklund[1].* 1) 1Department of Medical Epidemiology and Biostatistics, Karolinska Institute, Stockholm, Sweden; 2) Department of Biochemistry and Molecular Biology, Colorado State University, Fort Collins, USA.

Introduction: Established prognostic factors perform poorly in predicting disease relapse among patients treated for prostate cancer. Identification of novel biomarkers improving the prognostic information is of great importance to guide individual therapy. Materials and Methods: Post-treatment serum samples from a nested case-case design comprising 269 prostate cancer patients with lethal outcome and 269 patients with non-lethal outcome were used. All patients were diagnosed between year 2001 and 2003 in Sweden and followed up for survival until December 2010 through record linkage with the national Cause of Death Registry. Untargeted gas- and ultra performance liquid chromatography (GC, UPLC) coupled with mass spectrometry (MS) were employed to screen for novel prostate cancer biomarkers. Normalized and log-transformed metabolite concentrations were explored for association with prostate cancer-specific survival in time-to-event analysis using death from prostate cancer as endpoint. Results: Untargeted metabolomic profiling of prostate cancer serum samples revealed a total of 9836 GC/MS and 5301 LC/MS profiles. Univariate analysis of individual normalized feature levels indicated 23 peaks to be study-wide significant associated with prostate cancer-specific survival. Of note, at the 1×10-8 significance level we observed 11 associated peaks as compared to 1×10-4 expected peaks under the null hypothesis of no association. Further assessment exploring pair-wise ratios between metabolomic peaks revealed additional features significantly associated with prostate cancer prognosis. Conclusion: Untargeted metabolomic profiling of prostate cancer serum samples have identified a considerable number of molecular features strongly associated with disease prognosis. Further analysis is underway to identify these profiles molecular identity and to explore molecular pathways involved.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1190F

Analysis of TCGA patient and curated public genomic data identifies differentially regulated processes in breast cancer metastases. *R. Wisotzkey, A. Umesh, J. Park, J. Shima, J. Delaney, E. Kelly, E. Chiu, M. Shekar, I. Kupershmidt.* NextBio, 475 El Camino Real, Santa Clara, CA.

To identify biological processes involved in breast cancer metastases, public genomic data was analyzed using NextBio's Research platform. Meta-analysis of datasets comparing primary tumors that metastasized (M1) versus primary tumors that did not metastasize (M0) identified biological processes, or biogroups, that are significantly differentially regulated. For example, compared to primary tumors from M0 patients, tumors from M1 patients had upregulation of cell cycle-related genes, while those involved in the immune response and extracellular matrix were predominantly downregulated. Amongst the genes involved in the immune response, meta-analysis identified CCR7 and TRAT1 as two of the top-ranked differentially regulated genes, with a fold change of -5.6 and -3.0 in M1 versus M0 samples, respectively.

We then mined individual patient data from the TCGA breast cancer dataset using NextBio's Clinical platform for the expression of CCR7 and TRAT1. Meta-analysis results were corroborated in TCGA breast cancer patients, where those with M1 disease had downregulated CCR7 (p=1.5E-3) and TRAT1 (p=2.3E-2) compared to those with M0 disease.

Both CCR7 and TRAT1 have potential roles in tumor progression: CCR7 is a potential biomarker for metastasis of gastric cancer, and TRAT1 is associated with patient survival in melanoma patients. Our analysis adds support to these genes having a role in the progression of metastatic breast cancer.

Using NextBio's Research and Clinical platforms to analyze curated public genomic data, we demonstrate that metastatic and non-metastatic tumors can be differentiated based on gene expression patterns. We identify the immune response as one biological process that is important for breast cancer progression, and CCR7 and TRAT1 as specific genes within this category that may constitute a signature for metastatic breast cancer.

1191W

Extreme breast/ovarian cancer phenotypes in non BRCA1 & BRCA2 families of Greek ancestry. *F. Fostira, A.V. Stavropoulou, I. Konstantopoulou, D. Yannoukakos.* NCSR DEMOKRITOS, MOLECULAR DIAGNOSTICS LABORATORY, ATHENS, Greece.

During the last decade, mutations in more than 20 genes have been linked with hereditary breast/ovarian cancer. Despite that fact, BRCA1 & BRCA2 are still the reference genes of breast/ovarian cancer susceptibility. Mean age of onset in BRCA1 & BRCA2 carriers is around 42 years for most populations. However, there are cases diagnosed with breast cancer before the age of 30. We report three probands with extreme breast cancer phenotypes that were initially referred for BRCA1 & BRCA2 genetic testing. Two were diagnosed with breast cancer at the age of 24 and 25 respectively, while one was diagnosed with unilateral breast cancer at the age of 32. All three cases had extensive family history. In more detail, in the first family 9 individuals have been diagnosed with primary cancers (2 breast cancers-ages 24 and 27-, 2 bone cancers-ages 18 and 22-, 1 leukemia, 2 lung, 1 jaw, and 1 liver cancer) in three successive generations. In the second family, 6 individuals were diagnosed with 9 types of cancer. The proband was diagnosed with adrenal and breast cancer at the age of 1 and 25 years, respectively. Other family members were diagnosed with testicular, lung, osteosarcoma, bone, breast cancer and leukemia. All cancers manifested between the first and fifth decade of their life. The third family consisted of 2 sisters diagnosed with unilateral breast cancer at the age of 32 and breast cancer at the age of 37, while their mother developed ovarian cancer at the age of 34. The genetic background of this-at first glance-typical HBOC pedigree became more evident after the observation of melanin spots around the patients' lips and hands, which are one of the characteristic phenotypic features of Peutz-Jeghers syndrome.Genetic analysis revealed two TP53 pathogenic mutations, located on exons 8 and 6 of the gene respectively, and one STK11 pathogenic mutation, located on exon 1 of the gene.This is the first report of a complete characterization, clinical and genetic, of Peutz-Jeghers and Li-Fraumeni syndromes in Greece. Although both hereditary syndromes are individually rare, the underlying genetic defects can explain the breadth of extreme cancer phenotypic expressions, as well as the remarkable accelerated ages of onset. Moreover, in the following years, there will be a certain benefit from the genetic information as more targeted therapies will be available which can be used on the selection of the appropriate treatment for patients who are mutation carriers.

1192T

Modeling carcinogenesis in BRCA1 and BRCA2 heterozygous mammary epithelial cells. *M.E. Keith, M. Tenniswood.* Cancer Research Center, SUNY-University at Albany, Rensselaer, NY.

Mutations in the breast cancer susceptibility genes BRCA1 and BRCA2 account for half of all hereditary breast cancers, and are estimated to account for over 10% of the breast cancer burden in the US. The current standard of care for BRCA1 or BRCA2 mutation carriers is based on limited understanding of specific mutations, and thus treats all carriers as if they have the same relative risk for development of breast cancer. Several lines of evidence suggest that this is not true: BRCA1 and BRCA2 tumors have very different phenotypes, and specific mutations grant certain differential risk profiles for lifetime cancer incidence. Thus, we have created a novel model system of BRCA1 and BRCA2 heterozygous mammary epithelial cell lines, called MCK cells, from women opting for prophylactic mastectomy. These lines have been isolated from pre-neoplastic breast tissue that is not yet transformed. To date, we have collected 12 breast samples, with equal distribution between BRCA1 and BRCA2; this is the only model system which allows for comparisons between mutation sites within the BRCA genes.

The MCK cell lines are being immortalized for long-term use in cell culture. The effects of BRCA1 or BRCA2 mutation and the impact of different mutations within each gene on DNA repair and transformation will be assessed using these immortalized lines. Ionizing radiation is used as a DNA damaging agent, in order to mimic the effects of increased mammographic screening on BRCA mutation carriers. We will also examine the impacts of common dietary chemopreventive agents, such as vitamin D and EGCG, on growth and transformation in the MCK cell lines.

Our MCK panel of mammary epithelial cell lines will fill an immediate need in hereditary cancer research. The ability to compare and contrast mutations in differing sites within each gene will allow us to make more focused recommendations to BRCA mutation carriers, depending on the location of their particular mutation. In this way, we hope to “personalize” the management and treatment of BRCA1 and BRCA2 mutations, to enhance the quality of life of women who harbor deleterious mutations.

1193F

Integrated transcriptional and functional screening identifies novel regulators of lymphoid lineage determination. *E. Laurenti[1], S. Doulatov[1], I. Plumb[1], M. Doedens[1], C. April[2], J.B. Fan[2], J. Dick[1].* 1) Ontario Cancer Institute, Princess Margaret Hospital, University Health Network,, Toronto, Ontario, Canada; 2) Illumina Inc, San Diego, CA.

Introduction: The hematopoietic system is a highly regulated cellular hierarchy, responsible for the day-to-day production of mature blood cells which can be divided in two major lineages, myeloid and lymphoid. In contrast to mice, very few transcription factors (TFs) regulating early lineage decisions have been identified in human hematopoiesis. Materials and Methods: We previously demonstrated that single Multi-Lymphoid Progenitors (MLPs) retain the capacity to differentiate into all lymphoid cell types as well as monocytes and dendritic cells, indicating that molecularly, the choice between any of these fates must be determined at this stage. To identify TFs responsible for lymphoid vs. myeloid commitment, we interrogated MLP gene expression in the context of a high-resolution transcriptome analysis of 10 human hematopoietic stem and progenitor populations. Consistent with their multipotentiality, MLPs displayed balanced stem, lymphoid and myeloid transcriptional programs. We then used a combination of computational approaches to extract 10 lineage-modulating TFs, the function of which we tested in a knock-down screen based on a novel clonal assay for simultaneous read-out of B, NK and myeloid cell. Results: Altered cell fate decisions from single MLPs, in particular decreased commitment towards the B cell lineage, were observed upon knock-down of half of our candidates. Interestingly, 4 of these (BCL11A, BCL6, EBF1, SOX4) have been so far implicated in later stages of murine B cell development, suggesting that these TFs not only contribute to maturation of B cells, but also favor initial commitment. Preliminary data also provides the first report for a function of TEAD1, a transcriptional mediator of the Hippo pathway, in a hematopoietic context. In summary, we have developed a genome-wide strategy which successfully identified novel regulators of adoption of human lymphoid fate in primary human cells.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1194W

microRNA in biofluids - robust biomarkers for disease, toxicology or injury studies - The case of minimally invasive colorectal cancer detection. *P. Mouritzen, T. Blondal, D. Andreasen, N. Tolstrup, M.W. Teilum, A. Baker, S.J. Nielsen.* R&D, Exiqon A/S, Vedbaek, Greater Copenhagen, Denmark.

microRNAs represent the most well described class of small RNAs (21–23nt) and have been shown to function as post-transcriptional regulators of gene expression. The high relative stability of microRNA in common clinical source materials (e.g. FFPE blocks, plasma, serum, urine, saliva, etc.) and the ability of microRNA expression profiles to accurately classify discrete tissue types and specific disease states have positioned microRNA quantification as a promising new tool for a wide range of diagnostic applications. Furthermore microRNAs have been shown to be rapidly released from tissues into the circulation with the development of pathology. We are currently screening for and validating microRNAs as biomarkers for cancer with the aim of developing minimal invasive tests that may be applied in early detection population screens. Procedures have been developed to control pre analytical variables, for quality checking and qualifying biofluid samples in particular serum and plasma but also urine and other biofluids. In this process we have developed a genome-wide LNA™-based microRNA qPCR platform with unparalleled sensitivity and robustness even in biofluids where microRNA levels are extremely low. Only a single RT reaction is required to conduct full miRNome profiling thereby facilitating high-throughput profiling in important clinical sources without the need for pre-amplification. Thousands of biofluid samples have been profiled including blood derived plasma/serum and urine to accurately determine normal reference ranges for circulating microRNAs in several biofluids. This has allowed development of qPCR panels containing the relevant subset microRNAs present in various biofluids together with tissue specific microRNA markers. The panels will support development of robust biomarkers in disease, as well as toxicology and injury studies. An extensive QC system has been implemented in order to secure technical excellence and reveal any unwanted bias in the dataset. We will demonstrate individual behaviour of serum and plasma, with a focus on how hemolysis and cellular contamination can affect miRNA profiles. Furthermore, we will demonstrate that when close attention is focussed on these important pre-analytical parameters the system can be quickly and robustly applied in biomarker discovery and validation projects using the specific case of colorectal cancer early detection in blood.

1195T

Detection of KRAS, BRAF, NRAS and PIK3CA cancer markers by castPCR technology. *J. Au-Young[1], D. Keys[2], B. Ching[2], M. Mouanoutoua[2], D. Merrill[1].* 1) Medical Sciences, Life Technologies, Foster City, CA; 2) Genetic Analysis, Life Technologies, Foster City, CA.

Mutation detection is becoming increasingly important for making therapeutic decisions in oncology. Companion diagnostic tests can dictate whether cancer drugs such as tyrosine kinase inhibitors (TKIs) will be effective. Efficacy of TKIs is dependent on the presence of activating EGFR mutations. However, efficacy of TKIs is also dependent on absence of activating KRAS mutations as well as absence of EGFR drug resistance mutations. Thus, mutation detection tools are needed that have high sensitivity and high specificity. DNA sequencing and standard genotyping are limited to 5–20% sensitivity for detecting mutations in tumor samples. We have recently developed TaqMan® Mutation Detection Assays using competitive allele specific TaqMan® PCR (castPCR) technology for cancer mutation research. We have previously established that castPCR can detect mutations in the background of wild type with results that show the technology can robustly detect mutations as low as 0.1% with >99% concordance to other technologies including PCR-based technology and sequencing. The high sensitivity enables 10 copies of a specific mutation to be detected in excess of 10,000 copies of wild type gDNA from FFPE tumor samples. KRAS, BRAF, NRAS and PIK3CA exon 20 mutations are associated with low response rates to several chemotherapeutic drugs. In this study, we tested the sensitivity of castPCR for absence of mutations and confirmation of wild type by measuring the ratio or relative percentage of wild-type to mutant allele. We show that castPCR can be used for confirmation of wild type alleles in the absence of specific mutant alleles of KRAS, BRAF, NRAS and PIK3CA. The results obtained by castPCR assays will also be compared to those reported by standard DNA sequencing methods. Our data will demonstrate how castPCR technology can discriminate the mutated vs. wild-type where confirmation of wild type status may predict likelihood of response to targeted therapies.

1196F

Identifying the Causes of Dyskeratosis Congenita: the Prototypical Telomere Biology Disorder. *S.A. Savage, N. Giri, B. Ballew, B.P. Alter.* Div Cancer Epidemiology & Genetics, Natl Cancer Inst, Bethesda, MD.

Dyskeratosis congenita (DC) is an inherited bone marrow failure syndrome (IBMFS) caused by germline defects in telomere biology. The classic triad of dysplastic nails, skin pigmentation, and oral leukoplakia is diagnostic. Individuals with DC are at very high risk of bone marrow failure (BMF), leukemia, myelodysplastic syndrome, and head/neck cancers. Esophageal or urethral stenosis, liver disease, pulmonary fibrosis, developmental delay, and other medical conditions may be present. The clinical consequences manifest at variable ages and in different patterns, even within the same family. Leukocyte telomere lengths <1st percentile for age are diagnostic of DC. The inheritance of DC is variable with X-linked (*DKC1*), autosomal dominant (*TERC, TERT,* or *TINF2*), and autosomal recessive (*NOP10, NHP2*, or *WRAP53*) inheritance patterns. Our IRB-approved longitudinal cohort study of DC includes detailed medical record review, comprehensive questionnaires, and thorough evaluations at the NIH Clinical Center (CC) of affected individuals and their relatives. To date, 56 families with DC have enrolled, including 83 affected, and 204 unaffected relatives. Individuals with short telomeres and features similar to, but not diagnostic of DC, who lack a DC-associated mutation, are classified as DC-like. There are currently 19 DC-like families consisting of 24 affected and 48 unaffected individuals. DC-like families are re-classified as DC if a mutation is identified. Sixty-eight % of DC and 42% of DC-like families have been evaluated at the NIH CC. All probands are tested for germline mutations after genetic education and counseling. A total of 58% of families had a mutation in a DC-associated gene, comprised of 20% *DKC1*, 14% *TINF2*, 12% *TERT*, 9% *TERC*, and 3% *WRAP53*. No mutations were identified in *NOP10* or *NHP2*. Exome sequencing is underway in the mutation-negative DC and DC-like families. We combine clinical features with telomere length data to determine the likely inheritance pattern. Variants are called using a custom-calling algorithm and filtering is based on the expected modes of inheritance. Technical validation of rare variants (<1% in 1000 Genomes and other databases) is performed to rule-out false positive findings. Several variants in biologically plausible genes are currently being evaluated. In addition, genotype-phenotype analysis is ongoing in DC and the spectrum of overlapping telomere biology disorders.

1197W

Germline *BAP1* mutation - Additional case report and expanding clinical phenotype. *R. Pilarski[1], C.M. Cebulla[2], T. Rich[3], L. Strong[3], M.H. Abdel-Rahman[1,2].* 1) Clinical Cancer Genetics, Ohio State Univ, Columbus, OH; 2) Dept. of Ophthalmology, Ohio State Univ, Columbus, OH; 3) Clinical Cancer Genetics, U. of Texas M.D. Anderson Cancer Center, Houston, TX.

Objective: We studied a family with three cases of uveal melanoma (UM), three cases of mesothelioma, and multiple relatives with other cancers for mutations in the *BAP1* gene, which has recently been shown to cause hereditary UM, mesothelioma and other cancers. Results: We identifed a heterozygous germline truncating mutation (p.Q684*) in *BAP1* in two family members and four obligate carrier relatives, supporting the causative role of BAP1 in this family's cancers. We review the literature on reported cases, summarize the expanding spectrum of hereditary cancers associated with germline *BAP1* mutations, and suggest management recommendations for mutation carriers. Conclusion: Germline *BAP1* mutations cause a new hereditary syndrome which predisposes patients to an increasing spectrum of cancers, including UM, cutaneous melanoma and atypical melanocytic lesions, mesothelioma, lung carcinoma, meningioma, and probably other cancers. Our results indicate that *BAP1* testing is indicated in high risk families with UM, mesothelioma, and a range of other cancers.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1198T

Detection of human breast cancer gene fusions from RNA-sequencing analysis of formalin-fixed paraffin-embedded (FFPE) tissue specimens. *Y. Ma, J. Stephans, R. Ambannavar, J. Jeong, J. Morlan, A. Dei Rossi, M. Liu, D. Sinicropi, J. Baker, K. Qu.* Genomic Health Inc., Redwood City, CA.

A fusion gene is a hybrid created by the conjoining of two previously separate genes as a result of genomic aberrations. Oncogenic fusion genes such as BCR-ABL1 in chronic myelogenous leukemia, and TMPRSS2-ERG in prostate carcinogenesis are believed to drive disease pathology. Fusion genes may up-regulate expression of oncogenic genes by fusing a strong promoter to an oncogene. While gene fusions are most common in hematologic malignancies, with the advent of next-generation sequencing technology, rare gene fusion events, such as KIF5B-RET in lung adenocarcinoma, have now been identified in common solid tumors. The identification of rare gene fusions promises to play an important role in personalized cancer treatment decisions. Capability to detect gene fusions in formalin-fixed paraffin-embedded (FFPE) tissue would facilitate their assay in commercial laboratory settings, as it is the standard clinical practice to create FFPE tissue specimens from biopsies and surgical resections. Using a novel bioinformatics approach we have examined FFPE RNA-sequencing datasets from two breast cancer cohorts of 136 and 77 patients and detected multiple gene fusion events at base-pair resolution. These fusion transcripts may result from genomic rearrangements or transcript level rearrangements such as trans-splicing. A large portion of the identified fusion genes have been experimentally validated. These results demonstrate the power of next-generation sequencing as a tool for biomarker discovery in FFPE tissues.

1199F

Analysis of quantitative trait loci (eQTLs) in pancreatic cancer by RNA sequencing. *H. Parikh[1,2], J. Jia[1,2], W. Xiao[3], I. Collins[1,2], J. Hoskins[1,2], J. Powell[3], S. Thorgeirsson[4], J. Shi[5], G. Petersen[6], L. Amundadottir[1,2].* 1) Laboratory of Translational Genomics, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD 20877, USA; 2) Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD 20877, USA; 3) Bioinformatics and Molecular Analysis Section, Division of Computational Bioscience, Center for Information Technology, National Institutes of Health, Bethesda, MD 20892, USA; 4) Laboratory of Experimental Carcinogenesis, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA; 5) Biostatistics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD 20877, USA; 6) Department of Health Sciences Research, Mayo Clinic, Rochester, MN 55905, USA.

Pancreatic cancer is the fourth leading cause of cancer mortality in the U.S. Recent genome-wide association studies (GWAS) have identified four genomic regions associated with pancreatic cancer risk in populations of European ancestry on chromosomes 1q32.1, 5p15.33, 9q34.2 and 13q22.1. To explore the genetic mechanisms underlying natural variation in gene expression in the pancreas and their relationship to pancreatic cancer risk variants, we are conducting expression quantitative trait loci (eQTLs) analysis in normal and tumor derived pancreatic tissue samples. We performed massively parallel sequencing to catalog the transcriptome of 50 normal or tumor derived human pancreatic tissues. The median number of paired end reads (100 bp) was 124.4 M (58.1 to 203.6 M). Reads were mapped to the RefSeq database (National Center for Biotechnology Information (NCBI) build 37) using the Burrows-Wheeler Aligner (BWA) software. Reads that failed to map to RefSeq were mapped to the Ensembl database and then to the human genome assembly (NCBI build 37). Gene expression was calculated as tag counts. The EDASeq package was used to perform full-quantile normalization to adjust for differences between lanes and GC correction was performed. Blood derived DNA from the same subjects was genotyped using either the Illumina Omni1 or OmniExpress chips. As a first step in our eQTL analysis, we selected genes within 2 MB up- and downstream of the most significant SNPs in each of the four regions and analyzed SNPs in a region from 500 kb up- and downstream of the four above mentioned genomic regions associated with pancreatic cancer risk. We performed linear regression analysis between gene tag counts and SNP genotypes after correcting for age and sex. We identified 125 *cis*-eQTLs for chromosome 1q32.1, 2,757 *cis*-eQTLs for chromosome 5p15.33, 3,999 *cis*-eQTLs for chromosome 9q34.2 and 629 *cis*-eQTLs for chromosome13q22.1 with nominal *P*-value < 0.05. Together these results offer a unique picture of the functional aspects of pancreatic cancer risk loci. We plan to extend our analytical approach to identify *cis*- and *trans*-eQTLs at a genome-wide level which will provide a comprehensive resource for understanding regulatory variation in pancreatic tissues.

1200W

Identification of fusion transcripts and copy number variations in acute lymphoblastic leukemia by next generation RNA- and DNA sequencing. *M. Bonin[1], C. Schroeder[1], M. Sturm[1], U. Pflückhahn[2], M. Feldhahn[3], O. Kohlbacher[3], H.G. Rammensee[4], P. Lang[2], O. Riess[1], P. Bauer[1], M. Walter[1].* 1) Medical Genetics Department, University Tuebingen, Tuebingen, Germany; 2) Department of Hematology/Oncology, University Children's Hospital, Tübingen, Germany; 3) Applied Bioinformatics, Center for Bioinformatics and Quantitative Biology Center, University of Tübingen, Tübingen, Germany; 4) Department of Immunology, Tübingen, Germany.

Acute lymphoblastic leukemia (ALL) represents a highly heterogeneous entity on the basis of diverse cyto- and molecular genetic alterations with considerable influence on prognosis and therapeutic decisions. The different subtypes are characterized by different clinical features and specific laboratory findings. Recent clinical trials have confirmed the important impact of the underlying biology of each subtype for clinical outcome. Optimal therapeutic conditions are based on specific classification of the acute leukemia subtype at diagnosis and are guided by exact and sensitive quantification of minimal residual disease during complete hematologic remission. To gain new insights in ALL, e.g. by mutation load and modified pathways, we performed Exome- and RNA-sequencing methods using Illumina next generation sequencing technologies. Chromosomal Translocations and fusion transcripts have been a class of mutations in ALL. The development of RNA-sequencing methods provides an opportunity for deep characterization of cancer cell transcriptomes and exomes and the discovery of fusion transcripts arising from genomic rearrangements. Twelve patient samples (tumor cells and fibroblasts) were selected. All samples were characterized by a pathologist. From all samples RNA and DNA were extracted and analyzed by using 76nt paired-end reads with Illumina technologies. We identify 64 novel fusion transcripts in the 12 ALL samples. Supported by the GeneFuse algorithm, we were able to validate 51/64 (80%) fusion transcripts by qRT-PCR and Sanger-Sequencing. Most fusion transcripts were associated with Copy number transitions and were partially common in highly amplified DNA regions. Around 20% of the fusion transcripts partners have either been previously detected in oncogenic gene fusions, or reported by pathway analysis to be oncogenic. The results indicating that RNA-sequencing opens a great opportunity for discovering novel fusion transcripts in ovarian cancer and we are now able to analyze the impact of this alteration for the growth and resistance mechanisms of cancer cells.

1201T

Admixture Rate Estimation and Post-Call Correction for Complete Genomics Matched Tumor-Normal Whole Genome Sequencing data. *M. Chen[1,3], Z. Omay[2,5,6], A. Serin[2,5,6], M. Gunel[2,5,6], H. Zhao[1,2,4].* 1) Yale Center for Statistical Genomics and Proteomics,Yale University, New Haven, CT; 2) Department of Genetics,Yale University, New Haven, CT; 3) Program of Computational Biology and Bioinformatics,Yale University, New Haven, CT; 4) Departments of Epidemiology and Public Health,Yale University, New Haven, CT; 5) Department of Neurosurgery, Yale University School of Medicine, New Haven, CT; 6) Department of Neurobiology, Yale University School of Medicine, New Haven, CT.

Whole genome and exome sequencing of matched tumor-normal sample pairs is becoming routine in cancer research. The platforms of two companies, Illumina and Complete Genomics (CG), have become particularly commonplace, and >90% of the complete human genome sequences reported thus far have been sequenced using these platforms. However, most of current methods have been developed to analyze sequencing data generated on Illumina and similar instruments. Due to its novel gapped read technology and relative inaccessibility of data analysis pipeline, few studies have been designed either for CG data analysis or post-call improvement. In this study, we try to address three major questions: (1) Is there any CG platform-specific bias in somatic event calling? (2) What statistics can be used for sample quality diagnostics? (3) How to improve somatic events calling (especially CNV calls) by correcting effect of contaminated or impure samples? We have found that empirical joint distribution of tumor/normal genotype and somatic LAF(lesser Allele Frequency) are robust statistics that can be used for CG sample quality diagnostics. Admixture rate can be estimated from neutral LOH regions, which are identified based on CNV calls and LOH SNP calls. Estimated admixture rate can be further validated by somatic LAF in copy neutral regions. We're now developing methods to correct CNV calls and build additional quality control filter based on estimated admixture rate.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1202F

Transcriptome analysis of blast cells of patients with acute lymphoblastic leukemia in diagnosis and relapse. *M.A. Chiabai[1], G.R. Fernandes[1], L.H.T. Sakamoto[2], R. Pogue[1], R.W. Pereira[1].* 1) Ciências Genomicas e Biotecnologia, Universidade Católica de Brasília, Brasília, DF, Brazil; 2) Núcleo de Oncologia Pediátrica, Hospital de Apoio de Brasília, Brasília, DF, Brazil.

The transcriptome analysis using RNA-seq is a powefull tool for characterization of transcript levels and alternative splicing. RNA-seq analysis comparing patients with acute lymphoblastic leukemia (ALL) in diagnosis and at the relapse offers the potential to study the differences occurring between the two events. Acute lymphoblastic leukemia is the most common type of childhood cancer, corresponding to 25% of all cancers appearing in children. In this study we performed the RNA-seq analysis of transcriptome of two patients with diagnosis of ALL using the blats of the bone marrow collected at the diagnosis of the disease and at the relapse. The two patients has ALL-T and has the same immunophenotype of the blasts at the diagnostics and relapse. Paired-end RNA-sequencing was performed in separate four lanes. About 40 millions of reads were obtained for the samples on each lane, with 200x coverage. Sequences were mapped to the human genome and genomic analysis of genomic features was performed using several computational methods. Was identified 8055 differently expressed genes (DEGs) between the diagnosis and the relapse present in both patients; and 13057 SNPs, being 1725 SNPs validated in dbSNP and 11332 SNPs not validated. These results provide the transcriptome variation level between the diagnosis and the relapse of ALL, providing important clues to understand the molecular mechanisms of the ALL disease.

1203W

Comprehensive Germline Susceptibility Variant Discovery in Ovarian Cancer Using Exome Sequencing Data. *L. Ding, K. Johnson, K. Kanchi, M. Mclellan, C. Lu, Q. Zhang, D. Koboldt, C. Kandoth, T. Graubert, T. Ley, E. Mardis, R. Wilson.* The Genome Center, Washington University School of Medicine, St. Louis, MO.

Large-scale cancer genomics projects such as The Cancer Genome Atlas (TCGA) are producing a wealth of high throughput sequence data from a large number of cancer samples and their matched normals. These data hold great promise for discovering somatic changes that give rise to cancer and also for understanding the genetic basis of cancer susceptibility. We have analyzed germline alterations, including single nucleotide variants, insertions and deletions (indels), and copy number variations, in 429 TCGA ovarian cancer cases using both exome sequencing and SNP array data. Our analysis of 672 cancer genes identified 155 truncational variants from 133 cases in 55 genes. Comparison to the background truncational rate in human genomes from the 1000 genomes project indicated a significantly increased frequency of truncations in a few well-known ovarian susceptibility genes such as BRCA1, BRCA2, PALB2, CHEK2, PMS2, and BRIP1, as well as several novel genes such as NOTCH2, SBDS, SMO, CNTLN, FANCD2, and PCSK7. Tofurther explore additional novel germline susceptibility genes and pathways, we examined whether there was a significantly higher frequency of in silico predicted deleterious germline missense variants in TCGA ovarian cancer cases compared to non-case subjects who were part of the 1000 genomes and the NHLBI exome sequencing projects using a pooled burden test. Finally, we investigated the interaction between germline susceptibility variants and somatic mutations and found that genes bearing germline susceptibility variants were more likely to be somatically mutated and that truncationalgermline variants had a high tendency of undergoing loss of heterozygosity in the tumor. The systematic analysis of the ever-growing large-scale cancer sequencing data from various cancer types will facilitate the understanding of germline susceptibility variants in cancer.

1204T

Whole-exome sequencing of four patients with critera of familial hyperplastic polyposis. *C. Garrec[1,2], S. Küry[1], P. Lindenbaum[3], V. Vidal[4], F. Airaud[1], J. Chettrit[5], E. Cauchin[2], J.F. Mosnier[2], S. Bézieau[1,2].* 1) Service de génétique médicale, CHU Nantes, Nantes, France; 2) Université de Nantes, Faculté de Médecine, EA 4273 Biométadys, Nantes, France; 3) INSERM, UMR1087, L'institut du Thorax, Nantes, France; 4) Imaxio, Biopôle Clermont-Limagne, Saint-Beauzire, France; 5) Institut d'Histo-Pathologie, Nantes, France.

Background: Hyperplastic polyposis (HP) is a rare condition characterized by multiple or large hyperplastic polyps in the large intestine, typically with proximal localization, and associated with increased risk of colorectal cancer *via* the evolution of lesions through the serrated pathway of carcinogenesis. Some cases of familial hyperplastic polyposis have been described, but the familial transmission of this entity has not been elucidated by now. **Purpose:** We sought genes potentially responsible for familial transmission of HP by comparing exome sequences from constitutional and somatic DNAs. **Methods:** Four patients were selected, who fulfilled the diagnostic criteria of hyperplastic polyposis according to the WHO Classification of digestive tumours (2000). They showed either family history of HP and/or colorectal cancer, or early onset at disease. We assessed the carcinogenesis pathway involved by morphological characterization, immunohistochemistry and molecular biology analysis. Whole-exome sequencing was conducted on both DNA derived from blood samples and DNA extracted from FFPE (Formalin-fixed paraffin-embedded) hyperplastic polyps and sessile serrated adenomas. Genomic DNA was enriched for coding region using SureSelectXT Human All Exon v4 (Agilent technologies, Inc., Santa Clara, USA), and paired-end sequencing was realized with Genome Analyzer IIX (Illumina, Inc., San diego, USA). Reads were aligned on the hg19 reference using CASAVA software. KNIME software enabled us to apply several successive filters in order to identify functional variants. **Results:** Molecular phenotypes of hyperplastic polyps were consistent with the serrated pathway of carcinogenesis. Constitutional exome sequencing enabled us to select a preliminary list of 137 candidate genes, among which several strong candidates with relevant function to carcinogenesis. This first list will be refined using somatic exome sequencing, to identify tumor suppressor genes with a double hit event. These final results will be presented at the ASHG meeting.

1205F

RAD51 paralogs mutation screening in breast and ovarian cancer families. *L. Golmard[1], V. Caux-Moncoutier[1], G. Davy[1], E. Al Ageeli[1], B. Poirot[1], C. Tirapo[1], D. Michaux[1], C. Barbaroux[1], C. Dubois d'Enghien[1], L. Castéra[1], M-H. Stern[1,2], C. Houdayer[1,3], D. Stoppa-Lyonnet[1,2,3].* 1) Service de Génétique Oncologique, Institut Curie, Paris, France; 2) Unité INSERM U830, Institut Curie, Paris, France; 3) Université René Descartes, Paris, France.

Most currently known breast cancer susceptibility genes play a role in DNA repair. Recently, genetic studies were conducted on RAD51 paralogs, involved in homologous recombination repair. Germline mutations conferring breast and ovarian cancer susceptibility were published in RAD51C and RAD51D. Then, mutations in XRCC2 were identified in breast cancer families. No pathogenic mutation was found in RAD51B and XRCC3. We analyzed the five RAD51 paralogs (RAD51B, C, D and XRCC2, 3) in 142 breast and ovarian cancer family cases with no BRCA1/2 mutation. Our cohort was enriched in ovarian family cases since RAD51C and RAD51D mutations were detected essentially in ovarian cancer families. Mutation screening was performed on genomic DNA by EMMA (Enhanced Mismatch Mutation analysis). Detected variants were characterized by sequencing analysis. We found 1 splicing mutation and 1 missense variant in RAD51B (IVS05+3A>G and p.Arg159Cys), 2 splicing mutations in RAD51C (IVS04-2A>G and IVS08+5_IVS08+7del) and 1 missense variant in RAD51D (p.Ser-207Leu). We found no mutation in XRCC2 and XRCC3. Splicing alterations were confirmed by mRNA analysis for the two RAD51C mutations, resulting in exon skipping. No RNA was available to study the RAD51B splicing mutation IVS05+3A>G but this substitution is likely to be deleterious since it is located close to the consensus splice donor site and alters splicing in silico. Splicing reporter minigene assay are planned to confirm its effect on splicing. RAD51B missense variant p.Arg159Cys is likely deleterious: it changes a basic amino acid to an hydrophobic one, in a highly conserved functional domain and it is predicted deleterious in silico. As previously described, RAD51C mutations were detected in breast and ovarian cancer family cases. RAD51B IVS05+3A>G was found in a woman who had a breast cancer at 34 years and a 3rd degree relative with an ovarian cancer at 29 years. Interestingly, RAD51B p.Arg159Cys was found in a family with three breast cancer cases without ovarian cancer. To our knowledge, this is the first report of a deleterious RAD51B mutation in a breast and ovarian cancer family. Our study also confirms RAD51C as a breast and ovarian cancer susceptibility gene. Since RAD51 paralogs mutations frequency is low, this screening should be performed in larger different populations to better estimate their penetrance and their contribution to breast versus ovarian cancers, and establish clinical guidelines.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1206W

***MLH1*, *MSH2* and *MSH6* mutations identified by DNA sequencing and MLPA analysis: Experience from the colon Cancer Family Registry (CFR).** *S.R. Gunawardena[1], M.S. DeRycke[2], N.M. Lindor[3], M.A. Jenkins[4], J.L. Hopper[4], D.D. Buchanan[5], S. Gallinger[6], P. Newcomb[7], L. LeMarchand[8], R.W. Haile[9], S.N. Thibodeau[1], Colon Cancer Family Registry.* 1) Department of Laboratory Medicine & Pathology, Mayo Clinic, Rochester, MN; 2) Department of Health Sciences Research, Mayo Clinic, Rochester, Minnesota, USA; 3) Department of Health Science Research, Mayo Clinic, Scottsdale, Arizona, USA; 4) Center for Molecular, Environmental, Genetic, and Analytic Epidemiology, The University of Melbourne, Melbourne, Victoria, Australia; 5) Cancer and Population Studies Group, Queensland Institute of Medical Research, Herston, QLD 4006, Australia; 6) Cancer Care Ontario, Mount Sinai Hospital, Samuel Lunenfeld Research Institute, University of Toronto, Ontario, Canada; 7) Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA; 8) Epidemiology Division, University of Hawaii Cancer Center, Honolulu, Hawaii,; 9) Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, California, USA.

Colorectal cancer (CRC) is the third most common cancer in both men and women and is responsible for the second largest number of cancer deaths in the US. Most CRC cases are sporadic, but 10–30% are familial and 4–5% are hereditary. The most common form of hereditary CRC, Lynch Syndrome, is caused by a germline mutation in an MMR gene (*MLH1*, *MSH2*, *MSH6*, or *PMS2*). Baseline molecular studies for the Colon CFR, an international resource for studies on the etiology of CRC, included testing of tumor tissue for the presence of defective DNA mismatch (dMMR) repair as assessed by tumor microsatellite instability or loss of protein expression by immunohistochemistry. Studies also included testing for the presence of germline mutations in *MLH1*, *MSH2*, *MSH6* and *PMS2* among selected patient groups. Cases in Group 1 had evidence of dMMR and were tested for mutations in the gene demonstrating loss of protein expression in the tumor. Cases in Group 2 had no available tumor tissue for study and were tested directly for mutations in *MLH1*, *MSH2*, and *MSH6*. Cases in Group 3 had relatives with a mutation in any of the MMR genes and were tested for the identical mutation. Mutation analysis was performed by conventional Sanger sequencing and MLPA to detect large insertion/deletions. Polymorphic variants were excluded and the remaining variants were classified as either deleterious (nonsense, frame shift, and splice site) or as variants of uncertain significance (VUS). Of 446 individuals tested in **Group 1**, 32% (93 of 295), 66% (78 of 119) and 56% (18 of 32) were positive for a variant in *MLH1*, *MSH2*, and *MSH6* respectively. Of 208 individuals tested in **Group 2**, 10% (21 of 208) were positive for *MLH1*, 9% (20 of 208) for *MSH2* and 11% (22 of 208) for *MSH6*. Finally, of 1327 individuals tested in **Group 3**, 33% (217 of 661), 29% (143 of 489) and 37% (66 of 177) of cases were positive for *MLH1*, *MSH2* and *MSH6* respectively. Results indicate that a mutation was not identified in a substantial fraction of all three MMR genes, suggesting that these may lie in regions not detected by conventional methods. One reason for the lower than expected positive results for *MLH1* is the presence of sporadic cases resulting from promoter hypermethylation, these cases were not excluded from this study. The reduction of positive cases in Group 2 compared to Group 1 highlights the value of IHC analysis to focus on which gene to test. A description of variants identified will be presented in more detail.

1207T

Whole-Exome Sequencing of a Rare Case of Familial Childhood Acute Lymphoblastic Leukemia. *J. Healy[1], C. Richer[1], J.F. Spinella[1], V. Saillour[1], R. Vidal[1], E. Bareke[1], S. Busche[2,3], B. Ge[3], T. Pastinen[2,3], D. Sinnett[1,4].* 1) Department of Hematology-Oncology, CHU Sainte-Justine Research Center, Université de Montréal, Montreal, QC; 2) Department of Human Genetics, McGill University, Montreal, QC; 3) McGill University and Genome Quebec Innovation Centre, Montreal, QC; 4) Department of Pediatrics, Faculty of Medicine, Université de Montréal, Montreal, QC, Canada.

Familial childhood acute lymphoblastic leukemia (ALL) is extremely rare and to date no leukemia susceptibility gene has been identified to explain this uncommon occurrence. We postulated that pure (non-syndromic) familial childhood ALL could result from the accumulation of disadvantaging rare DNA variants in predisposing genes or biological pathways. To address this hypothesis, we used next-generation sequencing technologies to deep-sequence the whole-exomes of a family consisting of two male non-twinned affected siblings and their parents. Using Agilent's SureSelect target enrichment system and the SOLiD 4 platform, we captured and sequenced a total of 29.5 Gb for the entire family, with a mean capture specificity of 97% and coverage >35X. In total, we identified 29,651 germline variants in sibling A that met our filtering criteria and 32,728 in sibling B, of which 1,847 (6.2%) and 1,823 (5.6%), respectively, were novel. We investigated non-synonymous homozygous and compound heterozygous positions shared between the siblings, as well as genes/pathways with increased burden of rare non-synonymous variants, and identified a combination of rare missense variants in *SLX4* (compound heterozygote at rs137976282 and rs79842542) and *GEN1* (homozygote variant at rs16981869). Both genes play a crucial role in resolving Holliday junctions originating from replication and recombination intermediates and from DNA damage. Deep re-sequencing of 30 additional childhood ALL quartets (matched normal-tumor, mother and father) and genotyping in a validation cohort of over 300 cases and controls confirmed that the affected-sibs are the only patients that carry this combination of predicted damaging missense variations. Through functional validation studies that are currently underway, we will be able to substantiate the role of these variants in childhood leukemia onset and show that concomitant inheritance of these rare variants in *SLX4* and *GEN1*, could lead to increased DNA damage and chromosomal instability, and thus play a key role in driving leukemogenesis in these siblings.

1208F

Comprehensive BRCA1 and BRCA2 Mutational Profile in Lithuania. *R. Janavicius, V. Rudaitis, L. Griskevicius.* VUH Santariskiu Clinics, Vilnius, VNO, Lithuania.

BACKGROUND. The germline mutations in BRCA1/2 genes are the most significant and well characterized genetic risk factors for breast and/or ovarian cancer. Detection of mutations in these genes is an effective method of cancer prevention and early detection. Different ethnic and geographical regions may have different BRCA1 and BRCA2 mutation spectrum and prevalence due to founder effect. The population of Lithuania has over several centuries undergone limited mixing with surrounding populations and is mostly of indigenous Baltic origin, which is different from Slavs. The aim of our study was to asses full BRCA1/2 mutational profile in Lithuanian population. METHODS. We performed comprehensive mutation analysis of BRCA1/2 genes in 605 unrelated breast and/or ovarian cancer patients (with/without family history) and predictive unaffected patients (with family history) using high resolution melting (HRM) screening (Light Cycler 480/ Light Scanner 384) followed by direct sequencing (ABI 3500) and MLPA for large genomic rearrangements (LGRs). RESULTS. Overall, we have identified 25 different mutations (16 in BRCA1 and 9 in BRCA2 genes). Seven frequent pathogenious mutations in BRCA1 gene (c. 4035delA, c. 5266dupC, c.181T>G, c.1687C>T, c. 5258G>C, c.3700_3704del5, c.4789C>G, del 1–3ex) comprised 48%, 29%, 7,7%, 3,5%, 2,1%, 1,4% and 1,4% respectively of all BRCA1 mutations; a single BRCA2 mutation (c.658delGT) comprised 43% of all mutations in this gene. Five novel BRCA1 (c.4516delG, c.2481delA, c.5560delC, del 16–19ex, del 3–12ex) and 4 novel BRCA2 genes mutations (c.6408_6414del7, c.5701_5714del14, c.6410delA, c.6999insT) were identified; 3 different LGRs (del 1–3ex, del 16–19ex, del 3–12ex) were found in BRCA1. The most common c.4035delA (48% of all BRCA1/2 mutations) appears to be true Lithuanian (Baltic) founder mutation and haplotype data confirmed ancient origin of this mutation c.a. 62 generations ago. CONCLUSIONS. Characterization of BRCA1/2 mutational profile in Lithuania enabled to develop screening protocol using HRM for 8 common BRCA1/2 point mutations, which comprise 88% of all mutations detected in our country. This knowledge will provide more efficient approach for the individualization of genetic testing affordable for all breast/ ovarian patients and their relatives.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1209W

A broad re-sequencing study of 409 genes in NCI-60 cell lines using the Ion Ampliseq™ Comprehensive Cancer Panel and Ion PGM™ semiconductor sequencing reveals previously unreported cell line-specific mutations. *B.S.G. Kong[1], M. Shannon[1], I. Casuga[1], D. Joun[1], S.M. Chen[1], C.Y. Li[1], D. Ruff[1], R. Bennett[2]*. 1) Ion Torrent a part of Life Technologies, South San Francisco, CA; 2) Ion Torrent a part of Life Technologies, Carlsbad, CA.

Highly multiplexed PCR is quickly becoming the new template preparation paradigm for targeted re-sequencing studies because it allows investigators to interrogate hundreds of genes at once with greatly simplified workflows and reduced consumption of rare samples. At the forefront of this transformation is the Ion Ampliseq™ technology, which enables the selective amplification of 10s to 1000s of target sequences in a single multiplexed PCR and meshes seamlessly with the Ion semiconductor sequencing platform. Here we describe features and applications of the Ion Ampliseq™ Comprehensive Cancer Panel (CCP), which provides ready-access to hundreds of genes, making it ideal for broad targeted re-sequencing studies aimed at understanding inherited and somatically-acquired genetic variations in cancers. The panel employs approximately 16,000 primer pairs distributed across 4 pools to amplify the coding exons of 409 genes, including known oncogenes and tumor suppressor genes, and requires just 40ng of sample DNA. In this study, we applied the panel to the characterization of five NCI-60 cell lines (MCF-7, MDA-MB-231, DU-145, PC-3, SK-MEL-28) derived from breast, prostate, and skin cancers. The CCP includes 61 genes that have been sequenced in one or more of the cell lines in previous studies using Sanger sequencing methods. In this study, we confirmed the presence of previously reported mutations in these genes in the five cell lines. For the 348 additional genes that were evaluated with CCP, we found previously unreported mutations, including missense and non-coding mutations, in several of the genes. Taken together, these findings not only confirmed the presence of known mutations in the NCI-60 cell lines from previous surveys, but also uncovered cell line-specific mutations in genes only evaluated by CCP. The ability to examine more than 400 high-profile genes at once empowers researchers to realize a more complete picture of the genetic variation that underlies different cancer types.

1210T

Germline Sequencing for Aggressive Prostate Carcinoma. *D. Larson[1], D. Koboldt[1], E. Appelbaum[1], M. O'Laughlin[1], R. Fulton[1,2], J. Haslag-Minoff[4], I. Borecki[2,3], A. Kibel[6], R. Wilson[1,2,5], E. Mardis[1,2,5]*. 1) Washington Univ Sch Medicine, The Genome Institute, Saint Louis, MO; 2) 2Department of Genetics, Washington University, St Louis, MO, USA; 3) Division of Statistical Genomics, Washington University, St Louis; 4) Division of Urology, Washington University, St Louis, MO, USA; 5) Siteman Cancer Center, Washington University, St Louis, MO, USA; 6) 6Depart of Urology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA.

Prostate cancer is the most common cancer in men and has the second highest cancer mortality rate in the U.S. Sixteen percent of men will be diagnosed with the disease, but only a fraction are believed to have disease which can cause harm. Upwards of 50% of patients diagnosed via PSA screening are likely to have indolent tumors that do not require treatment. Unfortunately, the ability to determine who had indolent disease is imperfect and patients are over and undertreated. Thus, determining the genetic mechanisms underlying risk for aggressive prostate cancer would allow for patient treatment to be tailored to likely disease severity. Recent work has demonstrated that rare variants in coding genes contribute to susceptibility to aggressive prostate cancer. To identify rare variants that, in aggregate, may influence susceptibility to aggressive prostate cancer, we have sequenced the exomes of 300 Caucasians (150 cases and 150 controls) and 271 African-Americans (122 cases and 149 controls). Here we present preliminary results for single nucleotide variants (SNVs) and small indels. We predicted over 250,000 SNVs that were novel to dbSNP and the 1000 Genomes Project in African Americans and over 150,000 in Caucasians. We examine the distribution of rare deleterious variants within the Cancer Gene Census list of known cancer genes. Testing on the number of rare variants in cases versus controls indicates a rare variant contribution to susceptibility in the TET2 tumor suppressor gene in African Americans, but not Caucasians. In addition, we see evidence that rare variants in SETDB1 and LIFR may play a protective role.

1211F

A next-generation sequencing diagnostic panel to test all cancer susceptibility genes. *S.S. Mahamdallie, E. Ruark, K.-W. Lau, A. Renwick, S. Seal, E. Ramsay, S. Hanks, J. Douglas, N. Rahman.* Division of Genetics & Epidemiology, The Institute of Cancer Research, Sutton, United Kingdom.

Extensive efforts are being directed towards using exome sequencing in clinical diagnostics, but there are considerable technical, analytical, logistical and financial hurdles that are currently not easily circumvented. By contrast, panels of select genes can offer cost and practical advantages and can provide data of diagnostic standard. A panel of all cancer susceptibility genes is ideally suited to this strategy as many genes predispose to multiple cancers and many cancers are associated with multiple genes. We have designed a custom pulldown panel that targets all 85 cancer predisposition genes, including 2649 probes targeting 1375 exons. We undertook a pilot study in 48 samples, 46 with mutations in one or more genes, and two controls (1 duplicate and 1 negative). There were 53 mutations in total, 40 base subs or small indels and 13 large exonic duplications/deletions. We generated data from a HiSeq; 101bp paired-end (PE) reads, 24 samples per lane, a MiSeq; 101 and 151bp PE reads, 12 samples per lane and we compared these data to 48 exomes (HiSeq, 6 samples per lane). Sequence reads were mapped, called and annotated using Stampy, Platypus and an in-house algorithm, respectively. The HiSeq data were excellent. Only 3 probes failed. Mean coverage was 906x and at 50x an average of 99% exonic bases were covered and 97% of 1375 exons passed (defined as at least 90% covered in 90% of samples). The MiSeq data were much better for 151bp than 101bp, but were inferior to the HiSeq. For 151bp reads the mean coverage was 116x and at 50x an average of 82% exonic bases were covered but only 47% of exons passed. In the exome data 25 exons were not targeted at all and at 50x only 24% of exons passed. Our software detected all base subs and small indels in the HiSeq data, but missed one in the MiSeq data. The exonic duplications/deletions were visible in the HiSeq data, though automated calling with acceptable sensitivity and specificity is challenging and currently under optimisation. With regard to practicality and cost, up to 384 samples on a HiSeq could be run in 25 days at a consumable cost of £100 per sample. For the MiSeq, two lanes of sequencing should give sufficient coverage, and could provide faster turnaround time for fewer samples, albeit at higher cost. Our data show that a diagnostic assay for all cancer susceptibility genes is realistic in the near future and will greatly increase the number of people that can benefit from cancer gene testing.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1212W

New germline *MET* variants in Hereditary Papillary type 1 Renal Carcinomas within the French population. *E. Rouleau[1], C. Lefol[1], S. Caputo[1], V. Verkarre[2,3], C. Guy[1], F. Copigny[1], C. Maugard[13], O. Caron[8], F. Eisinger[4], P. Berthet[6], Y-J. Bignon[11,3], J. Chiesa[5], A. David[15], T. Frebourg[7], S. Giraud[9,3], S. Lejeune[12], J-M. Limacher[10], H. Zattara[14,3], S. Deveaux[2,3], I. Bieche[1], S. Richard[2], R. Lidereau[1].* 1) Laboratoire d'oncogénétique somatique et constitutionnelle, Hôpital Huguenin, Institut Curie, 92210 St Cloud, France; 2) Centre Expert National Cancers Rares PREDIR, INCa-AP/HP, Service d'Urologie, Hôpital de Bicêtre, 94275 Le Kremlin-Bicêtre-France; 3) Faculté de Médecine, Université Paris Descartes, Service D'Anatomie et de Cytologie Pathologiques, Hôpital Necker-Enfants Malades, 75015 Paris-France; 4) Institut Paoli-Calmettes, 13273 Marseille; 5) Groupe Hospitalo-Universitaire Caremeau, 30029 NIMES; 6) Oncogénétique, Centre François Baclesse 14076 Caen France; 7) Département de génétique, CHU de Rouen, Inserm U1079, IRIB Faculté de Médecine et de Pharmacie, 76031 Rouen France; 8) Institut Gustave Roussy, Consultation de Génétique, 94805 Villejuif France; 9) Hospices Civils de Lyon, Service de Génétique Moléculaire et Clinique, Lyon 69431 France; 10) Consultation de Génétique Oncologique, Hôpital Pasteur, COLMAR, France; 11) Laboratoire Diagnostic Génétique et Moléculaire, 63011 Clermont-France; 12) Service de Génétique Clinique, Hôpital Jeanne de Flandre, CHRU Lille, 59038 Lille-France; 13) Hôpitaux Universitaires de Strasbourg, Consultation d'Oncogénétique et laboratoire de diagnostic génétique, 67091 Strasbourg France; 14) Hôpital de la Timone, 13385 Marseille France; 15) CHU Nantes,Service de Génétique Médicale,44093 Nantes, France.

Background. Germline mutations in the proto-oncogene *MET* are responsible for hereditary papillary type 1 renal carcinoma. We describe the French National Cancer Institute renal cancer network experience for the detection and characterization of *MET* mutations. Methods: Since 2001, 108 index-cases were screened from all over France with a familial history of papillary renal cell carcinoma (RCC), multiple papillary RCC or early-onset RCC with papillary histology (under 50 year-old). Lymphocyte DNA was analyzed either dHPLC or qPCR-HRM and Sanger sequencing from exon 2 to exon 21. In 2 mutated cases, somatic molecular profiles from RCC were documented using CGH-array. Results. We found 6/108 index cases with the following causal and activating *MET* variants (p.His1112Arg x2, p.Val1236Ile, Val123-8Ile x2, p.Tyr1248Cys). Three cases belong to kindred with family history of RCC and 3 cases did not have any family history. The age of onset was from 36 (in one case) to over 50 year-old in 5 cases. All documented mutated patients had papillary type 1 RCC (P1RCC). The CGH-array profile obtained in two tumors from carriers identified the same pattern of autosomal trisomy 7 and trisomy 17. Fifteen index cases were identified with 11 rare unclassified variants : 4 intronics variants, 3 exonic variants at the proximity of the tyrosine kinase domain (p.Thr1010Ile x2, p.His1086Leu, p.His1097Arg), 1 within the tyrosine kinase domain (p.Ile1102Thr) and 4 in the SEMA domain (p.Glu168-Asp x2, p.Ile316Met x2, p.Ala347Thr, p.Asn375Ser). Three of them (p.His10-86Leu, p.His1097Arg) close to and within the tyrosine kinase (p.Ile1102Thr) were modelized as potentially causal and involved in papillary type 1 RCC. Discussion. Half of the germline *MET* mutations were identified in patients without any family history of renal cancer. The mean age of onset of P1RCC was not younger than in general population. Tumor characterization could help to better identified carriers with germline *MET* mutation, particularly in sporadic papillary type 1 RCC. The role of some potential causal variants should be further analyzed. Conclusion: Germline *MET* mutation screening should be proposed in the case of papillary renal carcinoma type 1 even with an age of onset superior to 50 years-old, with or without familial history.

1213T

Rare mutations in XRCC2 confer increased risk of breast cancer. *M. Southey[1], D.J. Park[1], F. Lesueur[2], T. Nguyen-Dumont[1], M. Pertesi[2], F.A. Odefrey[1], F. Hammet[1], S.L. Neuhausen[3], E.M. John[4,5], I.L. Andrulis[6], M.B. Terry[7], M. Daly[8], S. Buys[9], F. Le Calvez-Klem[2], A. Lonie[10], B.J. Pope[10], H. Tsimiklis[1], C. Voegele[2], F.M. Hilbers[11], N. Hoogerbrugge[12], A. Barroso[13], A. Osorio[13,14], G.G. Giles[15], P. Devilee[11,16], J. Benitez[13,14], J.L. Hopper[17], S.V. Tavtigian[18], D.E. Goldgar[19], kConFab, BCFR.* 1) Dept Pathology, Univ Melbourne, Melbourne, Victoria, Australia; 2) Genetic Cancer Susceptibility Group, International Agency for Research on Cancer, 69372 Lyon, France; 3) Department of Population Sciences, Beckman Research Institute of City of Hope, Duarte, CA 91010, USA; 4) Cancer Prevention Institute of California, Fremont, CA 94538, USA; 5) Department of Health Research and Policy and Stanford Cancer Center Institute, Stanford, CA 94305, USA; 6) Department of Molecular Genetics, Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, ON, Canada, M5G 1×5; 7) Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY 10032, USA; 8) Fox Chase Cancer Center, Philadelphia, PA 19111, USA; 9) Huntsman Cancer Institute, University of Utah Health Sciences Center, Salt Lake City, UT 84112, USA; 10) Victorian Life Sciences Computation Initiative (VLSCI), Carlton, Victoria 3010, Australia; 11) Department of Human Genetics, Leiden University Medical Center, Leiden, 2300 RC Leiden, The Netherlands; 12) Department of Human Genetics, Radboud University Nijmegen Medical Center, 6525 GA, Nijmegen, The Netherlands; 13) Human Genetics Group, Human Cancer Genetics Programme, Spanish National Cancer Center (CNIO), E-28029 Madrid, Spain; 14) Spanish Network on Rare Diseases (CIBERER), Valencia 46010, Spain; 15) Centre for Cancer Epidemiology, The Cancer Council Victoria, Carlton, Victoria, 3052, Australia; 16) Department of Pathology, Leiden University Medical Center, Leiden, 2300 RC Leiden, The Netherlands; 17) Centre for Molecular, Environmental, Genetic and Analytical Epidemiology, School of Population Health, The University of Melbourne, Victoria 3010, Australia; 18) Department of Oncological Sciences, Huntsman Cancer Institute, University of Utah School of Medicine, Salt Lake City, UT 84112, USA; 19) Department of Dermatology, University of Utah School of Medicine, Salt Lake City, UT 84132, USA.

Currently, only approximately 30% of the familial risk for breast cancer has been explained, leaving the substantial majority unaccounted for. Recently, exome sequencing has been demonstrated to be a powerful tool for identifying the underlying cause of rare Mendelian disorders. However, diseases such as breast cancer present substantially increased complexity in terms of locus, allelic and phenotypic heterogeneity, and relationships between genotype and phenotype. As part of a collaborative multiple-case breast cancer family exome capture followed by massively parallel sequencing project involving the Leiden University Medical Centre, the Spanish National Cancer Research Center and The University of Melbourne we identified a family with a protein-truncating and a family with a likely deleterious missense mutation in XRCC2. Based on the finding from exome sequencing, the rarity of these variants and the biochemical plausibility of XRCC2, we conducted two further studies in parallel: (i) Case-control mutation screening of XRCC2 and (ii) mutation screening of XRCC2 in a series of index cases from multiple-case breast cancer families. The population-based case-control mutation screening study identified six likely pathogenic coding variants in 1,308 early-onset breast cancer cases versus none in 1,120 controls (with severity grading, $P<0.02$). The additional mutation screening in 689 multiple-case families identified further families with protein-truncating or likely deleterious rare missense variants in XRCC2. Our identification of XRCC2 as a breast cancer susceptibility gene thus increases the proportion of breast cancers that are associated with homologous recombination DNA repair dysfunction and Fanconi Anemia and therefore could benefit from specific targeted treatments such as PARP inhibitors. This study demonstrates the power of massively parallel sequencing to discover susceptibility genes for common, complex diseases.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1214F

Identification of Genetic Susceptibility Loci for Familial Prostate Cancer through Whole Exome Sequencing: The ICPCG Study Design. *S.N. Thibodeau[1], M. DeRycke[1], S. McDonnell[1], S. Gunawardena[1], Y.W. Asmann[1], S. Middha[1], L.A. Cannon-Albright[3], J.L. Stanford[4], E.A. Ostrander[5], W.B. Isaacs[6], J. Xu[7], J. Schleutker[8,25], K.A. Cooney[9], E.M. Lange[10], J.D. Carpten[11], J.E. Bailey-Wilson[2], O. Cussenot[12], G.G. Giles[13,14], G. Severi[13,14], C. Maier[15], A.S. Whittemore[16], C.L. Hsieh[17], F. Wiklund[18], W.J. Catalona[19], W.D. Foulkes[20], D. Mandal[21], R. Eeles[22], D. Easton[23], D. Seminara[24], D. Schaid[1] on behalf of The International Consortium for Prostate Cancer Genetics*. 1) Mayo Clinic, Rochester, MN; 2) National Human Genome Research Institute, National Institutes of Health, Baltimore, MD; 3) University of Utah School of Medicine, Salt Lake City, UT; 4) Fred Hutchinson Cancer Research Center, Seattle, WA; 5) National Human Genome Research Institute, National Institutes of Health, Bethesda, MD; 6) Johns Hopkins Medical Institutions, Baltimore, MD; 7) Wake Forest University School of Medicine, Winston-Salem, NC; 8) University of Turku, Turku, Finland; 9) University of Michigan Medical School, Ann Arbor, MI; 10) University of North Carolina, Chapel Hill, NC; 11) The Translational Genomics Research Institute, Phoenix, AZ; 12) Hopital Tenon, Paris, France; 13) The Cancer Council of Victoria, Melbourne, Victoria, Australia; 14) University of Melbourne, Melbourne, Victoria, Australia; 15) University of Ulm, Germany; 16) Stanford University School of Medicine, Stanford, CA; 17) University of Southern California, Los Angeles, CA; 18) Karolinska Institutet, Stockholm, Sweden; 19) Northwestern University Feinberg School of Medicine, Chicago, IL; 20) McGill University, Montreal, Quebec, Canada; 21) Louisiana State University Health Sciences Center, Department of Genetics, New Orleans, LA; 22) Institute of Cancer Research and Royal Marsden NHS Foundation Trust, Surrey UK; 23) University of Cambridge, Cambridge, UK; 24) National Cancer Institute, National Institutes of Health, Bethesda, MD; 25) University of Tampere and Fimlab Laboratories, Tampere, Finland.

Prostate cancer (PC) is the most common male cancer in western societies. Although the variation of PC incidence results from several factors, a large body of literature strongly implicates a genetic etiology. Disease heterogeneity, however, has hampered efforts to identify PC susceptibility genes. Next-generation sequencing provides a powerful tool to uncover disease-causing mutations and, when combined with a family-based study design, increases the chances of finding rare mutations. We hypothesize that much of the “missing heritability” of PC is due to rare variants and that these can be identified using whole-exome sequencing (WES). The International Consortium for Prostate Cancer Genetics (ICPCG) is in the unique position of having identified and sampled the most informative high-risk PC pedigrees known throughout the world. We outline a three-stage study design developed by the ICPCG to identify PC susceptibility loci utilizing this extraordinary collection of families coupled with WES. In stage 1, WES is performed on ~575 familial cases of PC derived from 375 independent families: 3 cases/family from 100 families and 1 case/family from 275 families. Two complementary approaches are utilized for the initial data analysis: a family-based and a case-control analysis. The pedigree-based analysis avoids the potential complication from population stratification, while the familial case-control analysis avoids the complication of phenocopies in pedigree data. Those genes with variants that are most damaging (e.g., nonsense, frame-shift, splice site variants and selected other variants), those that co-segregate with PC within the tested families, and those that are found to be significantly more frequent among our familial cases compared to control data are then prioritized for further study. In stage 2, the top 1000 candidate genes identified in stage 1 are further analyzed by re-sequencing the coding regions in an independent set of 500 hereditary PC cases and 500 controls, ranking genes based on those containing multiple damaging variants, co-segregation with disease and allele frequency among cases compared to controls. In stage 3, the most likely PC susceptibility loci are identified by exon and intron re-sequencing of the top 100 ranked genes identified in stage 2 in an independent set of 1000 hereditary PC cases and 1000 controls. Preliminary results from the WES of the ~575 men will be presented.

1215W

Discovery of novel long noncoding RNAs and aberrant alternative splicing events by RNA-sequencing in childhood acute lymphoblastic leukemia. *R. Vidal[1], J.F. Spinella[1], V. Saillour[1], C. Richer[1], J. Healy[2], E. Bareke[1], S. Busche[2,3], B. Ge[3], T. Pastinen[2,3], A. Droit[4], D. Sinnett[1,5]*. 1) Department of Hematology-Oncology, CHU Sainte-Justine Research Center, Université de Montréal, Montreal, QC; 2) Department of Human Genetics, McGill University, Montreal, QC; 3) McGill University and Genome Quebec Innovation Centre, Montreal, QC; 4) Department of Molecular Medicine, Centre de Recherche du CHUQ, Université Laval, Québec, QC; 5) Department of Pediatrics, Faculty of Medicine, Université de Montréal, Montreal, QC, Canada.

Acute lymphoblastic leukemia (ALL) is the most frequent pediatric cancer, accounting for approximately 25% of cases. Advances in our understanding of the pathobiology of ALL have led to risk-targeted therapy and increased survival rates. However, approximately 20% of patients do not respond to current treatment protocols and over two-thirds of survivors experience long-term treatment-related health problems. Despite progress in treating the disease, very little is known of its underlying etiology. Therefore, in an attempt to investigate the complex landscape of mutations linked to changes in expression profiles and alternative splicing events, the whole transcriptome profiles of 30 primary ALL tumor samples and a subset of 8 matched normal samples were generated using a RNA-Sequencing approach on the ABI SOLiD platform. An average of 45M high quality paired-end tags were obtained per sample after aligning to the human genome reference sequence. Using a multi-step analysis pipeline including the Lifescope Genomic Analysis Software (Life Technologies), edgeR and Cufflinks, we investigated differential transcript expression between tumor and normal samples. We found a total of 2112 differentially expressed transcripts and isoforms, including 435 putative new transcripts related to the disease. Using a protein-protein network approach we identified hub genes related to DNA damage control (CCNA2, PARP2), apoptosis (ADAM9, FLT3) and involved in disease-relevant signaling pathways (CX3CR1, EPOR) that were shown to be up-regulated in over 60% of the cohort. We also found novel recurrent tumor-specific alternative splicing events involving genes that overlap these disease-associated pathways, as well as long ncRNAs whose expression in tumor samples correlate with well known ALL oncogenes such as MME, CDK6 and FLT3. These results suggest that long ncRNAs could be new potential therapeutic targets. Interestingly, we also found heterogeneity across ALL tumor transcriptomes suggesting that common but also primary tumor-specific transcriptional events are involved in childhood ALL onset and disease outcome. Integration of these results with exome sequencing, methylation profiling, and high-density genotyping data from these patients will allow us to identify bona fide driver events in leukemogenesis and provide further insight into the inherent complexity of this childhood disorder.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1216T

Association of polymorphisms in the *FOXE1* gene, but not *NKX2-1*, with familial papillary thyroid carcinoma. *S.G. Wilson[1,2,3], C-Y. Yan[1], L. Ward[1], S. Chew[1], V. Panicker[1], S.J. Brown[1], S. Chiripal[1], J. Goldblatt[2], T.D. Spector[3], J. Walsh[1,2].* 1) Endocrinology and Diabetes, Sir Charles Gairdner Hosp, Perth, Western Australia; 2) School of Medicine and Pharmacology, University of Western Australia, Crawley, Western Australia; 3) Department of Twin Research and Genetic Epidemiology, King's College London, London, United Kingdom.

Thyroid cancer is among the ten most common cancers in populations of European ancestry. The vast majority of primary thyroid cancers are papillary thyroid carcinoma (PTC) and a small proportion (~5%) of these show strong familial association (FPTC). The genetic factors that determine susceptibility to PTC remain uncertain, however genes encoding thyroid transcription factors are strong candidates. Among these the FOXE1 and NKX2-1 genes have previously been implicated for a role in PTC (1,2). We performed a sequencing-based association study of these two genes and FPTC. Genetic analyses used conventional Sanger sequencing of the genomic DNA of 15 individuals aged 34.7 +/− 10.9 (mean ±SD) years at diagnosis, from 11 pedigrees. PTC families were identified on the basis of a proband with PTC containing at least one additional first degree relative also with PTC. The BigDye Terminator v3.1 Cycle Sequencing method with subsequent capillary electrophoresis performed on a ABI3730 sequencer was used for the analyses of nucleotide sequence of the genes in DNA from the study subjects. A control group of 118 healthy females selected randomly from the TwinsUK cohort with no previously reported thyroid disease, served as a control group. SNP genotype data for the control group was available from analyses of genomic DNA using Hap610 SNP arrays. Statistical analyses used a Chi-square test and UNPHASED Software was used for haplotype analysis. We found a significant association between FPTC and the SNP rs907577 (p = 0.002) located 5' of the FOXE1 gene. There was no association of rs71369530, a common polyalanine deletion in exon 1, with the disease. Haplotype analysis of common variants in FOXE1 provided further evidence for association with thyroid cancer (p = 0.0008). In comparison, we found no association of genetic variants identified through sequencing in NKX2-1 with PTC. Amongst our study subjects a single missense mutation (NM_004473.3:c.743C>G; Ala248Gly), predicted to be of a benign nature, was detected in the FOXE1 gene of one subject. No rare variants were detected in the data for the NKX2-1 gene. These data confirm previous findings of a role for common genetic variation in the FOXE1 gene region in the risk for PTC and extends the finding to FPTC.

1) Gudmundsson J, et al. (2009) Nat Genet 4: 460–464. 2) Landa I, et al. (2009) PLoS Genet. 5(9):e1000637.

1217F

XPC gene founder splicing site mutation is common in Brazilian Xeroderma Pigmentosum patients. *M.W. Achatz[1,6], K.M. Santiago[1], F.P. Vairo[2], P. Ashton-Prolla[2], P.F.V. de Medeiros[3], R.M. Rocha[4], S.R. Rogatto[1,5].* 1) Centro Internacional de Pesquisa, Hospital A.C. Camargo, São Paulo, Brazil; 2) Universidade Federal do Rio Grande do Sul, Hospital de Clínicas de Porto Alegre, Rio Grande do Sul, Brazil; 3) Hospital Universitário Alcides Carneiro, Universidade Federal de Campina Grande, Paraíba, Brazil; 4) Departamento de Anatomia Patológica, Hospital A.C. Camargo, São Paulo, Brazil; 5) Departamento de Urologia/Faculdade de Medicina, Universidade Estadual Paulista, São Paulo, Brazil; 6) Departamento de Oncogenética, Hospital A.C. Camargo, São Paulo, Brazil.

Xeroderma Pigmentosum (XP) syndrome is a rare autosomal recessive disease clinically characterized by cutaneous hypersensibility to ultraviolet radiation that accounts for the induction of pigmentary alterations and to early onset of skin cancers. XP is genetically heterogeneous and germline mutations are found in one of seven autosomal genes encoding proteins of the Nucleotide Excision Repair (NER) pathway or mutations in a polymerase eta encoding gene, active in the translesion synthesis (TLS) pathway. In this study 18 Brazilian families apparently unrelated who received clinical diagnosis of XP were included. Clinical characterization and molecular profiling through direct sequencing of XPA and XPC gene for the screening of germline mutations was performed. Germline mutations on XPC gene were found in 55,5% (10/18) of the families and one family (5,5%) was detected as carrier of XPA p.Arg207Ter mutation. Five new genetic variants occurring in the coding region and splice site of XPC gene were identified. Five apparently unrelated families were found to be carriers of a common splice site mutation in homozygozity, c.2251-1G>C, which correlates with the absence of immunoexpression. The same mutation has recently been observed as a founder effect in the Mahori XP population, generating alternative aberrant transcripts and lack of protein expression on fibroblasts cell extracts. This study shows that the Brazilian XP patients share a high incidence of the same founder c.2251-1G>C as the Mahori population. Ancestry molecular studies are needed to define if there is a common origin involving these two populations.

1218W

Rare variants in genes from esophageal squamous cell carcinoma genome-wide association studies identified by exome/whole genome sequencing of high-risk upper gastrointestinal cancer families. *A.M. Goldstein[1], N. Hu[1], K.J. Jacobs[1,2], L.-J. He[3], X.-Y. Han[4], M. Rotunno[1], M. Cullen[1,2], J. Boland[1,2], H. Su[1], L. Wang[1], C. Wang[1], L. Burdett[1,2], M. Malasky[1,2], A. Hutchinson[1,2], M. Yeager[1,2], T. Ding[4], C. Giffen[5], M.A. Tucker[1], S.J. Chanock[1], M. Lee[6], P.R. Taylor[1].* 1) Division of Cancer Epidemiology and Genetics, National Cancer Inst, NIH, Bethesda, MD; 2) Core Genotyping Facility, SAIC-Frederick, Inc., NCI-Frederick, Frederick, MD; 3) YangCheng Cancer Hospital, YangCheng, Shanxi, PR China; 4) Shanxi Cancer Hospital, Taiyuan, Shanxi, PR China; 5) Information Management Services, Inc., Silver Spring, MD; 6) Center for Cancer Research, NCI, NIH, Bethesda, MD.

Gastric cancer and esophageal cancer cause more than 700,000 and 400,000 deaths, respectively, each year, and represent the second and sixth leading causes of cancer-related death worldwide. These upper gastrointestinal (UGI) cancers occur in the Taihang Mountains of North-Central China at some of the highest rates reported for any cancer; more than 20% of all deaths in this mountainous area have been attributed to these cancers. Risk factors for UGI cancer in this region include family history and dietary deficiencies and suggest a complex interplay of genetic and environmental factors. Multiple genome-wide association studies (GWAS) of esophageal squamous cell carcinoma (ESCC) in cases and controls from the Taihang Mountains as well as other Chinese and Asian populations have identified numerous loci encompassing approximately 14 genes associated with ESCC. The reported genes include PLCE1, C20orf54, ALDH2, ACAD10, ADH1B, RPL6, PTPN11, C12orf51, PDE4D, RUNX1, UNC5CL, CASP8, TRAK2, and ALS2CR12. Several of these loci have been observed in more than one study population whereas others have been reported only once. To examine whether rare variants (defined as minor allele frequency < 0.05) in any of these 14 genes were observed in high-risk families with multiple members with UGI cancer from the Taihang Mountains, we evaluated exome sequencing data from multiplex UGI families from this region. Families were selected for study if they had at least two members with UGI cancer and available blood-derived DNA. Exome sequencing was conducted in 2–4 affecteds/obligate carriers from 10 multiplex families. Nine families had at least one member with ESCC. Analysis of exome data revealed 6 rare substitutions and one possible intronic splice variant in 5 of the genes (PLCE1, c12orf51, RUNX1, CASP8, and TRAK2). Five of the rare variants had not been reported in dbSNP, 1000 Genomes, or NHLBI's Exome Variant Server and were considered novel. None of the variants, however, showed complete co-segregation with all UGI cases within the families in which they were observed. In fact, only two substitutions, one in C12orf51 and one in PLCE1, were observed in multiple affected family members. Future studies are needed to examine the function of these variants and to determine whether they increase risk for ESCC/UGI cancer. Additional studies to further evaluate the frequencies of these variants in large numbers of Chinese UGI cases and controls are also essential.

1219T

Identification of breast cancer susceptibility genes. *L. Guidugli[1], J.N. Weitzel[2], X. Wang[1], S. Hart[1], F.J. Couch[1], C. Szabo[3].* 1) Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN; 2) Clinical Cancer Genetics, City of Hope Cancer Center, Duarte, CA, USA; 3) Department of Biological Sciences, University of Delaware, Newark, DE, USA.

Breast cancer has a significant genetic component. The strongest epidemiologic risk factor is family history of the disease. Familial studies have identified a number of genes contributing to increased risk. Of these, the highly penetrant BRCA1 and BRCA2 genes account for the majority of families in whom disease-causal mutations have been identified. However, mutations in other genes have been associated with predisposition to breast or breast and ovarian cancer including CHEK2, ATM, and PALB2. Together these genes account for less than 20% of excess familial breast cancer risk and do not explain ~50% of high-risk families referred for genetic testing. This significant gap in our understanding of the genetic basis of breast cancer poses a significant barrier to our ability to identify women at increased risk, who would benefit greatly from appropriate clinical management. To identify additional breast cancer predisposition genes whole exome sequencing of familial high-risk breast cancer patients was undertaken. DNA samples from pairs of breast cancer patients from 70 high-risk breast cancer families with no pathogenic mutations in the BRCA1 or BRCA2 breast cancer predisposition genes were subjected to whole exome sequencing. After exclusion of common variants, candidate mutations in genes encoding proteins involved in the cellular response to DNA repair were validated. Segregation of validated candidates with breast cancer was evaluated in families carrying the mutations. Results of these ongoing studies will be presented.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1220F

Mutations in *SDHA* are a common cause of paragangliomas and pheochromocytomas and give rise to a diverse tumor spectrum. *A.R. Mensenkamp[1], J.U. Rao[2, 3], K.L.I. van Gassen[1], B. Kusters[4], J.W.M. Lenders[5, 7], H.P.M. Kunst[6], H.J.L.M. Timmers[3], M.J.L. Ligtenberg[1, 4].* 1) Human Genetics, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands; 2) Department of Endocrinology, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands; 3) Department of Laboratory Medicine, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands; 4) Department of Pathology, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands; 5) Department of Internal Medicine, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands; 6) Department of Otorhinolaryngology, Head and Neck Surgery, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands; 7) Department of Internal Medicine III, University Hospital Carl Gustav Carus Dresden, Germany.

Head and neck paragangliomas (HNPGL) and adrenal/extra-adrenal pheochomocytomas (PHEO) are highly heritable neural crest-derived tumors. Mutations in at least 10 genes are described to cause either HNPGL and/or PHEO. *SDHA*, encoding one of the subunits of succinate dehydrogenase, has only recently been recognized as a HNPGL/PHEO susceptibility gene. Although mutation analysis of *SDHA* is complicated by the presence of several pseudogenes, we have been able to design a primer set which exclusively amplifies all *SDHA* exons. Seventy patients who were negative for mutations in *SDHB*, *SDHC*, *SDHD*, *SDHAF2*, *RET*, *VHL*, *TMEM127* and *MAX* with no signs of NF1 were examined for mutations in *SDHA*. Tumors from patients with a mutation, if available, were subjected to immunohistochemistry of SDHA and SDHB, and analyzed for loss of heterozygosity. Analysis of SDH-activity is ongoing. In addition, metabolites were screened by ^{1}H-NMR and family history was recorded. In six patients an *SDHA* mutation was detected. The previously described nonsense mutation c.91C>T (p.(Arg31*)) was found in two patients with a HNPGL and in one patient with a right adrenal PHEO and a renal cell carcinoma. The pathogenic mutation c.1534C>T (p.(Arg512*)) was detected in a patient with a testis paraganglioma at the age of 23, who developed metabolically active retroperintoneal metastases three years later. His aunt, who also is heterozygous for the mutation, developed a tympanic HNPGL. The putative pathogenic missense mutation c.1753C>T (p.(Arg585Trp)) was found in a woman who developed a Wilms tumor at four years of age and a PHEO at age 20 years. Another putative pathogenic missense mutation, c.778G>A (p.(Gly260Arg)), was detected in a woman with a jugulotympanic HNPGL. All immunohistochemical and loss of heterozygosity analyses were in line with a loss of function of the latter mutations. Most patients with an *SDHA* mutation did not have relatives with HNPGL/PHEO, suggesting a reduced penetrance of *SDHA* mutations. In conclusion, *SDHA* mutations may account for about 8% of all patients with a HNPGL/PHEO without a mutation in any other known gene. The tumor spectrum appears to be highly variable. HNPGL, PHEO, and other tumors may develop at a young age and can be metastatic. Mutation analysis of *SDHA* should therefore be part of the genetic screening in all patients with a HNPGL/PHEO who are negative for *SDHB* and *SDHD* mutations.

1221W

Identifying novel cancer susceptibility genes through exome sequencing and copy number analysis of individuals with Li Fraumeni-like cancer phenotypes. *S.E. Plon[1,2,3], L.C. Strong[4], B. Powell[2], L. Jiang[3], H. Cheung[1], D. Ritter[3], D.M. Muzny[3], D.A. Wheeler[3], R.A. Gibbs[2,3].* 1) Dept. Pediatrics, Baylor College of Medicine, Houston, TX; 2) Dept. Molecular & Human Genetics, Baylor College of Medicine, Houston, TX; 3) Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX; 4) Department of Genetics, MD Anderson Cancer Center, Houston, TX.

Background: The identification of mutations in specific genes responsible for inherited cancer susceptibility impacts patient treatment, as well as allowing increased surveillance and prevention for at-risk family members. In addition, knowledge of inherited predisposition to pediatric malignancies has provided important insights into the mechanisms of cancer, in both children and adults. Current utilization of high-throughput sequencing methodologies for detecting coding changes in cancer has been focused on somatic mutations in tumor tissue, and conversely; genome-wide association studies of cancer susceptibility have focused on common variations that individually have small impacts on cancer risk. Methods: We are carrying out a project at the intersection of these two different approaches, using massively parallel sequence analyses of constitutional DNA from ethnically diverse childhood cancer patients with medical and family histories suggestive of inherited cancer susceptibility syndromes. Patients are enrolled through cancer genetics protocols at several institutions in Texas. We are analyzing small cohorts of families with consistent and specific phenotypes including probands with childhood sarcomas and second or third primary malignancies by age 40, highly suggestive of the Li Fraumeni syndrome but where thorough sequence and copy number analysis of TP53 and other well-characterized cancer-associated genes has been negative. Results: We have identified eight unrelated individuals who meet these criteria, the majority of which have either negative family history or a single affected first-degree relative consistent with autosomal dominant inheritance. Capture of coding regions using a vCrome 2.1 platform was followed by Illumina paired end exome sequencing performed on constitutional DNA from blood or lymphoblastoid line and analyzed for novel or rare mutations which are predicted to impact gene function. Rare copy number variants were determined based on Affymetrix high-density arrays. Given the limited availability of family samples, genes will be prioritized for further analysis by the identification of deleterious variants or CNVs in multiple independent probands. Conclusions: Genome-scale sequencing and copy number analysis should allow identification of causative mutations in individuals with p53-like cancer phenotypes in the absence of extended family information. Supported by grants 5R01CA138836 and CPRIT RP101089.

1222T

Development and validation of NGS-based molecular diagnosis of inherited forms of colorectal cancer. *J. Tinat[1,2], F. Charbonnier[1,2], S. Coutant[1], R. Marlin[1], G. Bougeard-Denoyelle[1], S. Baert-Desurmont[1,2], M. Tosi[1], T. Frebourg[1,2], I. Tournier[1].* 1) Inserm, U1079, University of Rouen, Institute for Research and Innovation in Biomedicine (IRIB), Rouen, Normandy, France; 2) Department of Genetics, University Hospital, Rouen, Normandy, France.

Considering the number of genes involved in Mendelian predisposition to colorectal cancer and the overlapping phenotypes which require to sequentially analyse different genes, we have developed a one step NGS-based strategy. All the exons, introns and 15 kb of the 5' and 3' untranslated regions of 10 genes respectively involved in Lynch syndrome (*MSH2*, *MLH1*, *MSH6* and exons 6 to 15 of *PMS2*), adenomatous polyposis (*APC* and *MUTYH*) and hamartomatous polyposis (*STK11*, *SMAD4*, *BMPR1A* and *PTEN*) were captured using the Sure-Select technology from Agilent. We also included in this capture the *TP53* gene for logistic reasons, since our laboratory also performs this analysis at the national level. *TP53* results were subsequently masked for patients referred to our laboratory in the context of inherited forms of colorectal cancer. We multiplexed the samples prior to the capture (5 patients/capture) and after the capture (10 patients/flow cell lane) and performed a paired-end 2×75 bp sequencing on an Illumina GAIIx platform. This strategy allows us to analyse 80 patients in one run. The informatics pipeline includes the RTA and CASAVA 1.8 software from Illumina, and the ANNOVAR and Variant Effect Predictor software. The average sequencing depth was around 200 x, across the captured regions. In order to validate this strategy, we analysed 70 patients harbouring a total of 113 known variations including 70 point mutations, 31 insertions/deletions and 12 genomic rearrangements. All the point mutations were successfully identified. Among the 31 insertions/deletions, only 2 located within *STK11* and *MSH6* were missed because of a low read depth caused by a high GC content. In order to estimate the false positive rate, we selected 42 variations revealed by NGS and checked their presence by Sanger sequencing. All variations were confirmed, except two resulting from sequencing artefacts located in mononucleotide repeats. This evaluation shows that our capture set and sequencing strategy yield an excellent sensitivity and specificity and can be used in molecular diagnosis laboratories in order to optimise the molecular diagnosis of Mendelian forms of colorectal cancer.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1223F

Interpretation of Variants of Unknown Significance with a Large Database of Genotyped and Phenotyped Individuals. *B.T. Naughton, A. Chowdry, J.M. Macpherson, G.M. Benton.* 23andMe, Inc., Mountain View, CA.

The interpretation of variants of unknown significance (VUS) from whole-genome sequence data is a substantial challenge in genetics. VUS are usually too rare to be amenable to genome-wide association studies and so traditionally have been interpreted with reference to the primary literature (especially for high-penetrance or Mendelian mutations) or by computational methods (e.g., SIFT, PolyPhen). While these methods can provide useful insights, they are often limited by the presence of false positives in the literature or by imperfect prediction algorithms. 23andMe, a personal genomics company, has assembled a database of 150,000 genotyped individuals, over 90,000 of whom have consented to participate in research and answered at least one research question. Participants answer research questions on the 23andMe website on topics as diverse as their medical history, personality, lifestyle and exercise. Here we present data demonstrating that this database can be used to empirically determine the significance of variants found in human sequence data. As proof of concept we used the database to confirm that the BRCA1 mutations 185delAG and 5382insC and the BRCA2 mutation 6174delT are associated with greatly increased breast cancer risk. Conversely, we confirmed that the BRCA mutations R841W and S1040N are benign polymorphisms that are not associated with increased breast cancer risk. In a real-world example, we analyzed a VUS in MLH1 from a sequenced exome that was suspected to be cancer-causing. Using individuals' self-reported cancer status from the database, we determined that the variant is unlikely to be cancer-causing. Our finding agrees with previously reported results in the literature. Due to the extensive phenotyping of our cohort (over 50 million phenotypic data points) and the large number of rare variants on our custom genotyping chip (over 30,000 putatively disease-associated rare variants), this method is applicable to a large number of genes and phenotypes. We further discuss extending this method to variants not present on the genotyping chip by inferring the presence of mutations in individuals in the database based on identity-by-descent (IBD).

1224W

Exome sequencing identifies rare deleterious mutations in DNA repair genes FANCC and BLM as potential breast cancer susceptibility alleles. *E.R. Thompson[1], M.A. Doyle[1], S.M. Rowley[1], D.Y.H. Choong[1], G.K. Philip[2], A.H. Trainer[1], P.A. James[1,3], G. Mitchell[1,3], I.G. Campbell[1,3,4], kConFab.* 1) Peter MacCallum Cancer Centre, East Melbourne, Victoria, Australia; 2) Life Sciences Computation Centre, Victorian Life Sciences Computation Initiative, Carlton, Victoria, Australia; 3) Sir Peter MacCallum Department of Oncology, University of Melbourne, Parkville, Victoria, Australia; 4) Department of Pathology, University of Melbourne, Parkville, Victoria, Australia.

Despite intensive efforts using linkage and candidate gene approaches, the genetic etiology for the majority of families with a multigenerational breast cancer predisposition remains unknown. We completed whole exome sequence analysis of germline DNA from multiple affected relatives (n= 130) from 64 breast cancer families with the aim of identifying rare protein truncating and non-synonymous variants that are likely to include novel cancer predisposing mutations. Twelve individuals from nine families were found to harbour mutations in known breast cancer genes (BRCA2–7 individuals, CHEK2–2, BRCA1–1, PTEN - 1, ATM - 1) that were not identified by clinical genetic testing either because the index case was a phenocopy who did not carry the mutation, or because the gene is not routinely tested. Among the remaining families, a list of ~50 candidate genes has been identified on the basis of recurrence and biological function, and these are currently being assessed for segregation among family members and for prevalence among 1,000 BRCA1/BRCA2 mutation negative breast cancer families and controls. In particular, our analysis identified families with heterozygous, deleterious mutations in the DNA repair genes FANCC and BLM, which are responsible for the autosomal recessive disorders, Fanconi Anemia and Bloom's syndrome. In total, screening of all exons in these genes in 453 breast cancer families identified three with truncating mutations in FANCC, and two with truncating mutations in BLM. Additional screening of FANCC mutation hotspot exons identified one pathogenic mutation among an additional 957 breast cancer families. None of the deleterious mutations were identified among 464 healthy controls and are not reported in the 1000 genomes data. Given the rarity of Fanconi Anemia and Bloom's syndrome disorders among Caucasian populations, the finding of multiple deleterious mutations in these critical DNA repair genes among high risk breast cancer families is intriguing and suggestive of a predisposing role. Our data demonstrates the utility of intra-family exome sequencing approaches to uncover cancer predisposition genes but highlights the major challenge of definitively validating candidates where the incidence of sporadic disease is high, germline mutations are not fully penetrant, and individual predisposition genes may only account for a tiny proportion of families.

1225T

Cancer genetic testing panels: preliminary experience within an adult genetics practice. *C.G. Selkirk, S.M. Weiss, K.J. Vogel, A.C. Newlin, S.M. Weissman, P.J. Hulick.* Center for Medical Genetics, NorthShore University HealthSystem, Evanston, IL.

Background: The traditional clinical approach to genetic testing for suspected cancer syndromes has followed a single-gene testing approach, based on the constellation of cancers seen in a family history. With the recent clinical availability of cancer genetic testing panels, which includes next-generation sequencing for 14–22 genes, it is yet to be determined whether the paradigm for cancer genetic testing will shift. In some cases, ordering a panel may be more cost effective than performing testing in a step-wise approach of syndromes with overlapping phenotypes. Additionally, panels may identify mutations where the family does not have a classic presentation for the associated cancer syndrome. However, given the number of genes on each panel, it is anticipated to involve a high rate of variants of unknown significance (VUS). At present, the clinical utility is unknown for such panels and there are no set criteria or guidelines established for offering such testing. Methods: We have devised a clinical ranking system for determining appropriateness for cancer gene panel testing, balancing the potential yield of testing with the likelihood of using the information to guide medical management. Beginning April 2012, patients are ranked using a clinical risk ranking system based on their personal and/or family history. We are tracking our patient cohort who has had a discussion of cancer panel testing and tracking outcomes related to insurance preauthorization, testing uptake, and testing results. Results: We will report on our cohort of patients who has been introduced to panel testing. Based on the assigned clinical risk number, we will describe patients' decisions to pursue insurance pre-authorization, outcomes of pre-authorization, patient uptake of testing, test result outcomes, associated medical management recommendations, and the utility of our ranking system to guide recommendations for cancer panel genetic testing. Conclusions: Next-generation sequencing technology may be a cost-effective and time-saving approach for clarifying cancer risk in families with inherited cancer predispositions. In order to validate this approach to genetic testing, data is needed regarding patient interest and uptake in testing, insurance coverage for testing, test result outcomes (including variant rates), and utility of testing to guide medical management recommendations.

1226F

Thirty percent of breast cancer families negative by commercial BRCA1/BRCA2 testing are resolved by mutations in 13 other breast cancer genes, by BRCA1/BRCA2 phenocopies, and by BRCA1/BRCA2 CNVs. *T. Walsh, S. Casadei, A.M. Thornton, G. Bernier, C.H. Spurrell, S.M. Stray, J. Mandell, M.K. Lee, M-C. King.* University of Washington, Seattle, WA.

A challenge to the present practice of genetic testing for inherited risk of breast cancer is how to explain the illness in breast cancer patients with severe family history, but negative (wildtype) results by commercial testing for BRCA1 and BRCA2. We are studying this question for all breast cancer patients in >2000 severely affected families. Proband(s) of each family received negative results of commercial genetic testing of BRCA1 and BRCA2. We are using our recently developed approach of targeted capture of complete genomic loci and massively parallel multiplexed sequencing to identify all single base pair substitutions, insertion-deletions, and copy number variants in all known breast cancer genes. Here we report on sequence data from >1000 patients from the first 457 families. More than 90% of families are of European ancestry, including 9% Ashkenazi Jewish. We included here only germline truncating mutations, complete gene deletions, and missenses experimentally shown to be damaging. In 30% of families (135/457), mutations very likely to be causal were identified: 18% (81 families) by mutations in 13 known breast cancer genes other than BRCA1 and BRCA2; 2% (8 families) by conventional mutations in BRCA1 or BRCA2, with proband(s) proving to be phenocopies; and 10% (46 families) by CNVs in BRCA1 or BRCA2, generally in families without supplementary commercial CNV analysis. Genes other than BRCA1 and BRCA2 harboring germline loss-of-function mutations were CHEK2 (29 families), PALB2 (21), ATM (10), TP53 (5), PTEN (4), RAD51C (3), BRIP1 (2), NBN (2), BARD1 (1), CDH1 (1), MLH1 (1), RAD50 (1), and STK11 (1). All mutations were individually rare, with known founder alleles (e.g. CHEK2 c.1100delC) accounting for only a small proportion. In genes other than BRCA1 and BRCA2, 9% of mutations were gene-disrupting CNVs, in ATM, CHEK2, PALB2, RAD51C, and TP53. Subsequent sequencing of still-unresolved families has thus far revealed eight new candidate breast cancer genes (see abstract by Spurrell et al.). Our results indicate that families severely affected by breast cancer but with negative results from commercial genetic testing are well served by complete genomic sequencing of all known breast cancer genes. As such comprehensive testing becomes more widespread, it will be important to determine the risks associated with each of these genes so they can be incorporated effectively into clinical care.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1227W

Comparison of Next Generation Sequencing with Traditional Sequencing and MLPA for BRCA1 and BRCA2 Testing. *K. Chun[1,2], A. Brown[3], K. Ng[3], R. Denroche[3], J.D. McPherson[3,4].* 1) Genetics Program, North York General Hospital, Toronto, Ontario, Canada; 2) Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Ontario, Canada; 3) Ontario Institute for Cancer Research, Toronto, Ontario, Canada; 4) Department of Medical Biophysics, University of Toronto, Toronto, Ontario, Canada.

Objective: To compare next generation sequencing (NGS) with traditional sequencing and MLPA (multiplex ligation-dependent probe amplification) for BRCA1 and BRCA2 clinical testing in Ontario. Methods: 96 patients with different known pathogenic variants in BRCA1 or BRCA2 were identified. Full Sanger sequencing and MLPA data of all the exons of BRCA1 and BRCA2 was available for 87 of the patients, while the remaining 9 patients had information pertaining to the pathogenic variant alone. The HaloPlex PCR Target Enrichment and Library preparation protocol was used to select the regions of interest (BRCA1 and BRCA2) and construct barcoded sequencing libraries. Equimolar pools of the barcoded libraries were sequenced on the Illumina MiSeq platform using a paired end 150 bp read length protocol. Demultiplexed FASTQ files were generated from on instrument basecalls using Illumina's CASAVA software. The sequences were aligned to the human genome using the Novoalign short read aligner. Single nucleotide variants and short indels were identified using the Genome Analysis Toolkit and annotated with ANNOVAR. Results: 87 pathogenic variants and 18 variants of unknown significance were detected in the 96 patients. Sixty of these variants were frameshift mutations and 45 were single base pair changes. Furthermore, all of the polymorphisms (single base pair changes) present in these patients were also correctly identified. NGS was not, however, able to detect the 9 exonic deletions and one exonic duplication, although evidence of these large genomic rearrangements was observed upon closer examination by comparing average read depth across each exon. NGS was found to be more cost-effective and less labour intensive than traditional Sanger sequencing and MLPA. Conclusions: Next generation sequencing is an efficient and cost-effective method to potentially replace traditional sequencing and MLPA for BRCA1 and BRCA2 clinical testing in Ontario.

1228T

Genomic capture and massively parallel sequencing reveals inherited loss-of-function mutations in 8 genes in 19% of familial breast cancer patients from Greece. *A. Stavropoulou[1], T. Walsh[2], F. Fostira[1], M. Tsitlaidou[1], S. Casadei[2], S. Glentis[1], G. Fountzilas[3], I. Konstantopoulou[1], M.C. King[2], D. Yannoukakos[1].* 1) NCSR DEMOKRITOS, ATHENS, Greece; 2) Dept of Medicine (Medical Genetics), University of Washington, Seattle, WA, USA; 3) Aristotle University of Thessaloniki, Thessaloniki, Greece.

Inherited mutations in known and yet-to-be discovered genes predispose for breast and/or ovarian cancer. Mutations within BRCA1 and BRCA2 are the most common, but recently additional genes have been linked to breast and/or ovarian cancer susceptibility. We used BROCA to target, hybridize, and multiplex sequence 30 known breast and ovarian cancer genes. Coding exons, UTRs, introns, and several kb flanking each locus were captured and sequenced to median read depth >300x, so that >99% of basepairs have >10x coverage. Subjects were 88 Greek patients with familial, invasive breast cancer, for whom Greek founder mutations in BRCA1 and BRCA2 had been previously excluded. Eighteen nonsense, frameshift, multi-exon deletions, or missense mutations previously established as damaging were identified in 17 patients (19%). Genes harboring loss-of-function mutations were BRCA1, BRCA2, ATM, CHEK2, PALB2, PTEN, MLH1, and MSH2. One patient carried mutations in two genes (MLH1 and ATM). All mutations were different. All were confirmed by Sanger sequencing with diagnostic primers from patients' genomic DNA. Five of the 18 mutations were genomic deletions of sizes 1.8 to 12 kb in BRCA1, BRCA2, CHEK2, and MSH2 that were detected by read depth from BROCA data. These CNVs deleted from one to three exons. Four other patients carried mutations at or near splice sites in ATM, BRCA2, and RAD51C. These splice variants have been confirmed in genomic DNA; their effects on transcripts are being evaluated. The patient with both MLH1 and ATM truncating mutations had been diagnosed with both breast and colorectal cancer and had a family history of both cancers. The patient with a PTEN nonsense mutation was diagnosed with breast cancer at age 36; her mother died of breast cancer at age 29 and a maternal cousin died of childhood leukemia. We conclude that among Greek patients with familial breast cancer, the mutational spectrum is highly heterogeneous with respect to both loci and alleles. These patients are well served by an approach that detects all classes of mutations in all known breast cancer genes.

1229F

Elevated rate of somatic L1 retrotransposition in colorectal tumors. *S. Solyom[1], A.D. Ewing[2], E. Rahrmann[3], T. Doucet[1], H.H. Nelson[4], D.F. Sigmon[1], A. Casella[1], B. Erlanger[5], S. Wheelan[5], G. Faulkner[6], D. Haussler[2], D. Largaespada[3], H.H. Kazazian[1].* 1) McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD; 2) Center for Biomolecular Science and Engineering, University of California at Santa Cruz, Santa Cruz, California; 3) Department of Genetics, Cell Biology and Development and Pediatrics, Masonic Cancer Center, University of Minnesota, Minneapolis, Minnesota; 4) Division of Epidemiology and Community Health, Masonic Cancer Center, University of Minnesota, Minneapolis, Minnesota; 5) Department of Statistics and Oncology, Johns Hopkins University School of Medicine, Baltimore, MD; 6) Genome Plasticity and Disease Group, Mater Medical Research Institute, Aubigny Place, Raymond Tce, South Brisbane, Australia.

Long INterspersed Element-1 (LINE-1 or L1) retrotransposons comprise 17% of the human genome, where they are the only autonomous mobile elements. L1s move by a “copy and paste” mechanism via an RNA intermediate, a process called retrotransposition. They are also responsible for the mobilization of the non-autonomous Alu and SVA retrotransposons, as well as numerous other RNA species. Although, insertional mutagenesis by these mobile elements is a known cause of various human Mendelian disease cases, their mutagenic load on the cancer genome has so far been elusive. Using a next generation L1-targeted resequencing method on the DNA of 16 colorectal tumors and matched normal colon tissue, here we show for the first time that certain cancers are excessively mutagenized by human-specific L1 (L1Hs) elements, the only currently active family of human L1s. We provide the most extensive confirmation of de novo L1 insertions in malignancy to date by analyzing 112 insertion sites with PCR. 72 of these were successfully PCR amplified at the 3' junctions and validated to be tumor-specific, and 65 were sequenced. For 34 insertions altogether, both 5' and 3' junctions were retrieved, enabling us for the first time to characterize target site duplications (TSDs, a hallmark of retrotransposition) and L1 insertion size in cancer tissue. Interestingly, all of the validated tumor-specific insertions were severely 5' truncated, and we yet have to find a full length L1 in these tissues. This is in contrast to germline polymorphic insertions, where 30% were full length, in agreement with the length of L1 insertions in the reference genome. This difference in truncation rate is statistically significant, and potentially indicates the footprint of a DNA repair pathway in somatic tissues or cancer. The number of validated L1 insertions varies widely from tumor to tumor with up to 17 insertions in some and none in others. Among the insertions, numerous genes playing a role in tumorigenesis have been targeted. Surprisingly, the elevated rate of L1 retrotransposition does not correlate with global L1 methylation or with microsatellite instability. Thus, somatic L1 retrotransposition occurs frequently in colorectal cancers, and the truncated structure of these insertions may be useful in determining the etiology of 5' truncation in L1 insertion events.

1230W

Mutational analysis of splicing machinery genes SF3B1, U2AF1 and SRSF2 in myelodysplasia and other common tumors. *S. Lee, E. Je, M. Kim, J. Oh, N. Yoo.* Pathology, College of Medicine, The Catholic University of Korea, Seoul, South Korea.

Recently, a whole-exome sequencing identified frequent somatic mutations of genes encoding multiple components of the RNA splicing machinery, including splicing factor 3B subunit 1 (SF3B1), serine/arginine-rich splicing factor 2 (SRSF2) and U2 small nuclear RNA auxiliary factor 1 (U2AF1) in bone marrows of myelodysplastic syndromes (MDS) patients (45–85%). RNA splicing is a mechanism for gene regulation and proteome diversity, and is ubiquitous to all types of human cells, suggesting a possibility that mutational events in the splicing genes might be involved in other tumors. For this, we have analyzed SF3B1, SRSF2 and U2AF1 genes in a variety of human tumors in hematologic (400 cases) and solid tumors (900 cases). In 61 MDS, 22 MDS (36%) harbored at least one somatic mutation of these genes. 80% of the MDS with ring sideroblasts harbored the mutations, while 32% of those without remarkable ring sideroblasts harbored them. 7% of the acute myelogenous leukemias, 25% of the chronic myelomonocytic leukemias, 50% of MDS/myeloproliferative disorders harbored the mutations, frequencies of which were similar to preexisting data. We newly identified that childhood acute lymphoblatic leukemias (ALL) also harbored SRSF2 mutations in two cases (1%). We also observed two mutations in solid tumors (prostate and gastric cancers), but the mutations were not the hotspot mutations. Our data indicate that the splicing gene mutations are tissue-specific and suggest that the mutations play unique roles in the pathogenesis of MDS and less frequently other hematologic neoplasia, but not in common solid tumors.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1231T

Calling Allele or Haplotype-specific Copy Number in Tumor Sequence Data. *N. Dewal[1], Y. Hu[2], M.L. Freedman[3,4], T. LaFramboise[5], I. Pe'er[2], R.A. Gibbs[1], D.A. Wheeler[1]*. 1) Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX; 2) Department of Computer Science, Columbia University, New York, NY; 3) Dana-Farber Cancer Institute, Boston, MA; 4) Harvard Medical School, Boston, MA; 5) Department of Genetics, Case Western Reserve University, OH.

During tumor initiation and progression, cancer cells acquire a selective advantage, allowing them to out-compete their normal counterparts. Identification of the genetic changes that underlie these tumor acquired traits can provide deeper insights into the biology of tumorigenesis. Regions of copy number alterations (CNAs) and germline DNA variants are some of the elements subject to selection during tumor evolution. Integrated examination of inherited variation and somatic alterations holds the potential to reveal specific nucleotide alleles that a tumor "prefers" to have amplified or deleted. Next-generation sequencing of tumor and matched normal tissues provides a high-resolution platform to identify and analyze such somatic CNA regions. Within a CNA region, examination of informative (e.g. heterozygous) sites deviating from a 1:1 ratio may suggest selection of that allele. A naïve approach examines the reads for each heterozygous site in isolation; however, this ignores available valuable linkage information across sites. We therefore present an updated version of our Hidden Markov Model-based method - Haplotype Amplification in Tumor Sequences (HATS) - 2.0 that analyzes tumor and normal sequence data, along with training data for phasing purposes, to infer amplified or deleted alleles and haplotypes in regions of somatic copy number change. Our method is designed to handle rare variants and biases in read data. We assess the performance of HATS using simulated amplified or deleted regions generated from varying copy number and coverage levels, followed by CNA regions in real data. We demonstrate that HATS infers the amplified or deleted alleles more accurately than does the naïve approach, especially at low to intermediate coverage levels and in cases (including high coverage) possessing stromal contamination or allelic bias. Lastly, HATS 2.0 can now detect bi-allelic alterations, at the same time providing a better estimate of tumor purity, compared to the assumption of mono-allelic changes made by the original version of HATS. HATS will thus enable identification of the allele selected for copy number change.

1232F

Somatic Mutation Detection of Tumor and Matched-Normal Samples for Semi-Conductor Based Sequencing. *A.H. Joyner, S. Utiramerur, D. Brinza, H. Breu, B. Krishnaswami, A. Shukla, J. Zhai, Y. Lou, F. Hyland, D. Thomas, E. Beasley*. Bioinformatics, Life Technologies, Foster City, CA., USA.

The advent of next-generation DNA sequencing technology has enabled quick and low-cost variant analysis of tumor samples. One of the major challenges to the field of somatic mutation detection is in distinguishing novel tumor sample variants (somatic) from parentally-inherited variants (Germ-line). Ion Reporter™ Software is a hosted suite of informatics tools that streamline and simplify analysis, annotation, and archiving of semiconductor sequencing data. It is designed specifically for clinical researchers performing routine sequencing assays by automating variant analysis bioinformatics pipeline to connect detected variants with public annotations, allowing faster biological assessment. It is optimized for semiconductor sequencing for faster and more accurate results, and scales easily from a few to thousands of samples and from genes to genomes. Here is presented a method, integrated into the Ion Reporter™ Software suite, to confidently determine and annotate tumor-specific variants. This method analyzes normal sample reads and takes into consideration chance observations, informed from the error rate. In this way, evidence for the tumor allele in the normal sample is used to qualify a variant as a tumor-specific variant. User provided parameters allow this workflow to be modified for various experimental settings (i.e. asymmetrical coverage of tumor/normal, expected frequency of somatic variant), and provides functionality to make non-confident calls in the face of imperfect sequencing evidence.

1233W

Acute lymphoblastic leukemia - novel therapeutic approach based on an active patient-individualized multipeptide vaccination. *C. Schroeder[1], M. Sturm[1], U. Pflückhahn[2], M. Feldhahn[3], O. Kohlbacher[3], S. Stevanovic[4], H.G. Rammensee[4], P. Lang[2], O. Rieß[1], P. Bauer[1]*. 1) Department of Medical Genetics, Tübingen, Germany; 2) Department of Hematology/Oncology, University Children's Hospital, Tübingen, Germany; 3) Applied Bioinformatics, Center for Bioinformatics and Quantitative Biology Center, University of Tübingen, Tübingen, Germany; 4) Department of Immunology, Tübingen, Germany.

Acute lymphoblastic leukemia (ALL) is the most common cancer in childhood with an incidence of 3.5 / 100.000 at the age of 0–14 years. Treatment of ALL improved tremendously over the last decades, leading to a 5-year event free survival of more than 80 %. However, patients with high-risk acute lymphoblastic leukemia often relapse and in this case therapeutic options are limited. Thus, novel therapeutic concepts are needed. We established a workflow for an individualized multi-peptide vaccination in children suffering from high-risk ALL of B-Cell lineage. Tumor specific mutations are identified via the combination of whole-exome (tumor cells and fibroblasts) and transcriptome (tumor cells) sequencing. The variants found by differential analysis are validated on DNA-level using a custom-design deep-sequencing approach. All validated variants are further characterized by in-silico predictions (e.g. SYFPEITHI) and analysis of tumor tissue (HPLC plus mass spectrometry) to identify those mutations resulting in mutated peptides that are presented on the cell surface. The variants most likely to be immunogenic can be synthesized under GMP-conform conditions and administered to the patient. So far six patients have been included in this study and were sequenced according to the protocol (avg. percentage of reads on target: 70.1 %; avg. percentage of target regions at 10x: 90.41%). With stringent filter criteria an average of 83 tumor-specific mutations per sample pair were identified. The laboratory turnaround time for sequencing and validation is estimated to be around 8 weeks. The next steps ahead will be prediction and analysis of mutated peptides within the patient's tumor tissue. In summary, with this approach we should be able to detect tumor-specific and immunogenic mutations that are the basis for a successful individualized immunotherapy.

1234T

No evidence for the role of somatic mutations and promoter hypermethylation of FH gene in the tumorigenesis of nonsyndromic uterine leiomyomas. *S. Vaidya[1,2], K.P. Rao[3], Q. Hasan[2]*. 1) Vasavi Medical and Research, Hyderabad, India; 2) Kamineni Hospital, Hyderabad,Hyderabad; 3) Osmania University,Hyderabad,India.

Fumarate hydratase (FH) gene is reported to have specific involvement in syndromic uterine tumors, but its role in nonsyndromic forms is still unclear. Hence, the present study has aimed to screen the role of promoter methylation status and mutations in exon 2 and 7 regions of FH gene in the genesis of nonsyndromic uterine leiomyomas. Leiomyoma and myometrium tissues were collected from 85 hysterectomized uterine specimens. DNA from each of the biopsy was subjected to PCR, methylation-specific restriction assay, and DNA sequencing. In silico analysis was carried out to identify the impact of sequence variants on the protein structure. Chisquare ($\chi 2$) test was used to compare the promoter methylation proportions of leiomyoma and myometrium tissues. No sequence variants were observed in exon 2 region, but three novel heterozygous germ line sequence variants, i.e.,c.1010A>C, c.1021 G>A, and c.1066 T>C in exon 7 region of the FH gene were detected in 14/85 (16.5 %) of the cases examined. In silico analysis results showed that c.1010A>C and c.1021 G>A mutations damage the structure and function of FH, whereas c.1066 T>C mutation is mostly tolerant or neutral. No significant difference of FH promoter methylation status between the leiomyoma (11.76 %) and myometrium (5.88 %) tissues was observed (P00.176). Therefore, it is concluded that somatic mutations in FH do not show pronounced effect in nonsyndromic uterine leiomyomas compared to that of their syndromic counterparts. However, higher frequency of FH mutations in leiomyoma cases raises the need to conduct larger number of prospective case-control and family-based studies to assess them as risk markers to nonsyndromic leiomyomas. Keywords: Nonsyndromic uterine leiomyoma. Fumarate hydratase. Missense variants. SIFT scores. Promoter methylation.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1235F

Somatic mutation spectrum of metastatic melanoma through exome sequencing of 48 tumor-normal pairs. *M.D. Willard[1], J.N. Calley[1], S-S. Wong[1], R.H. Higgs[3], X. Ma[3], P.J. Ebert[1], S.M. Bray[1], I.H. Wulur[1], Y. Yue[1], Y. Lin[2], J. Wang[2], A. Aggarwal[1], S. Li[2], C.R. Reinhard[1], A.B. West[2], T.D. Barber[1]*. 1) Translational Therapeutics, Eli Lilly and Company, Indianapolis, IN; 2) Informatics, Eli Lilly and Company, Indianapolis, IN; 3) Statistics, Eli Lilly and Company, Indianapolis, IN.

Melanoma is a deadly form of skin cancer that results from the unregulated growth of melanocytes, and is an area of high unmet medical need. Although promising targeted therapies have recently emerged, it is clear that the high mutation rate and complex mutation spectrum of melanoma has contributed to the inconsistent improvements in overall survival rate. To further our understanding of the somatic mutation spectrum, we sequenced the exomes of 48 melanoma tumor-normal pairs. We identified mutations in well-established melanoma genes (*i.e.*, BRAF, NRAS, KIT), in genes recently described in melanoma (*i.e.*, PTK2B, GRM3, PREX2), and in genes whose role in melanoma progression remains to be determined. Additionally, utilizing our melanoma mutation data, as well as complementary datasets from recently published exome and whole-genome sequencing studies, we identified novel recurrent mutations that may play a role in driving melanoma tumorigenesis. Our comprehensive investigation expands the understanding of the mutation spectrum of melanoma and identifies mutation events that may underlie the etiology of this deadly disease.

1236W

RNA-Seq identifies differentially expressed genes and mutations in oligodendrogliomas. *E. Schrock[1], K. Szafranski[2], J.A. Campos Valenzuela[3], S. Schauer[1], D. Krex[4], A. Rump[1], K. Hackmann[1], G. Schackert[4], L. Kaderali[3], M. Platzer[2], B. Klink[1]*. 1) Institute for Clinical Genetics, Medical Faculty, University of Technology, Dresden, Germany; 2) Leibniz Institute for Age Research - Fritz Lipmann Institute (FLI), Jena, Germany; 3) Institut für Medizinische Informatik und Biometrie, Medizinische Fakultät Carl Gustav Carus, Technische Universität Dresden, Germany; 4) Klinik und Poliklinik für Neurochirurgie, Universitätsklinikum Carl Gustav Carus, Technische Universität Dresden, Germany.

Brain tumor patients have a very poor prognosis (survival time often below one year) and targeted therapies are not yet available. Therefore, we started an explorative study for the identification of therapeutic targets in 13 oligodendroglial tumors. Most oligodendrogliomas (O) and also some oligoastrocytomas (OA) are characterized by a unique and typical unbalanced translocation, der(1;19), resulting in a 1p/19q-codeletion. We studied nine Os/OAs with and four Os/OAs without the 1p/19q-codeletion in addition to three normal brain samples obtained commercially. We combined transcriptome next generation sequencing (RNA-Seq) using a 2×100 nt paired-end approach on the Illumina-HiSeq-platform, with high resolution array CGH using the 400K chip as well as miRNA array, 6×80k expression array and Exon array analyses (Agilent). Array CGH identified additional aberrations (e. g. -4, -9, -14, -18), which are known to be associated with an unfavorable prognosis. We could also detect additional smaller deletions and duplications, which have not been described so far, such as a gain on chromosome 2 (+2p22.1-pter) involving the MYCN oncogene. The comparison of the transcriptome data between tumors with and without the 1p/19q-codeletion as well as normal brain samples revealed significantly differentially expressed genes (adjusted p-value <0.05). The average expression level of genes on 1p and 19q was reduced by half in the 1p/19q-codeleted tumors compared to tumors without the codeletion. Consistent with this, most of the down-regulated genes were located on 1p. Interestingly, in normal brain samples, the average expression level of 1p-located genes was found to be placed in between the two tumor groups. Even more surprisingly, the average expression level of the genes on 19q in normal brain samples was as low as in the tumors showing the 19q deletion. These findings were confirmed by the RNA-expression data of the Agilent 8×60k Arrays. Using RNA-Seq we could also identify structural changes, like the known IDH1 c.395G>A mutation and mutations in CIC and FUBP1. Our ongoing integrative data analysis shows that our comprehensive approach is very promising for identifying the key molecular changes in oligodendrogliomas and we aim at identifying new candidate genes as molecular therapeutic targets in a larger cohort.

1237T

Identification and quantification of somatic mutations with the Ion AmpliSeq™ Cancer Panel. *M. Andersen[1], S. Roman[1], K. Rhodes[1], C. VanLoy[1], A. Broomer[1], D. Topacio[1], G. Roma[1], R. Bennett[1], T. Neff[2], C. Beadling[2], C. Corless[2]*. 1) Life Technologies, Carlsbad, CA; 2) OHSU Kinght Cancer Institute, Portland, OR.

Translational and clinical cancer research requires rapid approaches that are compatible with small input amounts of often challenging samples, such as archived FFPE tissue. Ion AmpliSeq™ technology is a new tool for enriching tens to thousands of genomic targets in a single tube from 10ng of input DNA. The purpose of this study was to validate the performance of the Ion AmpliSeq™ Cancer panel in identifying and quantifying somatic mutations in archived FFPE tumor tissue. The Ion AmpliSeq™ Cancer Panel consists of a pool of 190 primer pairs that are designed to amplify mutation hotspot regions (including 739 known somatic mutations) across 46 genes. Forty-five archived FFPE-derived samples of tumor DNA that were previously genotyped on a Sequenom MassARRAY® platform were amplified using this panel, and the products were then ligated with barcodes, subjected to emulsion PCR and sequenced on an Ion Torrent PGM using 316 chips (4 samples per chip). The average amplicon read length was 76 bp and the average read depth was 2000; more than 95% of the amplicons had a read depth >400. All 53 known point mutations were correctly identified by the Ion Torrent variant caller software, including mutations in two DNA samples prepared from tumor cells dissected by laser-capture microscopy (Arcturus XT, Life Technologies). The range of somatic mutation frequencies was 8% to 96% and correlated well with results from the MassARRAY® (R2= 0.84). There were 19 in/dels represented in the tumor samples, of which all were visible in the final sequence but only a subset were identified by the variant caller software. In addition to the known sequence variants, 27 new non-synonymous mutations were detected in gene regions not covered by the MassARRAY® panel. These included 24 single nucleotide substitutions and 3 deletions of 2 bp. All 27 were confirmed by Sanger sequence analysis, including one variant detected in both a tumor sample and a laser microdissected specimen prepared from the same tumor. The Ion AmpliSeq™ Cancer Panel represents a fast and easy approach to generating target sequence for next-generation sequencing. It requires very little input DNA, and high read depths can be generated to support good sensitivity. Custom panels based on this technology are now being tested.

1238F

Whole-genome and targeted sequencing analysis of early stage high-grade serous ovarian cancer. *M. Bibikova[1], S. Humphray[1], R. Grocock[1], J. Peden[1], F. Nielsen[1], Z. Kingsbury[1], J. Chien[2,3], E.L. Goode[2], V. Ho[1], C. April[1], S. Munchel[1], J. Cottrell[1], Y. Tarabishy[2], J.M. Cunningham[2], S. Kaufmann[2], L.C. Hartmann[2], K.R. Kalli[2], V. Shridhar[2], J.-B. Fan[1], G. Heath[1], D. Bentley[1]*. 1) Illumina, Inc., San Diego, CA; 2) Mayo Clinic, Rochester, MN; 3) University of Kansas Cancer Center, Kansas City, KS.

Ovarian cancer is an aggressive disease which is usually diagnosed at a late stage and has poor therapeutic outcome. There are currently no genetic tests available to screen for or diagnose ovarian cancer in its early stages. To identify potential biomarkers for the early detection of ovarian cancer, we designed an enrichment pool targeting exons from 260 candidate genes thought to be somatically altered in early stage ovarian cancer (~1.6 Mb). Design was based on whole genome sequencing of 25 Mayo Clinic fresh frozen tumors (17 early stage high grade serous carcinomas and 8 other histologic subtypes) with matching blood DNA and supplemented with genes identified from the literature and public databases for further biological validation. The candidate gene enrichment pool was screened in a large set of validation samples (N=230) that were sequenced to an average 400x coverage. Due to the lack of availability of high quality fresh-frozen ovarian cancer samples, the validation set was largely represented by archived formalin-fixed paraffin-embedded (FFPE) tissues. Therefore, we optimized our TruSeq® sequencing library preparation protocol for use with FFPE samples. Good concordance in variant calls was obtained between five matched FFPE and fresh-frozen samples, even though experimental noise was higher in FFPE samples. Discordant variant calls were mainly due to low depth of coverage in the regions where variant calls were made. Improvements to DNA sequencing methods for archived samples will significantly enhance cancer research and biomarker discovery, and will result in more reliable sequencing-based diagnostics.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1239W

Detection of somatic mutations in tumor genomes using *de novo* assembly with assembly to assembly mapping (DA-AM). *A.R. Carson[1], W. Pfeiffer[2], T. Schwartz[2], G. Oliveira[1], T. Nicholas[1], G. Zhang[1], M.A. Miller[2], E.J. Topol[1], S. Levy[1]*. 1) Scripps Translational Science Institute, Scripps Health, La Jolla, CA; 2) San Diego Supercomputer Center, University of California San Diego, La Jolla, CA.

Identification of somatic mutations in tumor genomes by standard variant detection strategies is currently not comprehensive. Typical sequencing approaches employ paired-end reads mapped against a reference genome and succeed in identifying a portion of all point mutations, indels, and structural variants in a sample. However, detection by such methods is limited by several factors, including increasing detection errors coupled with increasing mutation size. Here we propose a novel approach to contribute to the comprehensive detection of somatic mutations in tumors using *de novo* assembly, followed by assembly to assembly mapping (DA-AM). This strategy detects additional variants, including large indels, not found by traditional variant detection methods. We sequenced the exomes of a colon tumor-blood pair using the Illumina HiSeq platform, creating 100bp paired-end reads for *de novo* assembly using SOAPdenovo. Assembled contigs were then uniquely aligned to the human reference genome using ATAC, an assembly-to-assembly comparison tool. We then used unique alignments to identify indels in each exome. For comparison, we also mapped the reads using BWA and identified indels using the GATK variant detection strategy. Of the ~8,000 indels found in each tissue by DA-AM, 87% are also identified by GATK. Conversely, only 39% of the GATK variants are found by DA-AM. Interestingly, the size distribution of the DA-AM unique variants is significantly different from the DA-AM variants that match GATK. The DA-AM unique variants are 2–10x larger than their counterparts, and include a 927bp deletion found in both the blood and tumor. In contrast, GATK unique variants are typically smaller (2.5bp average versus 17.6bp for DA-AM unique variants). Additionally, DA-AM finds unique somatic mutations, including a 15bp tumor-specific insertion 5' of *SGK1*. Validation of these unique variants is currently underway, using multiple approaches including whole genome sequencing with Illumina and Ion Torrent platforms. In summary, the DA-AM methodology successfully identifies unique variants, including somatic mutations that often exceed the detection limits of traditional variant discovery approaches. The largest identified variants represent a size that, to date, has only been identified by arrays or paired-end mapping. As such, our strategy underscores the importance of *de novo* assembly in providing a more comprehensive set of somatic mutations in cancer genomes.

1240T

Identification of therapeutic targetable mutations in cancer by whole transcriptome and genome sequencing. *D.W. Craig, W. Liang, W. Tembe, A. Christoforides, J. Aldrich, T. Izatt, J.M. Trent, J.D. Carpten.* Dept Neurogenomics, TGen, Phoenix, AZ.

Recent technological advances in the form of Next-Generation Sequencing (NGS) technologies now provide us with platforms to interrogate entire human genomes at a fraction of the time and cost compared to more traditional sequencing technologies. Identifying targeting mutations or concepts is highly dependent on design of sequencing libraries, depending depth, regions sequenced, and type of material sequenced (RNA, DNA). In this work, we present a framework for integrated analysis of whole genome and transcriptome NGS data from matched tumor and normal cancer patients to identify genomic alterations occurring in advanced oncology settings. We compare methods for detection of actionable variants based on library preparation, such as between targeted sequencing and whole-genome sequencing. We show how use of multiple library preparation approaches can allow for greater resolution for detection of complex variants or low-frequency variants, such as use of through use of exomes and large-insert whole-genomes. For informing mutations, we show how use of transcriptomic data can assist in interpretation particularly in "N-of-1" cases and present an analysis pipeline for generating informative analyses for aiding interpretation of genomic events in the context of actionable targets.

1241F

Identification of Somatic Mutations in Parathyroid Tumors using Whole Exome Sequencing. *M.K. Cromer[1], L.F. Starker[2,3,5], M. Choi[1], R. Udelsman[2], C. Nelson-Williams[1], R.P. Lifton[1], T. Carling[2,3,4]*. 1) Genetics, Yale University, New Haven, CT; 2) Surgery, Yale University, New Haven, CT; 3) Endocrine Neoplasia Laboratory, Yale University, New Haven, CT; 4) Cancer Genetics and Genomics, Yale University, New Haven, CT; 5) Surgical Sciences, Uppsala University, Uppsala, Sweden.

The underlying molecular alterations causing sporadic parathyroid adenomas that drive primary hyperparathyroidism (pHPT) have not been thoroughly defined. Therefore, the aim of the study was to investigate the occurrence of somatic mutations driving tumor formation and progression in sporadic parathyroid adenoma using whole exome sequencing. Eight matched tumor-constitutional DNA pairs from patients with sporadic parathyroid adenomas underwent whole exome capture and high-throughput sequencing. Four of the eight tumors displayed a frameshift deletion or nonsense mutation in MEN1, which was accompanied by loss of heterozygosity (LOH) of the remaining wild type allele. No other mutated genes were shared among the eight tumors. One tumor harbored a Y641N mutation of the histone methyltransferase EZH2 gene, previously linked to myeloid and lymphoid malignancy formation. Targeted sequencing in an additional 185 parathyroid adenomas revealed a high rate of MEN1 mutations. Furthermore, this targeted sequencing identified an additional parathyroid adenoma that contained the identical, somatic EZH2 mutation that was found by exome sequencing. This study confirms the frequent role of LOH of chromosome 11 and MEN1 gene alterations in sporadic parathyroid adenomas and implicates a previously unassociated methyltransferase gene, EZH2, in endocrine tumorigenesis.

1242W

Exome sequencing of metastatic prostate tumors, GWAS and functional analysis of the methylation regulator TET2 in prostate cancer. *M. Dean[1], M. Nickerson[1], K.M. Im[1], K.J. Misner[1], W. Tan[2], H. Lou[2], D.W. Wells[2], K. Frederickson[3], T. Harkins[3], T. Naab[4], B. Gold[1], T. Andersson[2], B. Zbar[1], W.M. Linehan[5], G.S. Bova[6], H. Li[2], S. Anderson[1], M. Yeager[7]*. 1) Lab Experimental Immunology, NCI-FCRDC, Frederick, MD; 2) SAIC-Frederick, NCI-FCRDC, Frederick, MM; 3) Roche Diagnostics Corporation, Indianapolis, IN; 4) Department of Pathology, Howard University, Washington, DC; 5) Urologic Oncology Branch, NCI, Bethesda MD; 6) Department of Pathology, School of Medicine, Johns Hopkins University, Baltimore, MD; 7) Division of Cancer Epidemiology and Genetics, NCI, Bethesda, MD.

Metastatic prostate cancer (mPCa) has highly morbidity with poor control using existing therapies. The molecular basis of lethal mPCa is incompletely defined and markers that distinguish indolent from aggressive cancer are needed. We performed exome sequencing on metastatic tumors from an Ashkenazi Jewish patient who died of mPCa to identify genetic changes in metastatic disease and validate an experimental approach that avoided complications due to primary tumor heterogeneity. We show a germline mutation in *BRCA1* and somatic, nonsynonymous alterations in 22 genes in the primary tumor. An additional 31 genes were altered in all metastatic tumors defining a metastatic progenitor distinct from the primary tumor. Multiple, known cancer driver genes were altered, including *BRCA1*, *TMPRSS2-ERG*, *PBRM1*, and *TET2*. We sequenced 19 genes altered in this subject in metastases from additional mPCa patients and observed additional somatic, nonsynonymous changes in *TET2*, encoding a methylcytosine dioxygenase previously linked to PCa risk by genome-wide association studies. We characterized *TET2* germline missense variants and confirmed expression of the gene and the protein in prostate samples. We further mapped the location of the GWAS loci in the *TET2* gene and demonstrated multiple linked associated variants that are present in the promoter and in the first non-coding exons of the gene, and characterized alternative transcripts. Mobility shift assays identified novel DNA binding proteins, at the site of the associated variants, interacting in an allele-specific manner. The experimental approach provided one of the most comprehensive analyses to date of metastatic cancer.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1243T

Identification of Novel Driver Mutation in EGFR/KRAS/ALK-Negative Lung Adenocarcinoma of Never Smokers by whole exome sequencing. *S. Han[1], J. Yoon[4], J. Ahn[3], H. Jang[3], M. Lee[1], H. Kim[2], B. Cho[2], J. Lee[1].* 1) Department of Phamacology, Brain Korea 21 Project for Medical Science, Yonsei University, Seoul, South Korea; 2) Division of Medical Oncology, Yonsei University College of Medicine, Seoul, South Korea; 3) Department of Chemistry, Yonsei Univeristy, Seoul, Korea; 4) Hwasung Public Health Center, Hwasung, Republic of Korea.

Lung cancer in never smokers (LCINS), known for distinct pathogenesis and prognosis from lung cancer in smokers, are not fully understood. To identify novel driver mutation in lung cancer in never smokers (LCINS), we performed whole-exome sequencing of 20 normal and matched tumor samples from never smoked lung adenocarcinoma patients without EGFR, KRAS, and ALK mutations ("triple-negative adenocarcinoma"). The data revealed that candidate driver mutations in MLL3, YTDHF1, PIK3CG, ARID3A, and EPSTI1. Recurrently and frequently mutated genes were re-sequenced in-depth with capture probes. Validated mutations were further analyzed to detect significantly affected pathways. The whole-exome sequencing analysis provides new insights in unrevealed mutation profiles as well as the underlying mechanism of developing LCINS.

1244F

Detection of low frequency tumorigenesis variants by ultra deep sequencing (>1000x) across 409 oncogenes. *T.T. Harkins[1], C.C. Lee[1], M. Andersen[2], E. Levandowsky[1], M. Dindinger[1], J. Spangler[1], T. Ross[1], S. McLaughlin[1], V. Sheth[1].* 1) Life Technologies, Beverly, MA. 01915; 2) Life Technologies, Carlsbad, CA. 92008.

Understanding the importance of low frequency somatic mutations has been a barrier to understanding how tumors evolve and which mutations drive progression of the disease and which mutations are passengers. Most sequencing techniques to date have either used whole genome sequencing or exome (via hybridization enrichment) to sequence tumors to 30 to 100x average coverage. This depth of sequencing does not allow the detection of variants that occur below the 10 to 20% level. In this study, we have developed a method that combines ultra deep sequencing of a collection of normal, primary tumor and metastatic tumors with hierarchical clustering using allele frequencies. We have sequenced a collection of archived samples from individuals with stage III to V ovarian cancer. Sampling 8 to 20 sites per individual, including multiple sites from the primary tumor along with several metastatic sites, has allowed us to create a model to look at the relationships between samples based upon mutation profiles. To perform this study we targeted 409 oncogenes using Ion Torrent's AmpliSeq™ Comprehensive Cancer Panel. This panel uses 16,000 unique amplicons that are multiplexed in a total of 4 tubes, requiring a total of 40 ng of genomic DNA. This technique generated average sequence depths greater than 1000x coverage across all amplicons for all samples. Using this ultra deep sequencing method has allowed the identification of low frequency mutations and their allele frequency changes across the sample collection. Phylogentic relationships based upon hierarchical clustering within the sample sets were established using the somatic mutation allele frequencies for each sample.

1245W

Targeted sequencing of clinically relevant genes using TargetRich™ Gene Panels. *K. Jansen Spayd[1], I.A. Vasenkova[1], T. Shvetsova[1], D.A. Kloske[1], R.C. Bachmeyer[1], D.T. Moore[1], K.E. Varley[1,2].* 1) Kailos Genetics Inc., Huntsville, AL; 2) HudsonAlpha Institute for Biotechnology, Huntsville, AL.

Next-generation sequencing (NGS) holds great promise for many clinical applications, including personalized therapy selection, fetal DNA screening, and the diagnosis of rare genetic diseases. Most NGS approaches target the whole genome or exome and, for studies that are focused on clinically relevant regions of the genome, this results in low coverage of target regions or inefficient use of sequencer capacity. The transition of NGS into the clinic will require the ability to deeply sequence hundreds to thousands of targeted genomic regions from a large number of DNA samples. The drawbacks of current targeted enrichment methods include wasted sequencing capacity occupied by off-target DNA, requirement for large quantities of template DNA, high cost of reagents, low sample throughput, the need for specialized equipment, and inflexible design parameters. We developed TargetRich gene panels, based on Nested PatchPCR, for investigators to perform targeted deep sequencing of hundreds of genomic regions in large numbers of patient samples, including formalin fixed paraffin embedded samples (FFPE), to fully and efficiently utilize NGS platforms. Here, we describe the use of two TargetRich cancer research panels: a ten gene panel and a sixty three gene panel. We simultaneously amplified over 750 target sequences in a single tube from only 250 ng of patient DNA. The TargetRich Nested PatchPCR protocol can be performed on many samples in parallel in standard 96-well plates, can be completed in less than 9 hours, and requires no specialized laboratory equipment. We sequenced the targeted genes on three NGS platforms: Illumina GAIIx, MiSeq and Ion Torrent PGM. The high percentage of on-target sequence data and focused content allowed us to achieve deep coverage across the loci (average of 400X read depth). The accuracy of germline variant detection was tested and confirmed on samples that were sequenced as part of the 1000 Genomes Project. In addition, we tested the sensitivity of the method for detecting sequence variants introduced at lower frequencies, and found that we could detect variants present in less than 1% of sample, demonstrating that the method is sufficiently sensitive for detecting somatic mutations in heterogeneous tumor samples or fetal DNA mutations in circulating maternal DNA. Together, these results demonstrate the potential utility for TargetRich Nested PatchPCR for targeted sequencing in a variety of clinical applications.

1246T

Rapid and economical re-sequencing of hundreds of genes from sample-limited specimens using the Ion AmpliSeq™ Comprehensive Cancer Panel and Ion PGM™ semiconductor sequencing. *D. Joun, SM. Chen, B. Kong, I. Causga, CY. Li, D. Ruff, R. Bennett, M. Shannon.* Ion Torrent, Life Technologies, South San Francisco, CA.

The Ion AmpliSeq™ Comprehensive Cancer Panel (CCP) interrogates a set of more than 400 genes that have been associated with a broad spectrum of human cancers. CCP leverages the transformative Ion AmpliSeq™ and Ion PGM™ semiconductor sequencing technologies to amplify the coding exons of 409 genes and enable detection of known and de novo mutations within these genes. The panel employs approximately 16,000 primer pairs that are distributed across 4 pools to selectively amplify target regions in parallel in highly multiplexed PCRs and requires only 40ng of DNA sample. In addition, with an amplicon length of 125–175bp, the panel is tailor-made for analysis of highly fragmented DNA samples such as those obtained from formalin-fixed paraffin-embedded (FFPE) specimens. In this study, we demonstrated the broad utility of the panel for robust detection of genetic variation in sample-limited archived fresh-frozen and FFPE specimens. We specifically assessed the genetic variation present in FFPE and fresh-frozen sample pairs, including normal (breast or liver) tissue versus primary tumor and primary tumor versus metastatic tumor. We established that the CCP performs equivalently with FFPE and high molecular weight DNA samples in terms of target coverage uniformity and reproducibility of single nucleotide variant detection. We also identified 3 known coding variants in TP53 (P72R), ARID2 (S587G), and HNF1A (I27L) in an archived normal liver tissue FFPE sample. In addition, we investigated TP53 as a global biomarker across various tumor types as well as several markers associated with breast cancer, including ERBB2, and GATA3. Although we found no differences in these genes between the tumor and normal samples tested, we uncovered variants in several other genes in the CCP, including both coding and non-coding variants, in the same samples. These findings underscore the benefits in terms of speed, simplicity, and target coverage of using the Ion AmpliSeq™ Comprehensive Cancer Panel with the Ion PGM™ sequencer to characterize hundreds of genes, especially for sample-limited specimens.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1247F

Bayes-MutSig: A novel approach to determine significantly mutated genes. *N.B. Larson, H. Sicotte, J. Sinnwell, K.R. Kalari, E. Wieben, L. Wang, J.C. Boughey, M. Goetz, R. Weinshilboum, B.L. Fridley.* Mayo Clinic, Rochester, MN.

Cancer is a complex genetic disease that is the result of multiple genetic mutations. One approach to identifying etiologically relevant genes is determining which genes are disproportionately mutated in tumor samples relative to the rest of the tumor genome. If we consider somatic mutation rates to differ based on the sequence context of the base-pair, as in the case of methylation-related mutation of CpG sites, it is necessary to estimate context-specific mutation rates to more accurately assess statistical significance. We propose a Bayesian approach, referred to as Bayes-MutSig, to estimate single- and multiple-sample background somatic mutation rates for various mutation categories using a logit-linear binomial model. This approach not only allows us to share information across genes, but incorporate prior knowledge about background mutation rates. Point estimates derived from the posterior distributions of the model are then used to identify significantly mutated genes for that individual relative to background mutation, obtaining gene-wise p-values and false discovery rate estimates. We applied this method to whole exome sequencing data of two paired tumor-normal samples from MC1137 (BEAUTY), a prospective neoadjuvant breast cancer study. Breast cancer biopsies (tumor) and matched patient blood samples (normal) were obtained from two patients with locally advanced breast cancer. We evaluated 18,394 non-overlapping genes using our Bayes-MutSig approach and identified significantly mutated genes for each sample.

1248W

The transcriptional landscape and mutational profile of lung adenocarcinoma. *W. Lee[1,3], Y. Ju[4], J. Shin[1,5], J. Lee[1,6], B. Thomas[1], J. Lee[1], Y. Jung[8], J. Kim[9], J. Shin[9], S. Yoo[5], J. Kim[5], E. Lee[4], C. Kang[8], I. Park[8], H. Rhee[4], S. Lee[6,7], J. Kim[1,2,3,5], J. Kang[10], Y. Kim[1,8], J. Seo[1,2,3,4,5].* 1) Genomic Medicine Institute (GMI), Medical Research Center, Seoul National University, Seoul 110–799, Korea; 2) Department of Biochemistry, Seoul National University College of Medicine, Seoul 110–799, Korea; 3) Department of Biomedical Sciences, Seoul National University Graduate School, Seoul 110–799, Korea; 4) Macrogen Inc., Seoul 153–781, Korea; 5) Psoma Therapeutics Inc., Seoul 153–781, Korea; 6) Department of Internal Medicine, Seoul National University Hospital, Seoul National University College of Medicine, Seoul 110–799, Korea; 7) Department of Internal Medicine, Seoul National University Hospital, Cancer Research Institute, Seoul National University College of Medicine, Seoul 110–799, Korea; 8) Department of Thoracic and Cardiovascular Surgery, Cancer Research Institute, Seoul National University College of Medicine, Seoul 110–799, Korea; 9) Division of Medical Oncology, Research Institute of Medical Science, The Catholic University of Korea, Seoul 137–040, Korea; 10) Division of Medical Oncology, Seoul St. Mary's Hospital, The Catholic University of Korea, Seoul 137–040, Korea.

Understanding the molecular signatures of cancer is important to apply appropriate targeted therapies. Here we present the first large scale massively parallel RNA sequencing study of lung adenocarcinoma demonstrating its power to identify somatic point mutations as well as transcriptional variants such as gene fusions, alternative splicing events and expression outliers. Our results revealed the genetic basis of 200 lung adenocarcinomas in Koreans including deep characterization of 87 surgical specimens by transcriptome sequencing. We identified driver somatic mutations in cancer genes including EGFR, KRAS, NRAS, BRAF, PIK3CA, MET and CTNNB1. New cancer genes, such as LMTK2, ARID1A, NOTCH2 and SMARCA4, were also suggested as a candidate for novel driver mutation. We found 45 fusion genes, 8 of which were chimeric tyrosine kinases involving ALK, RET, ROS1, FGFR2, AXL and PDGFRA. Of 17 recurrent alternative splicing events, we identified exon 14 skipping of MET proto-oncogene as highly likely to be a cancer driver. The number of somatic mutations and expression outliers varied markedly between individual cancers and was strongly correlated with smoking history of cancer patients. Our findings broaden our understanding of lung adenocarcinoma and may also lead to new diagnostic and therapeutic approaches.

1249T

Genes with Single Nucleotide Variations in Early-onset Female Breast Cancer Patients Identified through Exome Sequencing. *C. Lee[1], N. Leng[2], W. Kuo[3], H. Yang[4], K. Nobuta[2], C. Lin[5], C. Chang[6], Y. Lu[5], K. Lo[1], L. Hu[7], H. Chu[7], W. Chou[7], C. Chen[7], W. Yao[7], K. Chiu[1], A. Cheng[5], C. Shen[7], K. Chang[3, 8], C. Haudenschild[2], C. Chen[1].* 1) Genomics Res Ctr, Academia Sinica, Taipei, Taiwan; 2) Genomic Services, Illumina Inc., Hayward, CA, US; 3) Department of Surgery, National Taiwan University Hospital, Taipei, Taiwan; 4) Molecular and Genomic Epidemiology Research Center, China Medical University Hospital, Taichung, Taiwan; 5) Department of Oncology, National Taiwan University Hospital, Taipei, Taiwan; 6) Department of Medical Research, National Taiwan University Hospital, Taipei, Taiwan; 7) Institute of BioMedical Sciences, Academia Sinica, Taipei, Taiwan; 8) Department of Surgery, Cheng Ching General Hospital, Taichung, Taiwan.

Female breast cancer patients diagnosed in young age showed different characteristics to patients diagnosed in old age. To investigate candidate cancer driver genes in early-onset female breast cancer, exome sequencing was performed by Illumina GAIIx system, which obtained using 2×75 bp pair-end sequencing. The average coverage in both blood and tissue samples were 127x, and the average on-target rate were 98%. Paired blood and tumor tissue samples were collected from 84 patients with breast cancer diagnosed at an age of 40 years or younger. Patients were classified into four subtypes according to the expression levels of ER, progesterone receptor (PR), human epidermal growth factor receptor 2 (HER2) and Ki67 antigen: luminal A [ER(+), PR(+/−), HER2(-), Ki67low] (36 cases), luminal B [ER(+), PR(+/−), HER2(-), Ki67high] (13 cases), HER2 positive [ER(+/−), PR(+/−), HER2(+)] (23 cases), and triple negative [ER(-), PR(-), HER2(-)] (12 cases). Using hg19 as a reference, single nucleotide variations (SNVs) in exons were identified by CASAVA1.7 and annotated by ANNOVAR software. With a cut-off value of 30x coverage, nonsynonymous SNVs and synonymous SNVs located in conserved regions of 46 species were included in the analysis. Comparing variations in all paired samples, SNVs only observed in tumor sample were defined as Pure Tumor only (PT) mutations, which were somatic mutations. SNVs observed in tumor samples and in both blood and tumor samples were defined as Tumor only and Same (TS) mutations, which were somatic mutations in some patients and might be germline mutations in other patients. A total of 2466 PT mutants in 2054 genes and 1601 TS mutants in 1340 genes were observed in 84 patients. After comparing with 47 controls to filter polymorphisms in normal population, there were 2326 PT or TS mutants demonstrated in 2323 sites of 1965 genes. Gene with the highest mutation number was observed in 20 (23.8%) patients, while 12 genes harboring mutation sites in more than 9 (10.7%) patients. Of 1965 mutated gene, only 19 (1.0%) genes showed significant distributions among four subtypes.

1250F

Pathway analysis of somatic mutations in the whole genome sequence of aggressive tumors in African American prostate cancer patients. *K. Lindquist[1], R. Kazma[1], J.A. Mefford[1], T.J. Hoffmann[1], N. Cardin[1], B.A. Rybicki[2], D.A. Chitale[2], A. Levin[2], J.S. Witte[1].* 1) University of California, San Francisco, San Francisco, CA; 2) Henry Ford Health System, Detroit, MI.

Prostate cancer is the most common cancer and the second leading cause of cancer mortality among men in the United States. African Americans have not been included in many cancer genome studies, though they are more likely to have aggressive tumors than other American men are. Previous studies have characterized the role of chromosomal rearrangements in aggressive prostatic tumors. However, the role of small (1–50 nucleotide) somatic substitutions or insertions/deletions (indels) in these tumors is poorly understood. Studying these mutations in a pathway context, rather than individually, may help to elucidate their role. In this study, we used whole genome sequence data generated by Complete Genomics from five African American patients with aggressive prostate cancer, and identified all genomic modifications that occurred only in the tumor cells. We selected potentially functional mutations that occurred in genes (including coding regions, introns, transcription start sites, and untranslated regions) for inclusion in pathway analyses. We excluded mutations in regions susceptible to false positive calls, such as regions overlapping known repetitive sequences. For each sample, we ranked the genes using a formula that took into account the confidence that each mutation is a true positive finding. Using the Kyoto Encyclopedia of Genes and Genomes (KEGG) and other pathway databases and tools, such as the Protein ANalysis THrough Evolutionary Relationships (PANTHER) classification system, we determined whether the top ranking mutated genes are overrepresented by particular pathways. Our preliminary results showed that these genes implicate several pathways, some of which were previously known to be involved in prostate cancer (such as p53 and growth factor signaling pathways), and others which may be worth further investigation. These results demonstrate that pathway analyses can be a powerful strategy for studying small mutations in aggressive prostate tumors. This approach may be also valuable for other cancers and complex diseases.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1251W

SeqWright's Next-Gen Sequencing of FFPE Tumor Tissue Allows for Analysis of Multiple Cancer Markers With High Sensitivity and Excellent Correlation to Other Validated Clinical Methods. *A.C. Pond[1], X.X. Tan[1], L.T. Szkotnicki[1], V. Venegas[1], E. Zhou[1], Y. Mou[1], K.B. Thomas[1], P. Choppa[2], K. Guekunst[2], F. Lu[1]*. 1) SeqWright, Inc. 2575 W. Bellfort St. Houston TX, 77054; 2) Clarient, Inc. 31 Columbia. Aliso Viejo CA, 92656.

Next-generation sequencing (NGS) is quickly becoming a proven and valuable tool in many clinical applications, including drug discovery, the development of companion diagnostics and patient screening. In light of this, we sought to demonstrate the clinical utility of NGS by sequencing key mutational hotspots within a number of oncogenes and tumor suppressors using two leading cancer sequencing panels; the Ion AmpliSeq™ from Life Technologies and the TruSeq™ Cancer Panel from Illumina. NGS on gene panels holds many advantages over other biomarker detection methodologies. Through deep, targeted sequencing, NGS can obtain similar sensitivity of proven methodologies, such as allele-specific PCR (AS-PCR), while interrogating thousands of loci in parallel. This, along with the ability of NGS to discern individual reads, allows for the quantification of allele frequency and gives these systems the power to resolve complex mutations otherwise undetectable by other leading methods. We were able to successfully sequence multiple genes in parallel, with nanogram quantities of formalin-fixed DNA from several clinical samples. The data demonstrate a high concordance between the two platforms' (MiSeq™ and Ion Torrent™) ability to quantitate allele frequency, an indication of their overall accuracy. In addition, NGS demonstrated one hundred percent concordance with clinical AS-PCR genotyping and Sanger sequencing results and proved exceptional in several key aspects, including the ability to calculate percent allele frequency and to resolve complex mutations otherwise undetectable by AS-PCR and Sanger sequencing. An overall analysis of the workflow and sample requirements illustrate the feasibility of applying NGS in a clinical setting with nanogram quantities of micro-dissected, formalin-fixed samples. Overall, these studies directly demonstrate the fundamental advantages of NGS in clinical mutational screening and illustrate SeqWright's proven ability to leverage the full potential of leading technologies to help our customers discover clinically-relevant biomarkers and successfully develop new diagnostic tools.

1252T

Amplicon sequencing of tumors from xenograft, FFPE and fresh frozen samples. *B. Riley-Gillis, C. Qiu, O. Puig, R. Benayed.* Genetics and Genomics Laboratory, Translational Reasearch Sciences(TRS), Pharma Research and Early Development, Hoffmann-La Roche, Nutley, NJ.

Formalin-fixed, paraffin-embedded (FFPE) tumor samples are an invaluable resource for cancer research, drug development, and diagnosis. However, utilizing these samples is complicated by the nucleic acid degradation and base modification that occurs during the fixation and storage processes. In this study, we use the TruSeq Amplicon Cancer Panel and the MiSeq platform to assess the feasibility of sequencing tumors from FFPE, fresh frozen (FF), and xenograft samples. The Cancer Panel is a pre-designed Illumina panel that targets >35kb of sequence with 212 amplicons in 48 tumor suppressor and oncogenes. The quick turnaround time, ability to detect low frequency mutations in FFPE samples, and coverage of critical cancer genes make this panel ideal for rapid, sequence analysis of DNA from different types of tumor tissue. As a pilot study, we initially sequenced 4 matched FF and FFPE samples, 2 xenograft samples, and 1 HapMap control. Including the Illumina control sample, we had a total of 8 pooled samples with mean depth coverage for FF samples 4687x, FFPE samples 2249x, and for xenograft samples 4257x. After filtering variants based on frequency and quality score, the 2 pairs of FF/FFPE samples had 48 and 53 overlapping variants, respectively. Of these overlapping variants, 2/48 and 1/53 are missense mutations reported in COSMIC, the remaining are dbSNPs or non-coding variants. In a follow-up experiment, we sequenced the matched FF/FFPE samples in a 6-plex format and observed deeper coverage, as expected, but also higher quality scores for the low frequency mutations. Additionally, we identified 48 and 42 variants in 2 separate xenograft samples and are currently comparing the data to exome sequencing data. The results indicate that DNA from FFPE tumor samples can be successfully sequenced and yield high quality data. We are currently sequencing 20 additional matched FF/FFPE samples to thoroughly investigate the quality of variants detected from FFPE samples and define the coverage needed to call low frequency variants with high confidence. The ability to capture the significant information available in FFPE samples will be critical to deepen disease understanding and enhance cancer therapy. Additionally, baseline assessment of tumors from patients in clinical trials using targeted sequencing platforms is necessary to define tumor mutations, stratify patient populations and evaluate/predict therapy response.

1253F

Loss of the canonical Notch mediator RBPJ is recurrent in oligoastrocytomas. *P. Salo[1], E.I. Gaál[1], O. Tynninen[2,3], M. Niemelä[4], A. Laakso[4], A. Karppinen[4], A. Paetau[2,3], H. Mäenpää[5], J. Hernesniemi[4], M. Perola[1,6,7]*. 1) Public Health Genomics, Natl Inst Health & Welfare, Helsinki, Finland; 2) Departmentof Pathology, Haartman Institute,University of Helsinki,Helsinki,Finland; 3) Helsinki University Central Hospital, Helsinki, Finland; 4) Department of Neurosurgery, Helsinki University Central Hospital, Helsinki, Finland; 5) Department of Oncology, Helsinki University Central Hospital, Helsinki, Finland; 6) Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland; 7) University of Tartu, Estonian Genome Centre, Tartu, Estonia.

Gliomas are neoplasms of the central nervous system arising from cells of the glial lineage. WHO grade I gliomas are curable with surgical resection whereas grade II-IV gliomas cannot be completely resected, usually resulting in fatal recurrences. As a part of a personalized cancer care project, we sequenced the whole exome and transcriptome of a single oligoastrocytoma in order to gain insight into its genetic landscape. We detected and experimentally validated 34 nonsilent somatic mutations. Only one of these, R132H in IDH1, has a confirmed role in the pathogenesis of gliomas. However, intriguingly, the tumor had lost both alleles of the canonical Notch mediator RBPJ, one by a deletion of most of chromosome 4 and the other via a frameshift deletion. Notch signalling is a conserved pathway regulating apoptosis, proliferation and differentiation of cells in a variety of situations, including the differentiation of neural crest stem cells into glia. Several lines of evidence have suggested Notch signalling may contribute to the initiation or maintenance of gliomas but no conclusive data have been presented. Subsequently, we compared the tumor with a published series of 7 oligodendrogliomas and found another tumor which had lost the wild-type function of RBPJ, one allele via loss-of-heterozygosity (LOH) and the other by a somatic missense mutation. Based on these observations, we sequenced the protein-coding regions of RBPJ in an independent sample of seven oligodendrogliomas and three oligoastrocytomas. The tumors were selected to match the index case on the basis of carrying LOH at 1p/19q (a hallmark of oligodendrogliomas and oligoastrocytomas) and in the genomic region containing RBPJ on chromosome 4. Of these ten additional tumors, one carried a frameshift deletion and another one a nonsense mutation in the remaining copy of RBPJ. In effect, altogether 3 out of 11 or 27% of the gliomas we investigated had conclusively lost both alleles of RBPJ via a combination of LOH and nonsense or frameshift mutations. Furthermore, these three tumors had all been classified as oligoastrocytomas, so 3 out of 4 oligoastrocytomas had lost RBPJ. Our data add to the notion that Notch signalling is involved in the pathogenesis of malignant gliomas. They also suggest that the role played by Notch is most important in gliomas with an astrocytic component. This may aid molecular sub-classification of malignant gliomas and hence contribute to improved treatment.

1254W

Characterization of a Small Cell Prostate Cancer using Exome Sequencing. *A.F. Scott[1], D.W. Mohr[1], H. Ling[1], G.S. Liptak[2]*. 1) Inst Gen Med,Blalock 1034, Johns Hopkins Med Sch, Baltimore, MD; 2) Dept of Pediatrics, SUNY Upstate Medical Center, Golisano Children's Hospital, Syracuse, NY 13210.

We have performed genotyping and exome sequencing of a small cell prostate carcinoma (SCPC) and normal control. SCPC is a high-grade malignant tumor with neuroendrocrine differentiation. It accounts for fewer than 1% of all prostate cancers and has a poor prognosis. Unlike the more common prostatic adenocarcinomas, SCPC does not present with elevated PSA levels and is often discovered after the occurrence of metastases. No molecular characterization has been presented for this tumor type to our knowledge.

We isolated DNA from a metastatic Virchow node and used saliva (mouthwash) as the normal control sample. DNA was genotyped on a 1M Illumina array and exome sequenced on an Illumina HiSeq 2000 following library capture with an Agilent SureSelect All Exon v3 kit. Sequence data was processed using CIDRSeqsuite v2.3.0 and analyzed with interpretation software tools from Knome. The mean read depth for the tumor was 353 and 138 for the normal exome.

The genotyping array showed a large degree of apparent chromothripsis (copy number variation, loss of heterozygosity, etc.) with 17 of the 22 autosomes grossly affected. Although the number of nucleotide substitutions between tumor and normal were fewer than seen in many cancers we noted that the tumor contained a region of LOH affecting most of 17p including the tumor suppressor TP53. The retained allele contained a somatic substitution in the coding region not seen in the normal exome that was predicted to be damaging by SIFT and PolyPhen2 (H179R). A predicted “function changing” somatic nonsense mutation was also seen in RB1. Other potentially functional somatic mutations were observed in several genes including FOXA1, CNTD2, and ZNF8. Knome reported “phenotype-implicated” variants in the normal exome in the genes SRD5A2 and CREB3L1 although their role in this cancer remains to be determined. We conclude that the loss of a normal allele and the somatic mutation of TP53 are likely to be the major lesions in this tumor. Whether they occurred before or after chromothripsis remains unknown

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1255T

Profiling mutations in circulating tumor cells from breast cancer patients by targeted sequencing. *Y. Shen[1,2], X. Xu[3], W. DeWitt[2], N. Xu[1], F.Z. Bischoff[4], K.D. Crew[3], D.L. Hershman[3], M.A. Maurer[3], R. Parsons[3], K. Kalinsky[3].* 1) Center for Computational Biol & Bioinformatics, Columbia Univ, New York, NY; 2) Columbia Genome Center, Columbia Univ, New York, NY; 3) New York Presbyterian Hospital, Columbia University Medical Center; 4) Biocept Inc, San Diago, CA.

The genomic signature of circulating tumor cells (CTCs) may serve as a surrogate marker for accurate description of the metastatic tumors of interest, especially in the setting of treatment response. Here we describe a method to profile mutations in CTCs from breast cancer patients with metastatic tumors. Method: First we enriched CTCs from blood samples in metastatic breast cancer patients using a microfluidic device that contains a cocktail of antibodies against both epithelial and mesenchymal markers. We then use emulsion-based multiplex-PCR to generate amplicons of genes that harbor known mutations, and then perform massive-parallel semiconductor-based sequencing (Ion Torrent) of the amplicons. We sequenced at very high depth of coverage (>1000x) in order to detect cancer somatic mutations present in CTCs that are only the 1–3% range. We recalibrated the quality score of mismatch bases based on the context of sequence including homopolymers. We inferred the somatic mutations by testing statistical significance using a beta-binomial distribution model. Results: we were able to enrich CTC samples in 9 of 17 samples (CTC range 1–1063, mean=134), with CTC purity ranging from 0.3% to 6%. We found known missense mutations at E545D on PIK3CA, F354L on STK11, and Q61R on NRAS; we also found novel mutations of L540F and Q1033K within the hot spot regions of PIK3CA. We analyzed a healthy female concurrently as a negative control, in which none of the aforementioned mutations were observed. Conclusions: We developed a new method to enrich CTCs in breast cancer patients and characterize somatic mutations in CTCs. This provides a foundation for the future studies to describe the clonal lineage of primary tumors, CTCs and metastatic lesions.

1256F

Exome and Whole Genome Mutational Landscape in Pancreatic Ductal Adenocarcinoma. *L. Timms, A. Panchal, L. Mullen, J. Johns, R. Denroche, R. De Borja, F. Yousif, Z. Zha, M. Sam, A.M.K. Brown, T. Beck, J.D. McPherson.* Genome Technologies, Ontario Inst Cancer Research, Toronto, Ontario, Canada.

Pancreatic cancer is the fifth leading cause of cancer deaths with a poor prognosis and 5-year survival rates of ~2%, a result of late diagnosis, tumor resistance to current therapies and high metastatic potential. As a contributing member of the International Cancer Genome Consortium (ICGC), the Ontario Institute for Cancer Research (OICR) has committed to sequence 375 independent pancreatic adenocarcinoma tumors and their matched controls, in an effort to catalogue genomic abnormalities with the goal of identifying diagnostic and prognostic biomarkers and therapeutic targets. To date, whole exome sequencing has been performed on over 85 matched pancreatic primary tumor-normal and or xenograft-normal pairs using the Illumina Genome Analyzer IIX and HiSeq 2000 sequencers with a minimum of 100x target coverage. Sequence alignment and variant calling were performed using Novoalign and GATK, respectively. On average, 33 somatic non-synonymous single nucleotide variants (SNV) and 13 somatic small insertion or deletions (indels) (up to 15bp) were identified in each tumor sample. Verification of somatic coding, non-synonymous SNV and indels was performed using PCR amplification followed by sequencing on the Ion Torrent Personal Genome Machine (PGM). Our results confirm several known genomics abnormalities in PDAC, such as KRAS, p53 and SMAD4. In addition, ultra-deep KRAS (exons 2 & 3) sequencing of over 140 primary tumors was performed using the Ion PGM as a method of tumor cellularity estimation, as this is a near universal driver in PDAC. Whole genome sequencing (WGS) of primary tumors with cellularity over 20% and matched normal controls is underway, with over 25 pairs sequenced to complement the exome data; verification of variants identified from WGS is in progress. All somatic variant data are being released thorough the ICGC portal (www.icgc.org).

1257W

De novo germline mosaic BRCA1 exon deletion associated with bilateral breast cancer. *M. Tischkowitz[1,2], I. Delon[2], A. Taylor[2], A. Molenda[2], J. Drummond[2], K. Oakhill[2], A. Girling[4], H. Liu[3], J. Whittaker[2], R. Treacy[2].* 1) Department of Medical Genetics, University of Cambridge, Cambridge, UK; 2) East Anglian Medical Genetics Service, Addenbrooke's Hospital, Cambridge, UK; 3) Molecular Malignancy Laboratory and Department of Histopathology, Addenbrooke's Hospital, Cambridge, UK; 4) Department of Histopathology, Norfolk and Norwich University Hospital, Norwich, UK.

De novo BRCA1/BRCA2 mutations are very rare and to date only eight such cases have been reported. Here we present a case of a de novo BRCA1 mosaic deletion for exon 16. A 48-year-old woman with no family history presented with two synchronous right-sided grade III ER- PR- breast cancers at age 39 and a left sided grade III triple negative breast cancer with basal features at age 47. Complete sequencing of and MLPA analysis of BRCA1/BRCA2 revealed a low level heterozygous BRCA1 deletion detected by MLPA. This was confirmed independently by the presence of the p.Ser1613Gly polymorphism present at 30% compared to a heterozygous control (50%). The deletion was not present in either parent. Long-range PCR was used to identify the breakpoints of the deleted fragment and we propose a possible molecular mechanism for the deletion by recombination via a 38 nucleotides repeat present on both sides of the breakpoint. Further analysis of p.Ser1613Gly in both normal breast tissue and the cancers showed loss of BRCA1 heterozygosity in the right sided cancer but not the left cancer. Interestingly LOH status correlated with lymph node metastases (2/18 positive on the right, 0/17 on the left). In conclusion we believe that this is the first reported case of a proven de novo germline mosaic BRCA1 mutation and this can cause a similarly strong phenotype to a full heterozygote.

1258T

Single-nucleotide variant calling from low-coverage single-cell sequence data using tumor population structure. *S. Vattathil[1,2], N. Navin[3], P. Scheet[2,1].* 1) Human & Molecular Genetics Program, Graduate School of Biomedical Sciences, Univ. of Texas at Houston, Houston, TX; 2) Dept. of Epidemiology, Univ. of Texas MD Anderson Cancer Center, Houston, TX; 3) Dept. of Genetics, Univ. of Texas MD Anderson Cancer Center, Houston, TX.

Advances in single-nucleus sequencing (SNS) have created an unprecedented opportunity to study somatic mutations and clonal diversity at high resolution. Analysis of different types of mutations, such as single-nucleotide variants (SNVs), copy number variants, and indels, may provide complementary views of tumor progression. The SNS method uses next-generation sequencing technology, but requires additional whole-genome amplification which makes it more error-prone than traditional tissue sequencing. Further, a comprehensive survey of the tumor requires sequencing of many cells, making the experiments expensive. Statistical methods to improve the accuracy of mutation calling at realistic low sequencing depths allow the analysis to be applied to more samples, increasing power to define relevant mutations and describe the tumor genomic landscape.

Here we investigate how structure within the tumor cell population can be used to improve SNV calls from low-coverage single-cell sequence data. Multisample genotype calling methods have been popularly applied to low-coverage data in emerging large-scale surveys of human population genetics. However, tumor cell populations have a unique feature that may be exploited, namely that the natural history of a tumor cell population can be defined at the level of the entire genome, instead of at the local haplotype level as is the case for samples of germline genomes. That is, a much larger set of observed data from a sample of tumor cells may be used to inform the genotype calls at a given site in an individual cell.

In our method, which we call Single-Nucleus Genotype Likelihood-based Estimation (SiNGLE), we first define initial genotype estimates for each cell given the observed allele-specific read counts and the somatic mutation allele frequency at each site, estimated from the SNS data or from high-coverage sequencing of a matched tumor tissue sample. We use the genotype estimates to make inference on the structure underlying the sampled cells, and then redefine the genotype probabilities conditional on the inferred population structure. We apply the approach to experimentally-obtained single-cell and tissue sequence data from a breast tumor and a breast tumor cell line, and also evaluate the robustness of the method by testing genotype inference from simulated tumor sequence data generated under several models of tumor evolution.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1259F

Initial genomic analysis of a pure erythroid leukemia developing in association with hydroxyurea treatment for sickle cell anemia. *Z. Wang[1], D. Darbari[1,2], Z. McIver[1], I. Maric[3], L. Diaw[1], Y. Song[4], P. Johansson[4], J.B. He[4], J. Wei[4], A.J. Barrett[1], J. Khan[4], J.G. Taylor[1]*. 1) Hematology Branch, NHLBI, NIH, Bethesda, MD; 2) Children's National Medical Center, Washington, DC; 3) Laboratory Medicine, Clinical Center, NIH, Bethesda, MD; 4) Pediatric Oncology Branch, CCR, NCI, Bethesda, MD.

The antimetabolite hydroxyurea (HU) is the only widely available therapy for reducing complications of sickle cell anemia (SCA). It is not widely prescribed due to concerns about adverse effects, including cancer. Insufficient data currently exist regarding its leukemogenic potential. HU associated leukemia in SCA is rare, with only 6 literature reports. Whether such cases are coincidental or related to HU remains unanswered. We identified a 33 year old male with pure erythroid leukemia (formerly M6 AML) and a unique combination of SCA, HU treatment, and a short 4 year treatment latency. Cytogenetics showed a 5q rearrangement, 7q deletion, losses of chromosomes 15–22 and Y, suggesting a therapy related etiology. To date, no SCA AML cases have undergone in-depth genomic analysis to determine if HU has an etiologic role in therapy related leukemogenesis. High density SNP genotyping was performed to understand progressive disease changes. Analysis of germline DNA prior to HU and tumor DNA identified 10 mosaic aneuploidy events at diagnosis including losses at 5q, 7q and 17p and gains at 19p. High throughput parallel sequencing was also performed across a 38 megabase exome from the patient's germline, tumor and relapsed tumor DNA. We reasoned that somatic mutations in M6 AML might occur more frequently in genes regulating erythroid differentiation and proliferation. Somatic mutations were not present in commonly mutated AML genes like *NPM1*, *RUNX1* or *IDH1*, although 33 putative somatic missense mutations were detected including a validated somatic *TP53* mutation. Genome-wide cytogenetic band enrichment analysis of exome sequence data suggests the presence of a significant cluster of mutations along chromosome 19p. Together, the mosaic aneuploidy and exome sequence analysis suggest a region of 19p is potentially important to the pathogenesis of pure erythroid leukemia. This region extends at least 1.9 Mb and includes 58 genes. We are validating mutations in this region and exploring the role of 19p during stress erythropoiesis to further determine the significance of this finding. Investigation of this rare leukemia case offers a unique opportunity to elucidate the pathogenesis of erythroleukemia and possibly to determine if HU therapy is associated with leukemogenesis.

1260W

Novel L1 and Alu Retrotransposon Insertions in Cancer Related Gene Loci in the NCI-60 Cancer Cell Lines. *J. Zampella[1, 2], K. Burns[1]*. 1) Johns Hopkins, Baltimore, MD; 2) Howard Hughes Medical Institute, Chevy Chase, MD.

Long Interspersed Elements (L1 LINE) and Alu family Short Interspersed Element (SINE) are active Transposable Elements (TEs) in the human genome which have been implicated in oncogenesis. The functional roles of TEs remain under-explored due to the inability of current methods to locate these high copy number repeats in the genome-a barier recently overcome. Using a new genomic method, this study aims to discover cancer-specific L1 LINE and Alu SINEs in the NCI-60 panel.Using transposon insertion profiling microarray (TIP-chip) on genomic DNA from the NCI-60 tumor cell lines, vectorette PCR selectively amplified the 3' end of TEs to yield unique DNA fragments. These unique PCR amplicons were fluorescently labeled and hybridized to a custom genomic tiling microarray. Contiguous positive probes were detected as peaks. Data analysis included filtering peaks for `true' insertions and annotating their locations with respect to known TEs and cancer-related genes. We report 853 candidate L1 insertions (199 known and 654 novel) and 2,889 candidate Alu insertions (1,249 known and 1,640 novel) across the NCI-60 panel. There were 506 cancer genes within 5kb of the predicted L1 insertion intervals and 224 putative L1s within 500 bps of an exon. Relevant genes included RB1, ALK, IGF1R and BCL2. 1,027 cancer genes including BRAF, KRAS, BRCA2 and EGF were identified within 5kb of the predicted Alu insertions. 4 of the candidate Alu insertions were within 500 bps of an exon and 377 Alu insertion intervals mapped to the first intron of a gene. This is the first study to use a high through-put method to discover novel TE insertions in the NCI-60 tumor cell lines. L1 and Alu TEs may affect gene expression, alter genomic stability or act as insertional mutagens in neoplastic cells. Discovering L1 and Alu locations within the genome is the first step to understanding these TEs and their roles in oncogenesis. This study affirms that TIP-chip provides an effective high through-put method by which novel L1 and Alu insertions may be discovered. The discovery of these novel L1 and Alu insertions in proximity to known cancer genes provides specific targets to help elucidate the relationship between TEs and the development of different neoplasias. Ongoing work will evaluate the functional significance of these insertions.

1261T

Exome sequencing approach for identification of causative mutations in neurofibromatosis type 1 (NF1)-associated plexiform neurofibromas. *A. Pemov[1], H. Li[2], M. Wallace[2], D.R. Stewart[1], NISC Comparative Sequencing Program, NIH Intramural Sequencing Center, NHGRI, NIH, Bethesda, MD.* 1) NCI, NIH, Bethesda, MD; 2) University of Florida, Gainesville, FL.

Background. NF1 is an autosomal dominant tumor pre-disposition genetic disorder, caused by constitutive inactivation of one of two copies of the tumor suppressor *NF1*. Individuals with NF1 are prone to the development of both benign and malignant tumors including skin and plexiform neurofibromas (PNF). Although it is accepted that somatic inactivation of *NF1* is a necessary step in PNF development, little is known what other genes/pathways are disrupted/affected in the tumor. In this study, we performed whole-exome sequencing of 11 tumor samples matched with germline DNA obtained from 11 unrelated NF1 patients. **Methods.** Germline DNA was extracted from peripheral white blood cells and tumor DNA was obtained from primary Schwann cell (SC) cultures established from dissected plexiform neurofibromas. All cell cultures were of low passage and contained at least 70% SC. We used a capture kit from Illumina ("TruSeq") that targets roughly 60 million bases consisting of the CCDS annotated gene set. We sequenced the captured DNA on the Illumina HiSeq platform until we had sufficient coverage to call genotypes with an arbitrary quality score ("MPG") of 10 for at least 85% of targeted bases. We analyzed the data using an R-based statistical software designed to identify "driver" genes from "passengers." **Results.** First, we analyzed the *NF1* locus in both germline and tumor DNA. We identified pathogenic mutations in the *NF1* gene in all but one germline sample. All germline mutations were also confirmed in the tumor DNA. In addition, we identified somatic inactivation of *NF1* in nine out of 11 tumors. Second, we evaluated concordance of the exome sequencing data with genotyping data obtained from Illumina 2.5M SNP-arrays. Out of ~45,000 SNPs genotyped in both platforms, 97% were concordant. Finally, we analyzed the data to identify genes that could play an important role in PNF tumorigenesis. After correcting P-values for multiple testing we identified eight frequently mutated statistically significant genes with false discovery rate (FDR) below 0.05. Evaluation of biological functions of the genes revealed that such processes as cell cycle, nonsense-mediated mRNA decay and Notch pathway signaling might play an important role in PNF tumorigenesis. **Conclusions and further steps.** To our knowledge, this is the first attempt to identify "driver" mutations in NF1-associated PNF via exome sequencing. We are validating select mutations with the classic Sanger approach.

1262F

***DICER1* RNaseIIIb domain is mutated at two different sites in two different lesions in the same patient.** *M. Wu[1,2,3], N. Sabbaghian[1,3], N. Hamel[1,2], C. Choong[4], A. Charles[5], W. Foulkes[1,2,3]*. 1) Lady Davis Institute, McGill University, Montreal, Canada; 2) The Research Institute, McGill University Health Center, Montreal, Canada; 3) Program in Cancer Genetics, Departments of Oncology and Human Genetics, McGill University, Montreal, Canada; 4) School of Paediatrics and Child Health, University of Western Australia, Department of Endocrinology, Princess Margaret Hospital, Perth, Western Australia; 5) Dept Paediatric Pathology, Princess Margaret Hospital for Children, Perth, Western Australia.

DICER1 is an endoribonuclease responsible for the generation of microRNAs (miRNAs). miRNAs are responsible for the repression of many genes. *DICER1* has been found to be mutated in pleiotropic tumour predisposition syndrome (OMIM 601200). Families with germline mutations in *DICER1* are at increased risk of developing a variety of tumours associated with pleiotropic tumour predisposition syndrome. Here we present a patient who presented with a Wilms tumour and multinodular goiter. We discovered that the proband inherited one germline *DICER1* mutation (c.2117–1G>A) and acquired different somatic mutations in *DICER1* in her Wilms tumour (c.5429A>G) and in her multinodular goiter (c.5438A>G). We suggest that the acquired point mutations do not cause missense mutations but instead introduce exonic splice silencer (ESS) motifs which promote an alternative splicing event by which exon 25 is excluded from the resultant transcripts. If DICER1 protein is produced from such transcripts, the RNaseIIIb domain is almost entirely missing. We hypothesize that lack of RNaseIIIb domain would severely affect the variety and abundance of miRNAs in both the Wilms Tumour and multinodular goiter as compared to normal kidney and thyroid respectively to contribute to pathogenesis. These data are consistent with a variant of the classical two-hit model of tumour formation, in that the first hit is likely inactivating whereas the second hit is highly specific.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1263W

Multiplex PCR-based targeted deep sequencing of the comprehensive human lung cancer gene panel for the detection of KRAS and EGFR mutations. *Q. Peng, R. Gardner, N. Slepushkina, V. Devgan.* QIAGEN, Center of Excellence in Biological Research Content, Frederick, MD, USA.

Next-generation sequencing (NGS) technologies are revolutionizing cancer genomics research. The genetic spectrum of mutations in cancers, however, appears to be highly complex, with numerous low frequency somatic variations in a limited number of genes. Therefore, the deep sequencing (more than 1000X coverage) of selected cancer-related genes offers an effective solution for detecting somatic variants in heterogeneous samples and cancer cell subpopulations. A multiplex PCR target enrichment method for human disease-specific genes was developed and integrated with next-generation sequencing to provide an improved technique to identify biologically relevant mutations. In this study, a set of 20 lung cancer-related genes was deep-sequenced using genomic DNA isolated from formalin-fixed paraffin-embedded (FFPE) lung adenocarcinoma samples. The results showed more than 85% coverage for protein coding and UTR regions for all genes, with approximately 95% uniformity (>0.1X of median sequencing depth) and 90% specificity (on-target reads). Sequencing analysis revealed a number of mutations on various cancer-related genes, including previously unknown mutations in KRAS and EGFR. These mutations were confirmed by qPCR somatic mutation and pyrosequencing assays. The results demonstrate that the multiplex PCR target enrichment method provides a highly sensitive and specific solution for comprehensive sequence analysis of a limited number (tens) of potentially clinically relevant cancer genes in FFPE samples. The applications presented here are for research use only. Not for use in diagnostic procedures.

1264T

Highly sensitive detection of rare somatic mutations in tissue and plasma. *A. Marziali[1,2].* 1) Boreal Genomics, Vancouver, BC, Canada; 2) Engineering Physics, University of British Columbia, Vancouver, BC, Canada.

DNA sequence analysis is routinely used for detection of clinically actionable somatic mutations in tumor tissue. Current methods are limited to profiling tissues with mutation abundances above a few percent, particularly when detecting many mutations simultaneously. Since samples may only contain a small fraction of tumor content, and only a fraction of tumor cells may contain relevant mutations, false negatives and ambiguous results are common.

There is also a need for less invasive tumor profiling in diseases such as lung cancer where emergence of mutations such as EGFR T790M result in resistance to targeted therapies and necessitate the need for re-biopsies. Traditional biopsies are invasive and carry risks, preventing their use a monitoring tool. Profiling mutation status by analysis of circulating nucleic acids in plasma would be preferable, but poses the challenge that tumor DNA is present in low abundance in a large amount of normal “wild-type” sequence. Current methods are not sensitive enough to allow such detection reliably, primarily due to the high false positive error rates induced by DNA amplification on the high abundance of wild-type alleles.

We have developed a novel method of detecting up to 100 somatic mutations in tumor samples containing as little as 0.01% mutant DNA. The enabling feature of our method is removal of wild type DNA sequences prior to amplification and detection; thereby reducing false positive rates that result from PCR induced errors on wild type DNA.

Our method is capable of detecting a handful of mutant molecules in a background of millions of wild-type genomes and allows for the detection of low abundance somatic mutations in tissue and plasma samples.

We present validation data in FFPE and plasma samples, demonstrating 0.01% sensitivity. We also report on concordance between mutation incidence in tissue and matched plasma in human samples.

1265F

Integrative investigation on breast cancer by ER, PR and HE2-defined subgroups using mRNA and microRNA expression profiling and cancer core pathway analysis. *X. Dai[1], S. Khan[1], T. Heikkinen[1], P. Heikkilä[4], K. Aittomäki[2], C. Blomgvist[3], D. Greco[1,5], H. Nevanlinna[1].* 1) Department of Obstetrics and Gynecology, University of Helsinki and Helsinki University Central Hospital, Helsinki, Finland; 2) Department of Clinical Genetics, University of Helsinki and Helsinki University Central Hospital, Helsinki, Finland; 3) Department of Oncology, University of Helsinki and Helsinki University Central Hospital, Helsinki, Finland; 4) Department of Pathology, University of Helsinki and Helsinki University Central Hospital, Helsinki, Finland; 5) Department of Bioscience and Nutrition, Karolinska Institute, Stockholm, Sweden.

Breast cancer is a heterogeneous disease and understanding the core functional differences driving this heterogeneity is also of crucial clinical importance in seeking effective breast cancer treatment. Several layers of information, including immunohistochemical markers, mRNA and microRNA expression profiles and pathway analysis have been used for such purpose previously. However, an integrative approach of these is currently missing. We have studied here mRNA and microRNA expression profiles of breast cancer subgroups defined by immunohistochemical expression (IHC) of the major receptors defining breast cancer biology i.e. the hormone receptors ER and PR as well as human epidermal growth factor receptor 2 (HER2). The defined sets of features were validated in two independent data sets at multiple levels, including unsupervised clustering and supervised classification. A core basal group was further defined by epidermal growth factor receptor EGFR and cytokeratin 5/6 IHC expression and compared to triple negative [ER-/PR-]HER2- group. Moreover, the gene expression signatures of the tumor subgroups were screened by in-depth analysis of 12 cancer core pathways. We identified and validated a novel breast cancer gene expression signature composed of 976 mRNAs and 69 microRNAs. [ER+/PR+]HER2- and [ER+/PR+]HER2+ vs. [ER-/PR-]HER2+ tumors are shown to significantly differ both at the mRNA and microRNA levels. HER2 positive tumors are more closely related to triple negative tumors by mRNA profiling than by microRNA expression. Closely related microRNAs sharing the same targets may exert opposite roles in different subgroups. Besides the core cancer pathways, other pathways such as those controlling biomass synthesis are shown to be involved in the core basal group which may enable additional progressive nature compared to the other triple negative tumors. Some therapeutic strategies are proposed for breast cancer treatment, and the pathway analysis also reveals that the clinical strategy and pivotal targets need to be tuned according to different tumor subgroups. This study is the first attempt to elucidate mRNA and microRNA expression and cancer core pathways in breast cancer subgroups defined by ER, PR and HER2 expression. The results can avail functional studies of breast cancer with further translational potential for clinical use.

1266W

Novel Oligoheterocycles and potential ability as anti cancer therapy. *S. Al-Aqeel.* Department of Chemistry, Faculty of Science, King Saud University, Riyadh, PO Box 22452, Riyadh, Saudi Arabia.

There has been interest in heteocyclic systems as sequence-selective DNA minor groove binding agents (and some as RNA binding agents), such as dimeric benzimidazoles related to Hoechst33258 (and trimeric equivalents), and oligopyrroles (and benzimidazole-imidazole chimeras). More recently, some examples of symmetrical dimers or linked dimers, have been reported with similar targets, and at least some evidence of structures with potential for targeting G-quadruplexes. Developing new methods to generate novel bis- and oligobenzimidazoles and evaluation as nucleic acid binding agents is thus of interest. This contribution will describe synthesis of several families of novel oligoheterocyclic systems, designed to provide a modular approach to provide a range of new ligands types, categorized based on heterocyclic orientation, separation and various functionality of linkers incorporated, for example including compounds of the generic types shown. Additionally, other series of dimers, trimers and tetramers will be described. Data on evaluation of binding for an array of structures in the above classes to a series of target DNA sequences (containing various length (AT)n sections) will be presented, leading to identification of novel DNA-binding ligands in several structural families with sub-micromolar affinities and potential ability as anti cancer therapy.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1267T

Relative telomere length differs according to DNA extraction method. *L. Boardman[1], K. Litzelman[2], R. Johnson[3], M. Devine[1], R. Firl[1], A. Johnson[1], M. Vincent[1], J. Cunningham[3], C. Engelman[2], S. Seo[4], R. Gangnon[3, 4], D. Rider[5], G. Petersen[6], S. Thibodeau[3], H. Skinner[2]*. 1) Internal Medicine, Mayo Clinic,College of Medicine Rochester, MN; 2) Population Health Sciences, University of Wisconsin, Madison, WI Laboratory Medicine and Pathology, Mayo Clinic College of Medicine, Rochester, MN; 3) Laboratory Medicine and Pathology, Mayo Clinic College of Medicine, Rochester, MN; 4) Biostatistics and Medical Informatics, University of Wisconsin, Madison, WI; 5) Information Technology, Mayo Clinic College of Medicine, Rochester, MN; 6) Health Sciences Research, Mayo Clinic College of Medicine, Rochester, Mn.

Telomere length has been identified as a marker of biological aging. Both shorter and longer telomeres in peripheral blood leukocyte (PBL) DNA have been associated with diseases of aging, including multiple types of cancer and cardiovascular disease. However, associations have not been consistent across studies — even among studies of the same cancer type. In a study to assess the role of telomere length in colorectal cancer (CRC) patients, we examined PBL relative telomere length (RTL) measured by quantitative PCR (qPCR) in 1096 CRC patients and 1972 healthy controls ascertained through the Colon Cancer family registry (Colon CFR) recruitment locations (Hawaii, Australia, Toronto, Seattle, San Diego and the upper Midwest of the US). The centers used differing DNA extraction methods, including Phenol/Chloroform, PureGene and QiaAmp. We observed significant differences in RTL by DNA extraction method. PureGene extracted DNA had a mean RTL = 6219 base pairs (bp) with a range of 4548 -11050 bp compared to phenol/chloroform extracted DNA which had a mean RTL = 5593 bp with a range = 4184-16662 bp. DNA extracted by QiaAmp had a mean RTL= 4832 and a range = 4186-10544 bp. We subsequently compared RTL measured by qPCR from 25 CRC cases and 25 normal controls in PBL DNA extracted by each of the three DNA extraction methods used by the Colon CFR centers. The range of telomere length measured by qPCR from PBL DNA extracted by the QiaAmp column method was smaller than that which could be detected by either the Puregene or Phenol/Chloroform extraction methods (ranges: 4477-5266 bp, 4230-9259 bp, and 4761-9540 bp, respectively). Within-individual differences in RTL by extraction method supported this finding, with QiaAmp measuring telomeres as significantly shorter than PureGene (mean difference: 749.5 bp; 95% confidence interval (CI): [411.6, 1087.4]) or Phenol/Chloroform (mean difference: 908.3; 95% CI: [574.5, 1242.0]). These significant differences in PBL RTL based on DNA extraction method may contribute to the discrepancies between studies seeking to find an association between cancer or other diseases and RTL.

1268F

Hereditary breast/ovarian cancer in Algerian population: Molecular analysis of *BRCA1* and *BRCA2* genes. *F. Cherbal[1], R. Bakour[1], S. Adane[2], K. Boualga[3], N. Salhi[1], A. Chikh[1], P. Maillet[4]*. 1) Unit of Genetics, LBCM, Faculty of Biological Sciences, USTHB, Algiers, Algeria; 2) Central Hospital of Algiers, Algiers, Algeria; 3) Anti Cancer Center, Blida, Algeria; 4) SwissCheckup Laboratory, Yverdon Les Bains, Switzerland.

Background: Breast cancer is the leading cause of cancer death in women in Algeria. From year 2005 to 2010, we screened *BRCA1* and *BRCA2* genes in 121 individuals from an Algerian cohort with a personal and family history suggestive of genetic predisposition to breast/ovarian cancer. Methods: We analyzed 121 DNA samples from 90 breast/ovarian cancer families. 86 individuals from 70 families have been tested for complete *BRCA1* and *BRCA2* germline mutations screening. The approach used is based on *BRCA1* and *BRCA2* sequence variants screening by DGGE, SSCP or High-Resolution Melting (HRM) curve analysis followed by direct sequencing. Samples for which no pathogenic mutation was found were analyzed by MLPA. *In silico* analyses have been performed using different bioinformatics programs to individualize genetics variations that can disrupt the *BRCA1* and *BRCA2* genes function. Results: Three distinct pathogenic mutations c.83_84delTG, c.181T>G, c.798_799delTT and two new large genomic rearrangements involving deletion of exon 2 and exon 8 respectively, were detected in *BRCA1* gene. Interestingly, the *BRCA1* mutation c.798_799delTT was identified in two unrelated families. Two pathogenic mutations c.1310_1313delAAGA and c.5722_5723delCT were identified in *BRCA2* gene. In addition, 84 UVs and SNPs were detected in *BRCA1/2* genes (33 *BRCA1* and 51 *BRCA2*), 35 were new UVs (10 *BRCA1* and 25 *BRCA2*), 7 were rare UVs (4 *BRCA1* and 3 *BRCA2*) and 42 were SNPs (19 *BRCA1* and 23 *BRCA2*). Moreover, 9 new missense UVs identified in this study: two *BRCA1* (Q1356K and R1634M) and seven *BRCA2* (D367N, S879Y, S1013T, C1290S, V1810F, H2116D and G3086R) show a damaging PSIC score yielded by PolyPhen2 program and could be pathogenic. In addition, 6 new *BRCA2* missense UVs out of seven that were found to be damaging by PolyPhen2 program, also were deleterious according to SIFT program. Several SNPs of *BRCA1* P871L, K1183R, S1613G and *BRCA2* N289H, N372H, N991D, G2384K, have been identified with high frequency in patients who were tested negative for *BRCA1* and *BRCA2* mutations.These missense polymorphisms could have a role as susceptibility breast cancer markers in Algerian breast/ovarian cancer families where pathological *BRCA1* and *BRCA2* mutations were not present. Conclusions: The accumulating knowledge about the prevalence and nature of *BRCA1* and *BRCA2* variants in Algerian population will contribute in the future to the implementation of genetic testing and counseling for Algerian families at risk.

1269W

GENECAPP: Sequence-Specific *in vivo* Analysis of Protein-DNA Interactions in Human Cancer Cell Lines. *H. Guillen Ahlers[1], A. Ludwig-Kubinski[1], S. Tian[3], C. Anderson[1], A.M. Greene[1], J. Kennedy-Darling[2], M. Levenstein[2], R. Knoener[2], M. Chesnik[1], M. Scalf[2], Y. Yuan[2], R. Cole[1], M. Shortreed[2], L. Cirillo[1], R. Stewart[3], L.M. Smith[2], M. Olivier[1]*. 1) Wisconsin CEGS, Medical College of Wisconsin, Milwaukee, WI; 2) Wisconsin CEGS, University of Wisconsin, Madison, WI; 3) Morgridge Institute for Research at the University of Wisconsin, Madison, WI.

DNA-protein interactions are responsible for chromatin stability, activity, and transcription modulation of the encoded genetic information. Currently, no technologies exist to examine these protein-DNA interactions in a comprehensive genome-wide manner. To overcome this challenge, we developed **G**lobal **E**xo**N**uclease-based **E**nrichment of **C**hromatin-**A**ssociated **P**roteins for **P**roteomics (GENECAPP) which utilizes oligonucleotide capture to isolate selected chromatin regions in a sequence-specific manner and analyzes captured DNA-associated proteins using mass spectrometry.

We have used the GENECAPP technology in *S. cerevisiae*, targeting the Gal upstream activator sequence (UAS_{Gal}). In this study, GENCAPP is being used to target genomic regions in breast and ovarian cancer cell lines. Gene expression profiles of drug resistant cell lines (MCF-7_ADR, NCI_ADR) and their drug sensitive counterparts (MCF-7, OVCAR-8) were used to identify highly differentially expressed genes. We developed GENECAPP capture approaches to examine selected promoter regions of these genes to understand the underlying chromatin-proteome interactions that alter gene expression.

Our initial efforts focused on the promoter region of *ABCB1*, a gene implicated in multidrug resistance. Chromatin was extracted from cross-linked cells and fragmented using common restriction enzymes. Total capture efficiencies surpassed 10%, and the target region was enriched 15 fold compared to a negative control. Mass spectral characterization of protein binding before and after drug treatment is currently under way to identify DNA-binding proteins that may mediate the cell line-specific gene expression effect. Our initial data suggest that GENECAPP at the capture level works well in mammalian cells and has the potential to become a powerful tool to identify DNA-binding proteins bound to specific DNA regions of interest *in vivo*.Funded by the Wisconsin CEGS through NIH/NHGRI grant 1P50HG004952.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1270T

Examination of UVR-Induced DNA damage and repair and its association with apoptosis in human keratinocytes and fibroblasts. *M. Karbaschi[1], M.D. Evans[1], S. Macip[2], M.S. Cooke[1]*. 1) Department of Cancer Studies and Molecular Medicine, University of Leicester, Leicester, Leicestershire, United Kingdom; 2) Department of Biochemistry, University of Leicester, Leicester, Leicestershire, United Kingdom.

Skin cancer has an increasing incidence in countries with large populations of white skinned individuals. UVR, by initiating the DNA damage, can lead to mutagenesis and is regarded as the prime cause of most skin cancers. As a result, UVR protection is of primary importance to prevent UVR-induced skin cancers. Cyclobutane pyrimidine dimers (CPD) are an important form of DNA damage induced by both UVA and UVB and removed by nucleotide excision repair. The persistence of CPDs, compared to other forms of DNA damage, is understood to be a major contributory factor to their mutagenicity. Using the T4endonuclease V-modified comet assay on human keratinocytes and fibroblasts, we noted that there was rapid initial repair of CPDs over the first 6h post-irradiation, following either UVA or UVB treatments, but whilst this slowed significantly in the UVB-irradiated cells, it continued to be rapid in the UVA-treated cells with levels approaching baseline within 36h. This confirmed the widely accepted slow repair of UVB-induced cyclobutane thymine dimers, but we uniquely noted far more rapid repair of UVA-induced cyclobutane thymine dimers. There were no significant differences in cell viability between the two treatments over the first 6h post-irradiation, but at 24h post-irradiation viability had decreased significantly only in the UVB-irradiated cells. These data suggest that for at least the first six hours following UVB irradiation, the majority of cells were viable and capable of repair, after that time increasing numbers of cells enter apoptosis, and therefore fail to repair the DNA damage.

1271F

The IRS1 rs2943641 genetic variant and the protection against cancer. *S. Romeo, C. Maglio, J. Andersson Assarsson, K. Sjöholm, L. Sjöström, L.M. Carlsson.* University of Gothenburg, Gothenburg, Sweden.

INTRODUCTION: Obesity and insulin resistance are associated with an increased risk of cancer. Bariatric surgery and drugs that reduce insulin resistance have been shown to decrease cancer risk. The insulin receptor substrate 1 (IRS1) gene encodes the primary mediator of the insulin receptor. A single nucleotide polymorphism (rs2943641) in a non-coding region downstream of the IRS1 gene has been associated with insulin resistance in a genome-wide association study and this result has been widely replicated thereafter. However, the long term effect of this genetic variant on insulin-related comorbidities, such as cancer, remains unknown. The aim of this study was to determine whether the IRS1 rs2943641 T allele results in reduced insulin resistance and protects against cancer in the Swedish obese subjects (SOS) study cohort. METHODS: The SOS study is a prospective, controlled, intervention study involving 4,047 obese individuals; 2,010 individuals have undergone bariatric surgery and 2,037 conventional treatment (matched control group). A total of 2,988 non diabetic obese subjects (body-mass index, BMI=41±5) from the SOS study have been genotyped. Insulin resistance was measured by the homeostasis model assessment for insulin resistance (HOMA-IR). Data on the incidence of solid cancers were available from all individuals (median follow up: 15 years). SUMMARY OF RESULTS: At baseline, HOMA-IR levels were lower in T allele carriers from the SOS study cohort. However, at 2- and 10- year follow up, the association between the IRS1 T allele and HOMA-IR was blunted in surgery group. An interaction was observed between IRS1 genotypes and the surgical treatment on cancer incidence. In the control group, T-allele carriers had a significantly lower cancer incidence while there was no effect of IRS1 genotype on cancer incidence in the surgery group. CONCLUSION: the IRS1 rs2943641 T allele associates with protection against insulin resistance and cancer development in severely obese individuals. However, these associations are lost after bariatric surgery, suggesting that body weight modulates the effect of IRS1 genotype.

1272W

Lynch syndrome: Awareness among medical students at a United States medical school. *J. S. Taylor[1], M. K. Frey[1], M. Biewald[1], M. Worley, Jr.[2], S. Lin[1], K. Holcomb[1]*. 1) Obstetrics and Gynecology, New York Presbyterian Hospital Weill Cornell Medical Center, New York, NY; 2) Department of Gynecologic Oncology, Brigham and Women's Hospital, Boston, MA.

Objectives: Lynch syndrome was first described in the 1950s however until recently it was rarely included in medical school curricula. As a result, many practicing physicians have limited exposure, potentially contributing to significant under diagnosis. As identification of Lynch syndrome prior to malignancy allows for intensified screening, prophylactic surgery and improved patient outcomes, all physicians should be aware of the characteristics of affected families. We aim to determine the overall level of awareness of Lynch Syndrome among medical students at an American medical school Methods: A voluntary and anonymous questionnaire was delivered to students at an American medical school. The survey instrument assessed the respondent's perceived knowledge regarding the genetics and recommended screening for carriers of Lynch Syndrome mutations. Results: The questionnaire was distributed to the entire student body (405 students) with a response rate of 50%. Fifty-nine percent of students reported that they had learned about Lynch syndrome; 27% of first year students, 44% of second year students; 90% of third year students and 100% of fourth year students. Of the students familiar with Lynch syndrome, the reported knowledge of the underlying genetics was 46%, available genetic screening, 18%, criteria used to screen for the syndrome, 24%, recommendations for colon screening, 31% and recommendations for endometrial cancer screening, 17%. Conclusions: The majority of medical students surveyed had been exposed to Lynch syndrome and awareness increased over each year of education. Significantly more students were aware of recommendations for colon cancer screening than endometrial cancer screening (32% versus 17%, p = 0.01). Studies of the natural history of Lynch syndrome indicate that affected women are more likely to present with endometrial cancer than colon cancer and while there are no prospective data proving the efficacy of endometrial cancer screening in this high-risk population, the endometrium is easily accessible and can be sampled using simple office techniques. In addition, prophylactic hysterectomy and bilateral salpingo-oophorectomy are reasonable risk reducing interventions for the prevention of both uterine and ovarian cancer. Our findings suggest that increased emphasis must be placed on teaching the gynecologic manifestations of Lynch Syndrome in order to avoid the misconception that it is simply a colon cancer syndrome.

1273T

Characterization of active chromatin signatures in testicular germ cell tumour cell lines. *S.J. White[1], Y. van de Zwan[1,2], C.M. de Boer[1], F. Rossello[1], L.H.J. Looijenga[2], A.J. Notini[1]*. 1) Centre for Reproduction and Development, Monash Institute of Medical Research, Clayton, Victoria, Australia; 2) Erasmus MC, University Medical Center Rotterdam, Josephine Nefkens Institute, The Netherlands.

A fundamental goal of modern biology is understanding how genes are regulated both temporally and spatially. This control is not only dependent on DNA sequence; epigenetic factors such as CpG methylation and histone modifications also play a major role. A range of histone modifications are associated with different characteristics of transcriptional regulation, and techniques such as chromatin immunoprecipitation (ChIP) have started to reveal the complexity underlying this regulation. We have applied ChIP-seq to explore the epigenetic differences between the two major subtypes of testicular germ cell tumours (TGCTs), namely seminoma (SE) and non-seminoma/embryonal carcinoma (NS). SE consist of embryonic germ cells that are not overtly pluripotent, whereas NS are cancer cells that have reactivated pluripotency. Our initial studies focused on cell lines, with TCam-2 representative of SE and NCCIT of NS, especially the stem cell component; embryonal carcinoma (EC). Using antibodies specific for histone modifications such as H3K4me1, H3K4me3 and H3K27ac, we were able to identify active regions within the genome, including promoters and enhancers. Initial analysis matched the classification of the cell lines e.g. SOX17, a marker for SE, was strongly enriched for these modifications in TCam-2 compared to NCCIT, whereas the opposite was seen for SOX2, an EC marker. We are currently confirming the most significant differences between the two cell lines, and plan to expand this study into primary TGCTs. These findings will identify epigenetic differences between SE and NS, and will improve our understanding of pluripotency.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1274F

Single Nucleotide Polymorphisms (SNPs) and cancer risk in individuals with Costello Syndrome. *B.A. Thompson[1], G. Desachy[2], J. Quinn[2], A.E. Toland[3], L.A. Weiss[2], K.A. Rauen[1].* 1) Pediatrics, University of California San Francisco, San Francisco, CA; 2) Psychiatry, University of California San Francisco, San Francisco, CA; 3) Molecular Virology, Immunology and Medical Genetics, Ohio State University, Columbus, OH.

Background. The Ras/mitogen-activated protein kinase (MAPK) signal transduction pathway plays an important role in cell proliferation, embryonic development and tumorigenesis. Abnormal germline regulation of this pathway results in overlapping yet very distinct syndromes known as RASopathies, which are a class of human genetic syndromes caused by germline mutations in genes encoding components of the Ras/MAPK pathway. One feature observed in many RASopathies is an increased incidence of cancer. The type of cancer varies between each syndrome. Costello syndrome (CS), caused by germline mutations in *HRAS*, is one RASopathy whereby ~17% of individuals develop cancer including rhabdomyosarcoma, neuroblastoma and bladder cancer. It is unclear why some individuals with CS develop cancer and others do not. In population studies, SNPs in key oncogenes and tumor suppressor genes (TSGs) have been reported to be associated with an increased cancer risk. However, the association between these SNPs and cancer risk in CS individuals has not previously been described. **Methods.** Demographic information including ethnicity, sex and history of cancer was collected from 52 CS individuals. DNA was extracted from blood, fibroblasts or buccal swabs obtained from each patient. DNA was also extracted from a tumor obtained from one of the patients that was previously reported to exhibit loss of heterozygosity (LOH) for the wt *HRAS* allele. CS diagnoses were confirmed by sequencing *HRAS*. Eight SNPs previously reported to be associated with increased cancer risk were sequenced. These SNPs are located in the following genes: *MYC, AURKA*, *TP53*, *KRAS, MDM2* and *CDKN1A*. A statistical analysis of the allele frequencies was performed using a classic Fisher's exact test comparing cases and controls. **Results & Conclusion.** Of the 52 individuals with CS, 8 had a history of cancer (15%). Statistical analysis of the SNP allele frequencies of the cases (CS with cancer) vs controls (CS without cancer) yielded p-values >0.05 for all 8 SNPs. Sequencing of the tumor sample confirmed the LOH and presence of two copies of the HRAS G12S mutation, but did not show LOH at any of the 8 SNPs sequenced in this study. Although preliminary analysis of the data obtained from this study did not find evidence for association between these SNPs and cancer risk in this population, further analysis of other SNPs in TSGs or oncogenes may reveal an association impacting cancer risk assessment and future patient care.

1275W

Assessment of Individuals with BRCA1 and BRCA2 Large Genomic Rearrangements in High-Risk Breast Cancer and Ovarian Cancer Families. *L. Zhang[1], A. Arnold[2], M. Harlan[2], M. Robson[2].* 1) Molecular Diagnostics Service, Department of Pathology, Memorial Sloan-Kettering Cancer Center, New York, NY; 2) Clinical Genetics Service, Department of Medicine, Memorial Sloan-Kettering Cancer Center, New York, NY.

Background Standard BRCA1/2 genetic testing (BRACAnalysis) has consisted of full gene sequencing and a BRCA1 gene 5-site rearrangement panel (Myriad Genetics Laboratories). Detection of other large rearrangements (LRs) in both BRCA1 and BRCA2 has been available since 2006. Although it has been reported that LRs comprise a significant component of identifiable mutations in BRCA1 and BRCA2, the prevalence of these mutations has not been fully established. Testing for LRs in BRCA1/2 is often not covered by many insurance companies. Methods Our cohort consisted of 3301 probands seen for genetic counseling and BRCA1/2 genetic testing at Memorial Sloan-Kettering Cancer Center from September 2006 to September 2011. A proportion of individuals (1279) without deleterious point mutations were screened for LRs in BRCA1/2. The cancer family history and ethnicity of patients having LRs were reviewed. The pre-test probability each proband had for carrying a BRCA1/2 gene mutation was retrospectively calculated by applying the BRCAPRO, Penn II and Myriad II models. Patients were excluded from this review if they presented for testing for a known familial BRCA1/2 LRs. Results 150 probands carry a deleterious BRCA1/2 mutation (probands with Ashkenazi Jewish founder mutations were excluded). 19 patients were found to have LRs. 2 probands had LRs detectable by standard BRCA1/2 testing. 17 probands (17/150=11.3%) had LRs not detectable by standard BRCA1/2 testing. 7/19 (36.8%) patients with LRs had a less than a 10% pre-test probability of carrying a BRCA1/2 mutation using BRCAPRO, which outperformed the other two models. 13/19 (68.4%) of patients were diagnosed with breast cancer at an average age of 36.4 years. 5/19 of patients (26.3%) had a diagnosis of high-grade serous ovarian/fallopian tube cancer with an average age of 46.6 years. One patient was unaffected. 13/19 (68.4%) of patients were from Europe. 4/19 (21%) LR positive patients carried a deletion of exon 20 in BRCA1. One patient who is 1/8 non-Jewish carried a LR mutation (BRCA1*del ex20). Conclusions The sensitivity of risk assessment models of a priori risk of harboring a BRCA1/2 large rearrangement mutation is relatively low. Our data support the recently published NCCN guidelines that LRs should be part of BRCA1/2 testing for all patients that meet testing criteria for the BRCA1 and BRCA2 genes.

1276T

Genetic background of familial colorectal cancer type X. *TT. Nieminen[1], J-P. Mecklin[2], HJ. Järvinen[3], P. Peltomäki[1].* 1) Med Gen, Univ Helsinki, Helsinki, Finland; 2) Deparment of Surgery, Jyväskylä Central Hospital, Jyväskylä, Finland; 3) Second Dapartment of Surgery, Helsinki University Central Hospital, Helsinki, Finland.

BACKGROUND Hereditary non-polyposis colorectal cancer (HNPCC) consists of two separate categories, namely Lynch syndrome (LS), in which mismatch repair (MMR) gene mutations are present, and familial colorectal cancer type X (FCCX) which clinically resembles LS, but MMR gene mutations are absent. FCCX patients have an increased risk of colorectal cancer (CRC), but extracolonic cancers are relatively uncommon. The average age at onset is around 60 years. Despite intensive research efforts, the predisposing genes for FCCX have remained unknown. AIM, MATERIALS AND METHODS This study aims to dissect the genetic background of FCCX taking advantage of 19 FCCX families from a nation-wide registry. The age at onset for CRC varies between 36 and 72 years in these families. Predisposing genes are sought for by genetic linkage analysis, exome sequencing and Sanger sequencing of candidate genes. We have also attempted to explore the genetic background of FCCX by tumor investigations. RESULTS In 2 out of 19 FCCX families, the BMPR1A gene revealed novel mutations which affected the extracellular part of the BMPR1A protein. The mutation in family 20 was an in-frame deletion of three nucleotides (AGA) at position 264 (c.264–266del) and in family 68 a 24-bp deletion at a splice site involving the 3'portion of intron 1 and the 5'portion of exon 2 (c.68-10_68+14del). Family members at-risk were tested for the presence vs. absence of the mutation in question and there were 24/48 (50%) mutation positive members in family 20 and 2/9 (22%) mutation positive members in family 68. All the other FCCX families were negative for BMPR1A mutation. BMPR1A has previously been connected to juvenile polyposis syndrome (JPS), but in our FCCX patients none of the adenomas or polyps showed JPS- like findings. The remaining families have been subjected to exome sequencing and analysis of the results is in progress. CONCLUSION Our finding is a landmark observation considering the role of BMPR1A mutations in FCCX. Now it will be important to investigate whether FCCX families from other parts of the world have BMPR1A mutations as well and if so, to generate more information of mutation types and locations and the associated clinical phenotypes. Our results also suggest that FCCX families are genetically heterogeneous and additional susceptibility genes are likely to be found by available high-throughput methods.

1277F

Development of a robust method for establishing B cell lines using Epstein-Barr Virus. *I. Danjoh, R. Shirota, Y. Nakamura.* Cell Engineering Division, RIKEN BioResouce Center, Tsukuba, Ibaraki, Japan.

B lymphoblastoid cell lines (B-LCLs) are established from B lymphocytes by infection with Epstein-Barr virus (EBV). As their genomic structure is stable in culture, B-LCLs are a valuable resource for genome analysis. Although the efficiency of establishing B-LCLs from freshly obtained blood samples from healthy individuals is almost 100%, if blood samples were handled inappropriately after collection or peripheral blood mononuclear cells (PBMCs) were stored in liquid nitrogen for a long time, the efficiency gets considerably lower. We have established more than 550 B-LCLs from PBMCs that have been collected from indigenous peoples living in closed societies in South America and stored in liquid nitrogen for over 20 years (Detailed description of this B-LCL collection was reported in the reference below*). As some of these populations no longer exist, the collection is a unique repository of the human genome. B-LCL establishment from such samples are difficult and failure means the waste of valuable and rare samples.

Although the methodology for establishing B-LCLs was developed at the end of 1970s, surprisingly little attention has been given over the years to identifying the factors that determine the success rate of generating cell lines. It was therefore essential that the methodology used for establishing B-LCLs from these was improved in order to create a resource that would not be rapidly depleted. For this purpose, we sought to improve our success rate for establishing B-LCLs from the difficult and irreplaceable samples by a detailed examination of each step of the process.

The analysis showed that two parameters were particularly critical: the density of the PBMCs plated after EBV infection, and the EBV titer. These observations shed light on cases where establishment of B-LCLs was hard due to the small number and/or damaged PBMCs.* Danjoh et al.; (2011) The Sonoda-Tajima Cell Collection, a human genetics research resource with emphasis on South American indigenous populations. Genome Biol. Evol. 3: 272-283.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1278W

Association of single nucleotide polymorphisms in ERα, ERβ, CYP17A1 and CYP19A1 with breast cancer risk: A case control study from North India. *S. Chattopadhyay[1], SVS. Deo[2], NK. Shukla[2], SA. Husain[1].* 1) Human Genetics Lab, Department of Biotechnology, Jamia Millia Islamia, New Delhi, Delhi, IndiaHuman Genetics Lab, Department of Biotechnology, Jamia Millia Islamia, New Delhi 110025 Delhi, India; 2) DR. BRA- IRCH, All India Institute of Medical Sciences, New Delhi 110029 INDIA.

A complex microenvironment exists in breast cancer that affects the dynamic signaling in cancer cells. Estrogen is important in promoting the proliferation of both normal and neoplastic breast epithelium. Biosynthesis and metabolism of estrogen play a pivotal role in initiation, promotion and progression of breast cancer. Estrogen acts by binding to one of the two specific estrogen receptors (ERs), ERα (ESR1) and ERβ (ESR2) (both belonging to nuclear receptor superfamily). CYP17A1 and CYP19A1 (also known as aromatase), are the key genes involved in biosynthesis of estrogen and thus to its exposure. Single nucleotide polymorphisms (SNPs) in these genes are possible risk factors implicated in breast cancer. The present study explores the role of estrogen biosynthesis pathway related gene polymorphisms in susceptibility to breast cancer in North Indian population. Estrogen receptors, ESR1 (rs2234693) and ESR2 (rs2987983); CYP17A1 (rs743572) and CYP19A1 (rs700519) polymorphisms were selected for the present study. Genomic DNA was isolated from 290 sporadic breast cancer patients and 290 sex matched healthy controls according to the standard phenol/chloroform method. Genotyping of rs2234693 was done by Polymerase Chain Reaction-Restriction Fragment Length Polymorphism (PCR-RFLP); genotyping of rs2987983 by tetra-primer amplification refractory mutation system (ARMS) PCR and of rs743572 and rs700519 by direct sequencing. Hardy-Weinberg equilibrium testing of all SNPs ($\chi 2$ test, $P>0.05$) was consistent with the distribution. 3×2 $\chi 2$ test was used to compare the overall distribution of genotypes between patients and controls, achieving significance for all SNPs ($P<0.05$). ESR1 Intron 1 (rs2234693; TT vs TC+CC, P=0.005); ESR2 Promoter (rs2987983; CC vs CT+TT, P= 0.002); CYP17A1 5' UTR (rs743572; CC vs CT+TT, P=0.009) and CYP19A1 Exon7 (rs700519; TT vs TC+CC, P=0.016) were all significantly associated with breast cancer after correction with age. Our study signifies the role of these SNPs in affecting the risk of breast cancer. The enhanced risks imparted by these variants suggest the potential role of these low penetrance genes in the development and progression of breast cancer. This study provides an insight into the role of estrogen in hormone dependent breast cancer. Acknowledgment:Council of Scientific and Industrial Research, Govt. of India for providing fellowship to SC and University Grants Commission, Govt. of India for providing research fund to SAH.

1279T

Efficacy of Sequenom Sample ID Plus® SNP genotyping in identification of FFPE tumour samples. *J.K. Miller, N. Buchner, D. Pasternack, J.D. McPherson.* Ontario Institute for Cancer Research, MaRS Centre South Tower, Toronto, Ontario, Canada.

Misidentification of cell line and tissue DNA intended for next-generation sequencing (NGS) can be costly if not detected early. For this reason, it is necessary to have quality control measures in place for source DNA identification and authentication of matching tumor and normal pairs at the beginning of sequencing pipelines. Short tandem repeat (STR) analysis, such as AmpFISTR Identifiler®, is a standard, PCR-based human genotyping method used in the field of forensics. This genotyping method has also been effectively applied as an early quality control step for NGS to prevent processing of misidentified samples. A major issue in large-scale cancer research studies involving archival formalin fixed paraffin embedded (FFPE) tissues is that varying levels of DNA degradation can result in failure to correctly identify samples using STR genotyping. PCR amplification of STRs of several hundred base pairs is not always possible when DNA is degraded. The Sample ID Plus® panel from Sequenom allows for human DNA identification and authentication using SNP genotyping. In comparison to lengthy STR amplicons, this multiplexing PCR assay requires amplification of only 80–120 base pairs, and utilizes 47 SNPs to discriminate between individual samples. In this study, we evaluated both STR and SNP genotyping methods of sample identification, with a focus on paired FFPE tumour/normal DNA samples. The ability to successfully validate the identity of degraded samples can enable cost savings by reducing NGS rework.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1280W

Gene-specific accelerated transcriptional aging in Type 2 diabetes. *J. Kent[1], M. Almeida[1], J. Peralta[1], H. Goring[1], J. Curran[1], M. Johnson[1], T. Dyer[1], S. Cole[1], J. Jowett[2], A. Comuzzie[1], M. Mahaney[1], L. Almasy[1], J. MacCluer[1], E. Moses[1], R. Duggirala[1], J. Blangero[1], S. Williams-Blangero[1].* 1) Dept Gen, Tx Biomed Res Inst, San Antonio, TX; 2) Baker IDI Heart and Diabetes Inst, Melbourne, Australia.

Patients with long-duration Type 2 diabetes mellitus (T2DM) exhibit some of the physiological changes associated with normal biological aging, including increased risk for atherosclerotic cardiovascular disease and osteoarthritis. Using transcriptional profiles of primary peripheral blood mononuclear cells from 1218 Mexican American participants in the San Antonio Family Heart Study for whom age and T2DM status were available, we identified 4,034 of 22,413 detectable transcripts whose expression levels were correlated with age at a 5% false discovery rate (FDR). We then tested whether the age-related change in gene expression was the same in individuals with and without T2DM, using a mixed linear regression model, implemented in SOLAR, that included the random effect of kinship. Of 402 age-correlated transcripts that were also nominally ($p<0.05$) correlated with DM, 153 (38%) also showed at least nominal evidence of age by diabetes interaction. Thus, for specific genes, individuals with diabetes showed an "older" transcriptional profile than unaffecteds, even accounting for the overall change in gene expression with age. The set of differentially expressed genes was enriched for mitochondrial function, lipid transport, and immune response. Cohort characteristics - combined: 59.1% female, mean(SEM) age=39.3(0.48)y, median age(range)=37.6(15–92)y; with T2DM: N=188, 37.2% female, mean(SEM) age=54.9(1.05)y, median age(range)=52.5(20–92)y; without T2DM: N=1,030, 58.4% female, mean(SEM) age=36.4(0.48)y, median age(range)=33.9(15–88)y.

1281T

Integrating Metabolomic Information in Genome-Wide Association Studies. *R. Rueedi[1,6], M. Ledda[2], T. Corre[1,6], R. Salek[3], V. Mooser[4], P. Vollenweider[5], G. Waeber[5], U.K. Genick[2], Z. Kutalik[1,6], S. Bergmann[1,6].* 1) Department of Medical Genetics, Université de Lausanne, Lausanne, Switzerland; 2) Department of Food-Consumer Interaction, Nestlé Research Center, Lausanne, Switzerland; 3) European Bioinformatics Institute, Cambridge, United Kingdom; 4) Department of Pathology and Laboratory Medicine, Centre Hospitalier Universitaire Vaudois, Lausanne, Switzerland; 5) Department of Internal Medicine, Centre Hospitalier Universitaire Vaudois, Lausanne, Switzerland; 6) Swiss Institute of Bioinformatics, Lausanne, Switzerland.

Genome-wide association studies (GWAS) have played an instrumental role in discovering robust genetic associations with dozens of clinically relevant phenotypes. The discovered genetic determinants of complex traits, however, explain a disappointingly small fraction of the heritability estimated by twin studies. Clearly, a substantial part of the genetic architecture underlying diseases is propagated through the metabolome. The aim of our study is to establish these cascades of causal links from genotypes to phenotypes via metabotypes. Close to 1'000 participants from the Cohorte Lausannoise (CoLaus) study provided urine samples for unspiked urine NMR assessment. We found 37 significantly associated SNP to metabotype feature (i.e. NMR shift) pairs. Several associations involve SNPs which have been shown to associate to metabolites in previous, usually serum-based, GWA studies. Moreover, the majority of the genotype-metabotype associations have been confirmed by an equal-sized replication study. Interestingly, many of these features are in turn associated to phenotypes, such as blood-lipids, obesity, and lifestyle factors, suggesting the role metabolites may have as intermediate phenotypes in genotype to phenotype GWA studies. The associated NMR shifts require annotation to reveal the metabolites which produce them. To simplify the identification of the metabolites involved in the SNP to feature associations, we developed a statistical lookup procedure which ranks metabolites according to the coherence between their known spectrum and their association spectrum. While this lookup procedure cannot uniquely identify the metabolite involved in the association, it produces a list of candidates, from which biologically sound metabolites can possibly be inferred with the help of metabolomics experts. After this initial study, we will now focus on the use of metabotypes as intermediate phenotypes in GWA studies, and how such an intermediate phenotype can increase the statistical power to detect genotype-phenotype association. If valid, such an approach would be a far more cost-effective method of increasing GWA output than resorting to ever larger cohorts, and would make metabotyping a very attractive component of future genotyping projects.

1282F

Application of principal components analysis to the investigation of expression profile data. *H.H.H. Göring[1], A.R. Sanders[2], E.I. Drigalenko[1], W. Moy[2], J. Duan[2], J.E. Curran[1], M.P. Johnson[1], E.K. Moses[3], J. Blangero[1], P.V. Gejman[2].* 1) Dept. of Genetics, Texas Biomedical Research Institute, San Antonio, TX, USA; 2) Dept. of Psychiatry and Behavioral Sciences, NorthShore University HealthSystem Research Institute, and University of Chicago, Evanston, IL, USA; 3) Centre for Genetic Epidemiology and Biostatistics, University of Western Australia, Perth, Western Australia, Australia.

Gene expression profiling is becoming a widely used approach in human genetic research. We have performed a detailed investigation of the utility of principal components analysis (PCA) for analyzing transcriptional profile data. Microarray-based gene expression profile data were available for 885 lymphoblastoid cell lines (LCLs) from the Molecular Genetics of Schizophrenia Study and 1244 lymphocyte samples from the San Antonio Family Heart Study. Gene expression profile data was found to have substantial underlying structure, with the top 1, 2, 5, and 10 principal components (PC) explaining ~9, 17, 25m and 35 of transcriptomic variation in LCLs, respectively, with similar proportions in lymphocytes. We investigated the relationship of expression PCs to putative confounder variables in gene expression profile studies, including sex and age as well as various LCL characteristics (including transformation site, cell growth rate, energy status, and viral copy load). By regressing out PCs, the influence of these variables on gene expression levels can be largely eliminated, reducing both the risk to identify false positive findings and the power to detect true expression signatures characteristic of a condition of interest. We developed an empirical approach to decide how many PCs should be regressed out to maximize power to detect expression quantitative trait nucleotides (eQTNs), while maintaining proper false positive rates. Regressing out PCs accounting for 50% of gene expression variation is nearly optimal for detection of putatively *cis*-acting eQTNs, nearly doubling the number of significant associations. While regressing out expression PCs lowers residual heritability (which presumably comprises trans-regulatory variation), it nonetheless increases power to detect significant *trans*-acting eQTNs, and enables detection of several weak, putative master regulators. Our investigation shows that PCA has great utility and wide applicability for investigations of transcriptional profile data.

1283W

Prioritization of SNPs identified in microRNAs and their targeted genes by integrated analysis of high throughput datasets. *X. Chen[1], H. Zhao[1,2,3].* 1) Program in Computational Biology and Bioinformatics, Yale University, New Haven, CT, USA; 2) Department of Genetics, Yale University School of Medicine, New Haven, CT, USA; 3) Department of Biostatistics, Yale School of Public Health, New Haven, CT, USA.

Tens of millions of genetic variants have been identified from the sequencing of the human genomes, and many of these variants may play a role in disease etiology of various common diseases. Great efforts have been made to annotate these variants for their functional relevance. These annotations may provide useful information to prioritize candidates from genome wide association studies for replications and/or functional studies. A number of annotation tools, such as SIFT, Polyphen and ANNOVAR, have been developed, but they mostly focus on protein coding regions, since finding the relationship of diseases and non-coding regions are usually harder due to lack of enough knowledge. However, recent studies have shown that alterations in microRNAs (miRNAs) and their binding sites (3'UTRs) may play a crucial role in affecting disease susceptibility. Therefore, accurate annotations of variants in these regions are important. In this presentation, we describe our approach to annotating variants in miRNAs and their target regions through an integrated analysis of different types of high throughput datasets. A hierarchical Bayesian model is proposed to rank the variants (either novel or known) by their potential relevance to the disease development with various prior information in these non-coding regions. One of the most important priors is the miRNA-gene interaction network inferred from paired expression profilings of different diseases and then weighted by the disease similarities, which enables us to find the disease-specific interaction and meanwhile incorporate the information over all other types of diseases. Other priors include cross-species conservation scores, sequence context for miRNA targeting, minor allele frequency (especially for novel or rare SNPs) and other features available for miRNA and targeted regions. Our methods and results may facilitate the interpretation and utilization of genetic variants identified through next generation sequencing technologies, especially those variants in the non-coding regions to the human genome.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1284T

Cis- and _Trans_-eQTL analysis in 5,311 unrelated peripheral blood samples identifies novel disease pathways and helps to pinpoint causal variants. *H. Westra[1], T. Esko[2], M.J. Peters[3,12], C. Schurmann[4], H. Yaghootkar[5], J. Kettunen[6], M.W. Christiansen[7], R.S.N. Fehrmann[1], G.J. te Meerman[1], A. Hofman[8,12], F. Rivadeneira[3,8,12], E. Reinmaa[2], R.C. Jansen[9], J. Brody[7], S.A. Gharib[10], A. Suchy-Dicey[7], D. Enquobahrie[11], A.G. Uitterlinden[3,8,12], C. Wijmenga[1], B.M. Psaty[7], S. Ripatti[6,13], T. Frayling[5], A. Teumer[4], A. Metsepalu[2], J.B.J. van Meurs[3,12], L. Franke[1].* 1) Department of Genetics, University of Groningen, University Medical Centre Groningen, Groningen, The Netherlands; 2) Estonian Genome Center and Institute of Molecular and Cell Biology of University of Tartu, Tartu, Estonia; 3) Department of Internal Medicine, Erasmus Medical Centre, Rotterdam, The Netherlands; 4) Interfaculty Institute for Genetics and Functional Genomics, Ernst-Moritz-Arndt University Greifswald, Greifswald, Germany; 5) Genetics of Complex Traits, Peninsula College of Medicine and Dentistry, University of Exeter, Exeter, United Kingdom; 6) Institute for Molecular Medicine Finland FIMM, FI-00014 University of Helsinki, Helsinki, Finland; 7) Cardiovascular Health Research Unit, Departments of Medicine and Epidemiology, University of Washington, Seattle, WA, United States; 8) Department of Epidemiology, Erasmus Medical centre, Rotterdam, Netherlands; 9) Groningen Bioinformatics Center, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Groningen, Netherlands; 10) Center for Lung Biology Department of Medicine University of Washington, Seattle, WA, United States; 11) Department of Epidemiology, University of Washington, Seattle, WA, United States; 12) Netherlands Genomics Initiative-Sponsored by the Netherlands Consortium for Healthy Aging, Rotterdam, Netherlands; 13) Wellcome Trust Sanger Institute, Hinxton, United Kingdom.

Genome-wide association studies have so-far revealed over 4,500 SNPs that are associated with complex traits. However, due to linkage disequilibrium, it is often hard to identify the causal variant and gene, making it difficult to elucidate downstream molecular consequences. To resolve this, we performed expression quantitative trait locus (eQTL) mapping on 34,000 Illumina-based gene expression probes to identify *cis*-eQTLs (SNP and mid-probe distance < 250kb) and *trans*-eQTLs (SNP and mid-probe distance > 5Mb) in 5,311 peripheral blood samples across seven independent cohorts.

We confined our *trans*-eQTL analysis to 4,500 SNPs, associated with complex traits and detected over 350 SNPs with a *trans*-eQTL effect (FDR ≤ 0.05), reflecting 500 probes and 431 genes. 250 SNPs had both a *cis* and a *trans*-eQTL effect. In the genome-wide *cis*-eQTL analysis, we identified *cis*-effects for 15,376 unique SNPs, reflecting 16,076 probes and 10,129 genes. We observed that 38% of the 4,500 trait-associated SNPs have *cis*-effects.

Since the top 1000 *cis*-eQTLs showed a strong signal of association (having a P-value < 10^{-100}), we hypothesized that these *cis*-eQTLs may be enriched in causal variants. By comparing the top *cis*-eQTL SNPs with small-effect *cis*-eQTL SNPs, we observed a highly significant enrichment of SNPs that map to splice sites, introns, UTRs, miRNA binding sites and transcription factor binding sites. We furthermore observed that fewer than 6% of the top *cis*-eQTL SNPs were coding while 3% were non-synonymous, indicating that genetic gene expression regulation in peripheral blood is predominantly due to non-coding variation. Even though we studied peripheral blood, significant and biologically plausible *trans*-eQTLs were detected for SNPs that are associated with non-hematological traits. For example, rs174550 in *FADS1*, a SNP known to be associated with lipid and glucose levels, showed significant association in *trans* with *LDLR* gene expression levels (P-value = $4.27*10^{-14}$). We subsequently observed a significant correlation between *LDLR* gene expression and lipid levels.

Our results indicate that the analysis of *trans*-eQTLs in peripheral blood is applicable to identify downstream effects of trait-associated SNPs and that *cis*-eQTL analyses may aid in the identification of causal variants.

1285F

Ridge Regression for Genetic Data: A semi-automatic method to guide the choice of ridge parameter. *E. Cule[1], M. De Iorio[2].* 1) Imperial College London, London, United Kingdom; 2) Department of Statistical Science, University College London, London, United Kingdom.

We address the challenge of predicting disease risk using data obtained in a genome-wide association study (GWAS). Because genotypes are fixed at birth, but disease symptoms may not present until later in life, risk prediction using genetic data could suggest lifestyle or pharmaceutical interventions in higher risk individuals, which may reduce future disease susceptibility. In a GWAS, data are collected at hundreds of thousands or millions of genetic variants in hundreds or thousands of individuals. These data are typically analyzed one at a time (SNP-by-SNP). However, such an approach does not take into account the possible combined effect of multiple variants, a plausible hypothesis for complex diseases. Analyzing all SNPs simultaneously presents statistical challenges due to the high dimensions of the data and correlation among nearby SNPs. Traditional multiple regression fails under these circumstances. Current studies that incorporate genetic factors into risk prediction models use a small number of genetic variants most strongly associated with disease risk. However, the resultant risk prediction models offer little or no improved predictive ability compared to risk prediction models using clinical risk factors alone. Here we develop a method to fit a prediction model that incorporates all the variants in a GWAS simultaneously, with the aim of good predictive performance for complex diseases. Ridge regression (RR) is a penalized regression method that gives rise to stable coefficient estimates even when the number of predictors exceeds the number of observations. One challenge in the use of RR is the choice of the ridge parameter, which affects the model fit. A number of methods have been proposed in the literature to choose this parameter based on the data, but no consensus method provides a universally optimum choice. In this study we investigate a means of choosing the ridge parameter such that the degrees of freedom for variance is the same as that of a principal components regression with a specified number of principal components (PCs). We discuss ways to choose the number of PCs to use, and demonstrate that when the number of causal variables is large and effect sizes are small our method offers improved predictive performance over other regression methods. We apply our method to out-of-sample prediction using two Bipolar Disorder GWAS. We present an R package, ridge, which implements the method for genome-wide SNP data.

1286W

Defining the Functional Significance of Genes Using Natural Human Knockouts. *J.D. Hoffman, K.O. Polzin, A.E. Fish, P. Mayo, N.C. Schnetz-Boutad, J.L. Haines.* Center for Human Genetics Research, Vanderbilt University, Nashville, TN.

BioVU is Vanderbilt University's biorepository that was developed with the goal of linking DNA to de-identified data from electronic medical records to allow assessment of how genetic variation impacts clinical phenotypes. BioVU currently contains DNA samples on over 140,000 individuals and de-identified clinical data from over 2,000,000 records. Samples and data are available to Vanderbilt researchers with a primary requirement that the genetic data be redeposited into the BioVU genotyping database. Using the wealth of clinical and genetic data stored in BioVU we studied the role of whole gene-knockouts on the burden of disease within this population. Our study included 2,856 Caucasian subjects that were previously genotyped on the Illumina 660w chip and re-deposited in the BioVU database. Using PennCNV as our algorithm for assessing gene copy number, we examined homozygous gene deletions and their association with ICD-9 clinical diagnosis codes. To be categorized as a gene-knockout, the copy number variation (CNV) called by PennCNV needed to fully overlap any gene in that region and be called as a zero copy. Of the original 2,856 samples, 2,387 passed quality filtering and were included in downstream analysis. 1,704 of these subjects had at least one homozygous gene deletion, with the late cornified envelope 3B (LCE3B) gene being the most commonly homozygous deleted gene (frequency = 0.40). LCE3B has previously been confirmed to be a common deletion in the Caucasian population. Association analysis was carried out for 26 genes (with homozygous deletion frequencies ≥ 0.006) and 201 ICD-9 codes (with frequencies ≥ 0.05). Genes were tested against ICD-9 codes with age at time of diagnosis as a covariate in our regression analysis. Our most significant findings included the association of the human major histocompatibility complex HLA-DRB5 with the ICD-9 code for rheumatoid arthritis and other inflammatory polyarthropathies (p=0.0026, OR= 3.98 [1.61–9.79]). This locus also was found to be significant with the ICD-9 code representing type 2 diabetes (p=0.0045, OR=3.06 [1.41–6.61]. Although refinement of analysis methods and more sensitive grouping of ICD-9 codes will be necessary, our preliminary analysis suggests that there are common clinical implications for fully deleted genes.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1287T

Intronic single nucleotide polymorphism of CALM-1 gene is significantly associated with osteoarthritis knee. *S. Raj, R. Srivastava, D. Sanghi, S. Awasthi, A. Mishra.* Orthopaedic Surgery, KG Medical College, CSM Medical University, Lucknow, UP, India.

Objectives: Genetic exploration of genome has resulted in several susceptibility loci isolation confirming the genetic association of osteoarthritis. The Japanese population has shown higher incidence of osteoarthritis in patients having intronic and core promoter SNP in CALM-1 gene. At the same, Caucasian and Greek population showed absence of any such predisposition in their population with the CALM-1 gene SNP. Methods: We planned a case control study in osteoarthritis knee to study the presence of CALM-1 gene SNP, correlation of its presence with osteoarthritis and its correlation with clinico-radiological stage of the disease. 120 cases and 120 controls were enrolled. Clinicoradiological features were noted and symptomatic clinical scoring was done. Genetic polymorphism in relation to intronic region of Calm-1 gene was studied by DNA extraction, PCR and RFLP method. Statistical analysis was done using Stata software Results: 39 (32.50%) cases and 18 (15%) controls showed the presence of SNP which was significant (P value = 0.0022). Among SNP positive cases and controls, 5 (8.7%) cases and none controls were heterozygosis for the occurrence of SNP. On regression of affecting variables against SNP, taking the presence of osteoarthritis as dependent variable, we calculated the adjusted odds ratio of all the significant variables. Thereafter, on logistic regression to see the effect of variables on the occurrence of disease, we found age, sex, and presence of SNP affecting the occurrence of disease significantly (p value < .05). Conclusions: CALM-1 gene intronic SNP (rs3213718) is present in Indian Population. The target SNP is significantly affecting the disease as the difference between cases and controls is highly significant (p value = .0022). Females are more predisposed for OA. Mean age of presentation in cases was 53.31+/-9.5 years. Age is a significant factor in causation of disease. However it is not influenced by existence of SNP. Between cases and controls, height, weight and BMI did not show any significant difference.

1288F

Genetic risk factor in development and progression of osteoarthritis knee. *R.N. Srivastava, A. Mishra, S. Raj, D. Sanghi.* Orthopaedic Surgery, KG Medical College, CSM Medical University, Lucknow, UP, India.

Overview: Background: Genome sequencing resulted in identification of several susceptibility loci confirming the genetic association of disease. Objective: Investigate the association of SNP in GDF- 5(BSiE1) and ESR á (BtgI) gene with osteoarthritis knee (KOA). Methodology: A case-control study, 300 cases with knee osteoarthritis and an equal number of age, sex matched healthy controls were included. Cases were diagnosed using the ACR Guidelines of knee osteoarthritis (KOA). Clinical symptoms were assessed with the knee specific WOMAC index and VAS for knee pain. The severity of disease was determined by radiological KL grades (Kellgren-Lawrence). Blood was drawn from patients genomic DNA isolated from blood polymerase chain reaction coupled restriction fragments length polymorphism (PCR-RFLP) was carried out to identify the SNPs. Results: The variant genotype of GDF-5 were found to be present at significantly higher frequency in cases than in controls, resulting in about 1.62 fold increase of KOA risk (P Value=0.040). Likewise increase in frequency of heterozygous genotypes of ESR-á increased the risk to KOA (O.R. 1.49; 95% C.I. 1.06–2.08). Risk was also found to be increased in the cases that carried combination of variant genotypes of ESR á and GDF-5 (O.R.5.62; 95% C.I. 1.57–20.05). OA knee was found to be significantly associated with BMI (P Value= 0.00). A significant association was found with clinical score of knee OA (VAS, WOMAC). Cases with variant genotype of GDF-5 and ESR- á shows significant association with clinical score of knee OA, though no significant association of these SNPs was observed with radiological score. Conclusions: The results suggest that GDF-5 (BSiE1) and ESR-á (BtgI) gene polymorphisms are associated with osteoarthritis knee and gene-gene interaction may influence the development of KOA. Likewise association of these SNPs with clinical score have again demonstrated the role of genetic polymorphism on GDF-5 and ESR- á in the development and progression of KOA .

1289W

Functional variant within vitamin D metabolism gene *ACADSB* is associated with a more severe disease in multiple sclerosis. *M.F. George[1], F.B.S. Briggs[1], P.P. Ramsay[1], H. Quach[1], A. Bernstein[2], B. Acuna[2], L. Shen[2], E. Mowry[3], C. Schaefer[2], L.F. Barcellos[1].* 1) Epidemiology, University of California, Berkeley, Berkeley, CA; 2) Division of Research, Kaiser Permanente, Oakland, CA; 3) Johns Hopkins Hospital, Baltimore, MD.

Multiple sclerosis (MS) is an autoimmune disease affecting the central nervous system with genetic and environmental hypothesized predictors of disease progression. The progression of MS is heterogeneous, with varying manifestations in cognitive and physical systems. Vitamin D has been shown to have a protective effect on both disease onset and progression. We investigated the effect of functional variants within vitamin D metabolism genes on cognitive status and MS severity score (MSSS), as important clinical phenotypes in 1008 white non-Hispanic MS patients identified from a large MS registry within Kaiser Permanente Northern California. A total of 15 candidate genes involved in vitamin D metabolism were studied (total 65 SNPs with putative function derived from whole genome profiles, Illumina 660K BeadChip). Cognitive score for each MS case was determined using a validated telephone interview cognitive status assessment tool. MSSS was calculated for each case using the expanded disability status score (EDSS) at study entry and disease duration, which was defined as the elapsed time (in years) between the first symptom and EDSS assessment. Within *ACADSB*, rs1131430 was associated with increased risk of disease progression. The crude association with MSSS, as measured continuously, was significant (β=0.46; p=0.043). Results remained significant after adjusting for age, gender, history of smoking, history of infectious mononucleosis, education, age of onset, history vitamin D supplementation, disease course, family history of MS, cognitive score, use of disease modifying therapy, weighted genetic risk score, and *HLA-DRB1*15:01* status. MSSS was also analyzed as a dichotomous variable (<5 [low] vs. ≥5 [severe]). Rs1131430 was significantly associated with severe MSSS (OR=1.69; p=0.002), even when adjusting for above covariates. No vitamin D metabolism gene variants were associated with cognitive status. In our population of white non-Hispanic MS cases living in Northern California, cases with the missense change (rs1131430) within *ACADSB* on chromosome 10 at codon 316 Ile->Val, were more likely to have a more severe disease. Our results support a role for vitamin D metabolism in MS disease progression. Replication and functional studies are currently under way in our laboratory.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1290T

Homozygous c.14576G>A Variant of *RNF213* Predicts Early-Onset and Severe Form of Moyamoya Disease. *S. Miyatake[1], N. Miyake[1], H. Touho[2], H. Doi[1], H. Saitsu[1], K. Shimojima[3], T. Yamamoto[3], N. Okamoto[4], N. Kawahara[5], Y. Kuroiwa[6], M. Taguri[7], S. Morita[7], Y. Matsubara[8], S. Kure[9], N. Matsumoto[1].* 1) Department of Human Genetics, Yokohama City University, Yokohama, Japan; 2) Touho Neurosurgical Clinic, Osaka, Japan; 3) Tokyo Women's Medical University Institute for Integrated Medical Sciences, Tokyo, Japan; 4) Division of Medical Genetics, Osaka Medical Center and Research Institute for Maternal and Child Health, Osaka, Japan; 5) Department of Neurosurgery, Yokohama City University Graduate School of Medicine, Yokohama, Japan; 6) Department of Clinical Neurology and Stroke Medicine, Yokohama City University Graduate School of Medicine, Yokohama, Japan; 7) Department of Biostatistics and Epidemiology, Yokohama City University Medical Center, Yokohama, Japan; 8) Department of Medical Genetics, Tohoku University School of Medicine, Sendai, Japan; 9) Department of Pediatrics, Tohoku University School of Medicine, Sendai, Japan.

Moyamoya disease (MMD) is a cerebrovascular disease which is now a relatively common cause of pediatric strokes. MMD can lead to devastating neurological deficits and intellectual impairments if left untreated. Preoperative infarctions, early age of onset, intellectual impairment, seizure and progressive posterior cerebral artery stenosis are known prognostic factors. Surgical revascularization can improve the cerebrovascular hemodynamics, and prevent subsequent attacks for the ischemic type of MMD. Thus, early diagnosis and surgical intervention are very important. Recently *RNF213* was reported as a susceptibility gene for MMD. However its clinical relevance remained unknown. Our aim was to clarify the correlation between *RNF213* genotype and MMD phenotype. The entire coding region of the *RNF213* gene was sequenced in 204 MMD cases, and corresponding variants were checked in 62 pairs of parents, 13 mothers and 4 fathers of the cases, and 283 normal controls. Clinical information was collected from clinicians. Genotype-phenotype correlations were statistically analyzed. The c.14576G>A variant was identified in 95.1% of familial MMD, 79.2% of sporadic MMD, and in 1.8% of controls, thus confirming its association with MMD, with an odds ratio of 259 and P<0.001 for either heterozygotes or homozygotes. Homozygous c.14576G>A was observed in 15 cases, but not in the controls and unaffected parents. The incidence rate for homozygotes was calculated to be >78%. Homozygotes had significantly earlier age of onset compared with heterozygotes or wild-types (median age of onset=3, 7, 8 years, respectively). Sixty-percent of homozygotes were diagnosed with MMD before age 4, comparing 15% of heterozygotes, and 14.3% of wild-types, respectively, and all had infarctions as the first symptom. Furthermore, infarctions at initial presentation and involvement of posterior cerebral arteries were of significantly higher frequency in homozygotes than heterozygotes and wild-types. These features indicate that c.14576G>A homozygotes have more severe and wider vasculopathy in the brain. Variants other than c.14576G>A were not associated with clinical phenotypes. In conclusion, the homozygous c.14576G>A variant could be a good DNA biomarker for predicting the severe type of MMD, in which early medical/surgical intervention is recommended, and may provide a better monitoring and preventing strategy.

1291F

C reactive protein as a prognostic marker in melanoma progression. *S. Fang[1], Y. Wang[2], D. Sui[3], H. Liu[2], C. Schacherer[2], J. Gardner[2], M. Ross[2], J. Gershenwald[2], J. Reveille[4], L. Wang[5], Q. Wei[5], C. Amos[1], J. Lee[2].* 1) Genetics, UT M.D Anderson Cancer Ctr, Houston, TX; 2) Department of Surgical Oncology, UT M.D Anderson Cancer Ctr, Houston, TX; 3) Department of Biostatistics, UT M.D Anderson Cancer Ctr, Houston, TX; 4) Division of Rheumatology and Clinical Immunogenetics, The University of Texas Health Science Center at Houston; 5) Department of Epidemiology, UT M.D Anderson Cancer Ctr, Houston, TX.

CRP has been proposed as a prognostic marker in melanoma patients, but the relationship of serial CRP levels to patient outcomes during longitudinal follow-up has not been evaluated. This study aimed to investigate whether increased CRP levels are associated with melanoma stage of disease, recurrence (DFS) and overall survival (OS), and whether changes in CRP levels among patients undergoing serial blood draw are related to dynamic changes in disease status during follow-up. Blood samples were obtained from 1,804 melanoma patients at MD Anderson Cancer Center from 1997 to 2009. 587 initial and 557 confirmatory patients underwent plasma CRP tests. Kaplan-Meier methods and Cox regression models were used to evaluate the association of CRP with stage and disease progression. 115 patients underwent sequential blood draws for CRP; in these patients we evaluated the relationship between change in disease status and change in CRP level using non-parametric tests. Finally, using a genome-wide scan of the entire patient population, we examined the relationship of significant and candidate SNPs with CRP blood levels and melanoma susceptibility, DFS and OS. An elevated plasma CRP level was associated with a poorer OS in both the initial and confirmatory datasets. CRP contributed to the risk death by a factor of 1.40 (95% confidence interval, 1.25–1.57) per unit increase of logarithmic CRP in the combined dataset. As compared with CRP<10 mg/l, CRP>= 10 mg/l conferred an elevated risk of death in patients with all stages of melanoma, as well as in the subsets of patients with stage I/II or stage III/IV disease, and an elevated risk of cancer recurrence among stage I/II patients. Investigation of serial CRP levels with disease status demonstrated strong correlation between an increase in CRP level and melanoma progression. Finally, we detected evidence of association between CRP-region genetic polymorphisms, plasma CRP level and melanoma development and progression. Our study confirms that the elevated CRP levels are independently associated with a poorer survival duration in both early- and advanced-stage patients. Elevated CRP levels predict recurrence in patients without measurable disease, and are also dynamically associated with disease response in patients with metastasis. These results suggest that CRP may be a valuable prognostic marker in melanoma patients, and that targeting CRP-associated inflammation in melanoma patients may have therapeutic value.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1292W

***TGFBR1* variant associated with constitutively decreased TGF-β signaling and risk for colorectal cancer.** *M.J. Pennison[1], N. Bellam[1], J. Zimmerman[1], Q. Zeng[1], M. Wang[1], M. Sadim[2], V. Kaklamani[2], N. Yi[1], K. Zhang[1], J. Baron[3], D.O. Stram[4], B. Pasche[1], CCFR Investigators, Colon Cancer Family Registry, National Cancer Institute.* 1) Medicine, Hematology/Oncology, University of Alabama at Birmingham, Birmingham, AL; 2) Medicine, Hematology/Oncology, Northwestern University, Chicago, IL; 3) Medicine, Dartmouth School of Medicine, Hanover, NH; 4) Preventative Medicine, University of Southern California, Los Angeles, CA.

Purpose: Twin cohort studies show approximately one third of all colorectal cancer (CRC) cases are inherited, but the genetic cause for the majority of these cases is unknown. Germline variants within the TGF-β signaling pathway serve as strong candidates for CRC risk modifiers due to TGF-β's early anti-tumor signaling activation of SMAD proteins. We have previously shown that constitutively decreased *Tgfbr1* expression in mice leads to reduced SMAD activation and increased CRC development. These findings were validated in humans and we found that constitutively decreased *TGFBR1* signaling was associated with CRC risk. In this family-based study, we aimed at determining the association of *TGFBR1* and *SMAD7* with CRC risk. **Methods:** The study included 1,558 colorectal cancer cases and 2,640 healthy controls recruited by the Colorectal Cancer Family Registry (CCFR). A total of 14 *TGFBR1* haplotype tagging SNPs were genotyped using ABI's TaqMan SNP genotyping platform, as well as a *SMAD7* SNP identified in GWA studies and validated in several case control studies. Results were analyzed using logistical regression analysis utilizing the Cockerham genetic model (determines additive and/or dominant genetic effect) and a codominant model (determines relative risk of the rare homozygous and heterozygous genotypes for each SNP). Risk prediction accuracy (receiver operating characteristic curves) of the dataset was also determined. Differential *TGFBR1* gene expression and downstream protein expression was assessed between high and low-risk genotypes using groups of patient-derived lymphoblastoid cell lines (LCLs). **Results:** *TGFBR1* rs7034462 (OR, 2.85; CI, 2.02–4.01, P value, 5.62×10^{-13}) and *SMAD7* rs4939827 (OR, 0.80; CI, 0.70–0.91, P value, 7.0×10^{-4}) were both strongly associated with CRC risk. Gene expression profiling showed a significant reduction in *TGFBR1* gene expression in 8 LCLs carrying high-risk genotypes versus a group of 8 low-risk LCLs (P = 0.042). Protein expression analysis performed with 3 matched LCL pairs showed significantly reduced TGF-β mediated SMAD signaling in cells carrying the high-risk alleles of these two SNPs. **Conclusion:** Our study shows that the *TGFBR1* SNP, rs7034462, is strongly associated with constitutively decreased TGF-β signaling and risk for colorectal cancer. This is also the first evidence provided that links rs7034462 to decreased expression of *TGFBR1*.

1293T

Limited survival of MEN1 patients with mutations in the JunD interacting domain and first elements for intrafamilial correlation in MEN1 syndrome. A study from the GTE cohort (Groupe d'étude des Tumeurs Endocrines). *J. Thevenon[1, 2], A. Bourredjem[3], L. Faivre[1, 2], A. Costa[4], E. Gautier[3], C. Bonithon-Kopp[3], A. Calender[5], C. Binquet[3], P. Goudet[4, 6], the Members of the GTE.* 1) Centre de Génétique et Centre de Référence (Anomalies du Développement et Syndromes Malformatifs), Hôpital d'Enfants, CHU Dijon, France; 2) Equipe émergente GAD EA-4271, IFR Santé STIC, Université de Bourgogne, Dijon, France; 3) Centre d'investigation clinique -épidémiologie clinique/essais cliniques, CHU, Dijon, France; 4) Centre d'épidémiologie des populations, EA 4184, Faculté de médecine de Dijon, Université de Bourgogne, Dijon, France; 5) Laboratoire de génétique moléculaire et médicale, hôpital Édouard-Herriot, 5, place Arsonval, 69437 Lyon cedex 03, France; 6) Centre Hospitalier Universitaire de Dijon, Service de Chirurgie Endocrinienne, Dijon, France.

Background: Multiple Endocrine Neoplasia syndrome type 1 (MEN1) is an autosomal dominant rare affection predisposing to endocrine tumours. The disease is secondary to mutations in the tumour suppressive *MEN1* gene. Although genotype-phenotype studies failed to identify a statistical correlation, some families harbour recurrent tumoral patterns. The MENIN protein function was elucidated through the discovery of its interacting partners. Point mutations were shown to alter MENIN interaction with its partners leading to inappropriate transcriptional regulation and tumorigenesis. This is the first genotype-phenotype correlation study in MEN1 disease considering a time dependant statistical analysis, testing intra-familial correlation and evaluating the impact of the mutation localization in MENIN interacting domains in a large cohort of MEN1 patients.

Methods: We report on the cohort of MEN1 patients, issued from the GTE network. Patients with *MEN1* mutation and clinical follow up were included, i.e. 262 families, 806 patients. Mutations type, localization and 9 interacting factors of the MENIN protein were compared to occurring tumour types and death ages. Assessment of the intra-familial correlation and genotype-phenotype correlation was performed using a frailty Cox model.

Results: Significant interfamilial variability was evidenced in this cohort. There was no significant correlation between mutation type and the different phenotypes. Overall survival was significantly reduced when mutations affected the JunD interacting domain (HR=2.13 (CI-95%[1.26; 3.60]). Patients had a higher risk of death from cancer of the MEN1 spectrum (HR= 2.44 (CI-95%[1.25; 4.77]).

Conclusion: This large genotype-phenotype correlation study with a time to event analysis highlighted the intrafamilial correlation of the MEN1 syndrome. As previously known, there was no direct genotype phenotype correlation but the higher risk of death in the patients carrying mutations affecting the JunD interacting domain is the first argument for a genotype-phenotype correlation in MEN1 syndrome. This result suggests that loss of interaction between JunD and MENIN might result in the development of more aggressive tumors. Modifying factors might be the comprehensive key of MEN1 syndrome variable expressivity.

1294F

Association of vitamin D receptor gene polymorphism and vitamin D status in knee osteoarthritis. *D. Sanghi, R. Srivastava, S. Raj.* Orthopaedic Surgery, KG Medical College, CSM Medical University, Lucknow, UP, India.

INTRODUCTION: Evidence suggests that low serum levels of vitamin D may increase the severity of Osteoarthritis (OA). VDR gene polymorphism is known for its association with osteoporosis. The inverse relationship between osteoporosis and OA suggests that VDR gene polymorphism is a candidate gene to be associated with OA. OBJECTIVES: This study was done to analyze the association of vitamin D receptor gene polymorphism (Taq1 and Apa I) and serum vitamin D levels in knee osteoarthritis METHODS: This case control study consisted of 180 Osteoarthritis Knee patients and 150 controls. Cases were clinically diagnosed according to ACR criteria. Gradation of the disease was done by using KL grading system on the basis of radiological findings. The serum levels of vitamin D were assessed by using kit of Enzyme Linked Immunosorbent Assay. Detection of VDR gene polymorphisms (Taq1 and Apa I) were done by PCR-RFLP technique RESULTS: We observed an insignificant association for genotypes of TaqI(p=0.086) and ApaI(p=0.60) polymorphism between cases and controls.However, for TaqI marginal significant association (p=0.053, OR 1.4, 95% CI 1.008–1.945) was observed between wild type(T) allele and mutant type(t) allele, but for Apa I, there was no significant difference (p=0.334, OR 1.17, CI 1.367–1.867) between wild type(A) allele and mutant type(a) allele.We observed a significant association of low level of serum vitamin D levels in homozygous mutant(tt) genotypes as compared to heterozygous(Tt) and wild type(TT) genotypes. But In case of ApaI, we found higher serum vitamin D levels in homozygous mutant(aa) genotype in comparison to heterozygous(Aa) and wild type(AA) genotype but this difference was not statistically significant CONCLUSION: Though insignificant association was found between osteoarthritis and the genotypes of TaqI and ApaI polymorphism, a significant association with mutant allele of TaqI(t) was observed. Additionally, the association of mutant allele with reduced level of vitamin D was noted.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1295W

Association between neural genes DRD2, AVPR1a, and ASPM and Endophenotypes of Speech Sound Disorder. *C.M. Stein[1], F. Qiu[1], B. Truitt[1], R. Raghavendra[1], P. Joseph[1], A.A. Avrich[1], R.P. Igo, Jr.[1], J. Tag[2], L. Freebairn[2], H.G. Taylor[2], B.A. Lewis[2], S.K. Iyengar[1,3].* 1) Epidemiology & Biostatistics, Case Western Reserve Univ, Cleveland, OH; 2) Psychological Sciences, Case Western Reserve University, Cleveland, OH; 3) Genetics, Case Western Reserve University, Cleveland, OH.

Speech sound disorder (SSD) is one of the most common types of communication disorders, with prevalence rates of 16% at 3 years of age, and an estimated 3.8% of children continuing to present speech delay at 6 years of age. Several studies have identified promising candidate gene associations between communication disorders and genes in brain and neuronal pathways. In this study, we focused on three genes related to brain and central nervous system functioning: dopamine D2 receptor (DRD2), arginine-vasopressin receptor 1a (AVPR1a), and microcephaly gene ASPM. We genotyped tag SNPs in these genes in 490 children from 178 families ascertained through a proband with SSD. TagSNPs were selected with an R-squared threshold of 80% and 5% MAF for a Caucasian population; there were 4 SNPs in ASPM, 7 in AVPR1a, and 27 in DRD2. Our phenotypes of interest included measures of phonological memory (nonsense word [NSW] and multisyllabic word [MSW] repetition), vocabulary (PPVT and EOWPVT), and reading decoding (WRMT-AT and WRMT-ID); these traits are highly correlated (20–40%). We analyzed the data using GenABEL in order to model familial structure and polygenic variance. In DRD2, we found several SNPs that were statistically significant after correction for multiple comparisons: for reading decoding (rs10891552 $p=8\times10^{-5}$ and rs2734831 $p=1\times10^{-4}$), and both measures of phonological memory (rs12805897 $p=2\times10^{-6}$, rs17529477 $p=8\times10^{-7}$, and rs7125415 $p=4\times10^{-6}$). In AVPR1a, we observed statistically significant associations with both reading decoding measures (rs11174811 $p=3\times10^{-6}$ and rs1587097 $p=6\times10^{-7}$). Other SNPs within DRD2 and AVPR1a were associated with these traits, as well as vocabulary, at an uncorrected α=0.05. Finally, SNPs in ASPM were associated with vocabulary, reading decoding, and speech sound disorder, though these associations were not statistically significant after multiple testing correction. These results show that neural genes are associated with key domains of SSD, including phonological processing, vocabulary, and reading decoding.

1296T

Parent-of-origin effects and gender differences influence age-at-onset variation in FAP ATTRV30M kindreds. *C. Lemos[1, 2], T. Coelho[3], A. Martins-da-Silva[3], J. Sequeiros[1, 2], A. Sousa[1, 2].* 1) UnIGENe, IBMC, Porto, Portugal; 2) ICBAS, Instituto Ciências Biomédicas Abel Salazar, Univ. Porto, Portugal; 3) Unidade Clínica de Paramiloidose (UCP), Centro Hospitalar do Porto (CHP), Porto, Portugal.

Familial amyloid polyneuropathy (FAP) ATTRV30M is an autosomal dominant systemic amyloidosis, due to a point mutation in the transthyretin (TTR) gene. More than 100 mutations have been found in the TTR gene but V30M is by far the most common. A wide variability in age-at-onset (AO) has been uncovered, including among Portuguese patients [17–82 yrs]. Early (≤40) and late-onset (≥50) cases are not separate entities, often coexisting in the same family, with offspring showing anticipation - a much earlier AO than their affected parent. Historically, anticipation was mostly ascribed to ascertainment biases. Previous studies with Portuguese, Swedish and Japanese families have shown the presence of a true marked anticipation. Our aim was to study anticipation in a larger number of kindreds than assessed before, removing possible biases and to gain more insight into parent-of-origin effects. From the UCP-registry, we analysed 926 parent-offspring pairs, both clinically observed and with well-established AO. Women had a statistically significant higher AO than men, either for daughters (mean, SD - 33.7, 6.84) vs. sons (29.43, 6.08) or mothers (39.57, 11.75) vs. fathers (35.62, 11.62). Also, 291 parent-offspring pairs showed marked anticipation (≥10 years) and the transmitting parent was the mother in 203 pairs. Conversely, among the 22 offspring showing a 10 years higher AO, 19 had a transmitting father. Mother-son pairs showed larger anticipation (10.43, 9.34) while the father-daughter pairs showed only residual anticipation (1.23, 9.77). To remove possible biases, we repeated these analyses: 1) excluding the proband, 2) removing pairs with simultaneous onset and 3) excluding offspring born after 1960. Anticipation was found in all subsamples and the same trend of parent-of-origin effects was observed. Noteworthy, no parent with AO ≤ 40 had an offspring with AO ≥50. These findings confirm anticipation as true biological phenomenon. Furthermore, both parents and offspring's gender were found to be highly significant factors for anticipation. We also found that mother-son pairs showed higher anticipation. The study of genetic modifiers should focus on families, aiming to unravel mechanisms of anticipation that may have important clinical implications.

1297F

Genome-wide analysis of germline copy number aberrations and association with breast cancer susceptibility. *Y. Sapkota[1, 5], B.S. Sehrawat[1], S. Ghosh[2, 5], P.J. Robson[3, 4], C.E. Cass[2, 5], J.R. Mackey[2, 5], S. Damaraju[1, 5].* 1) Department of Lab Medicine and Pathology, University of Alberta, Edmonton, Alberta, Canada; 2) Department of Oncology, University of Alberta, Edmonton, Alberta, Canada; 3) Department of Agricultural, Food and Nutritional Sciences, University of Alberta, Edmonton, Alberta, Canada; 4) Alberta Health Services - Cancer Care, Edmonton, Alberta, Canada; 5) Cross Cancer Institute, Edmonton, Alberta, Canada.

Background Breast cancer susceptibility determinants include both genetic and non-genetic factors. DNA variations identified thus far in the germline account for less than 30% of genetic risk, and additional unidentified variants likely contribute to breast cancer predisposition. Recent studies have shown that copy number aberrations (CNAs) are the most common type of structural variation in human genomes and may be associated with complex diseases, including familial breast and BRCA1-associated ovarian cancers. Our interest is in the role of germline CNAs as genetic determinants in sporadic breast cancers (late age at onset of disease). Methods We accessed hybridization intensity files (Affymetrix SNP 6.0 array) for 302 sporadic breast cancer cases and 321 apparently healthy individuals from our previous studies. All subjects were women, predominantly of Caucasian origin. CNAs were identified using Single Nucleotide Polymorphism-Fast Adaptive States Segmentation Technique 2 (SNP-FASST2) algorithm (a hidden markov model-based approach) implemented in Nexus Copy Number 6.0. Frequencies of CNAs between cases and controls were compared using Fisher's exact test (association test, two-sided). P values were adjusted for multiple hypotheses testing using false discovery rate (Q value). Results SNP-FASST2 algorithm identified 21,257 germline CNAs in 623 subjects with an average of 149 copy number events (gains/losses) and 302 allelic events (copy-neutral loss of heterozygosities) per sample. Frequency comparisons between cases and controls indicated that 162 copy number events (62 gains and 100 losses) were significantly associated with breast cancer (Q<0.05), and a majority of these CNAs were reported in the Database for Genomic Variants. Of these, copy number losses at Xp22.33 (Q=9.7E-15) and 19q13.2 (Q= 2.7E-10) and a copy number gain at 14q32.33 (Q=2.3E-05) exhibited the largest differences in frequencies between cases and controls (>20%). Gene ontology analysis for genes within 162 copy number events revealed “regulation of epidermal growth factor receptor signalling pathway” as the most significant pathway. Genes identified in this pathway were rhomboid 5 homolog 1 (Drosophila) and cadherin 13, H-cadherin (heart). Conclusions We have identified candidate germline copy number events strongly associated with breast cancer susceptibility. Technical validation of the highly significant CNAs and independent replication of these findings is warranted.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1298W

Whole-Genome Detection of Disease-Associated Deletions or Homozygosity in a Case-Control Study of Rheumatoid Arthritis. *C.C. Wu[1], S. Shete[2], E.J. Jo[3], Y.E. Lu[4], Y. Xu[5], W.V. Chen[4], C.I. Amos[4]*. 1) Dept Epidemiology, Univ Texas MD Anderson CA Ctr, Houston, TX; 2) Dept Biostatistics, Univ Texas MD Anderson CA Ctr, Houston, TX; 3) Duncan Cancer Center, Baylor College of Medicine, Houston, TX; 4) Dept Genetics, Univ Texas MD Anderson CA Ctr, Houston, TX; 5) Biostatistics, School of Public Health, Yale University, New Haven, Connecticut.

Compared with genome-wide association studies, comprehensive studies of the contribution of copy number variations to complex human disease susceptibility are generally lacking. Deletion copy number variations have been known to cause microdeletion syndromes and are frequently observed in patients with neuron-developmental disorders. Because genomic deletions are abundant in humans and represent one of the least well studied classes of genetic variants and because known rheumatoid arthritis susceptibility loci explain only a small portion of familial clustering, we performed a genome-wide study of association between deletion or excess homozygosity and rheumatoid arthritis using high-density 550K SNP genotype data from a genome-wide association study. We applied a genome-wide statistical method that we recently developed to test each contiguous SNP locus between the 868 cases and 1197 controls from the North American Rheumatoid Arthritis Consortium for detecting deletion variants or excess homozygosity that influence susceptibility. Our method is designed to detect statistically significant evidence of deletions or excess homozygosity at individual SNPs for SNP-by-SNP analyses and to combine the information among neighboring significant SNPs for cluster analyses. In addition to successfully detecting the known deletion variants on HLA-DRB1 and C4 in HMC, we identified the respective 4.3-kb and 28-kb clusters on chromosomes 10p and 13q, which were significant at a corrected 0.05 nominal significance level, adjusted for multiple comparison procedure. Independently, we ran PennCNV, an algorithm for identifying and cataloging copy numbers for individuals based on hidden Markov model, to identify cases and controls that had chromosomal segments with copy number < 2. Using Fisher's exact test for comparing the numbers of cases and controls per SNP, we identified 26 significant SNPs (protective; more controls than cases) aggregating on chromosome 14 with p-values < 10-8 and 49 SNPs on chromosomes 2, 14, and 20 with p-values between 10-5 and 10-8. In addition, we found that 12 cases and 1 control commonly shared a 6.6-kb segment with copy number = 1, which lay between 2 adjacent significant SNPs on chromosome 19p (2,054,962-2,165,057) detected by our cluster method. The Fisher's one-sided exact test for comparing 12 cases and 1 control gives a p-value of 0.0117. The PennCNV method did not successfully detect the known deleterious genetic variants in MHC.

1299T

T1D risk score modeling using dense genotypes in autoimmune associated genomic regions in 6,670 cases and 9,416 controls and validation in 2,601 affected sibpair families and 69 trio families. *W.-M. Chen[1,2], S. Onengut-Gumuscu[1,3], P. Concannon[1,4], S.S. Rich[1,2], T1DGC*. 1) Center for Pubic Health Genomics, University of Virginia, Charlottesville, VA; 2) Department of Public Health Sciences, University of Virginia, Charlottesville, VA; 3) Department of Medicine Division of Endocrinology, University of Virginia, Charlottesville, VA; 4) Department of Biochemistry and Molecular Genetics, University of Virginia, Charlottesville, VA.

Type 1 diabetes (T1D) is the third most prevalent chronic disease of childhood and results from the autoimmune destruction of the pancreatic β cells. Although genome-wide association studies (GWAS) have identified ~41 distinct chromosomal regions associated with T1D, the effect on T1D from each individual SNP is small, and their joint effect on T1D and clinical relevance have not been extensively investigated. In a recent study, 196,524 variants in 186 distinct autoimmune associated genomic regions were genotyped using the ImmunoChip SNP panel in 6,670 T1D cases, 9,416 controls and 10,796 individuals from 2,601 affected sibpair (ASP) families and 69 trio families. The dense genotypes and the large sample size provided a unique opportunity to estimate the joint effect of currently known T1D candidate genes and to predict the T1D risk using a risk score model.

A SNP-based risk score may be used to predict the T1D status of an individual. We carried out a logistic regression analysis in the case-control data, modeling all statistically significant variants that were identified from the genome scan and the conditional analyses in associated regions, including 27 SNPs from the MHC region and 31 SNPs from non-MHC regions. A risk score for each individual was then derived as a linear combination of the associated SNPs with the weights being the regression coefficients. In this model, the MHC SNPs explained 5.7-fold more T1D variation than non-MHC SNPs. This risk score was able to predict the T1D status in the 6,670 and 9,416 controls with AUC=0.924. Due to the severely inflated risk scores among unaffected parents of the affected sibpairs, the prediction accuracy in the family data is underestimated as AUC=0.755. We developed a novel algorithm to estimate the effect size of risk scores in family data. After normalization to the observed distribution in the controls, the mean risk score in controls and cases was 0 and 1.538 respectively, with a difference 1.538 (se=0.014). The difference in risk scores between cases and controls in the random population was estimated using the family data as 1.446 (se=0.023) and the mean in cases was 1.592. The consistency of our MHC risk score estimate between the case-control and family data further validates the accuracy and utility of our risk score model in the prediction of T1D.

1300F

Variants in/near *DNER* are Reproducibly Associated with Type 2 Diabetes in Pima Indians. *Y.L. Muller, M. AbdusSamad, R. Hanson, W.C. Knowler, C. Bogardus, L. Baier*. Diabetes Molecular Genetics Section, Phoenix Epidemiology and Clinical Research Branch, National Institutes of Health, Phoenix, AZ., USA.

A prior genome-wide association study (GWAS) in 1149 Pima Indians using the Affymetrix 1-million SNP chip identified variants in/near *DNER* that were associated with type 2 diabetes (e.g. rs1861612, odds ratio=1.64 per copy of the risk allele, p=0.0002, adjusted for age, sex and birth-year). *DNER*, positioned on human chr.2:230222345-230579286, encodes a Delta/Notch like EGF-related receptor. Recent studies have shown that DNER mediates notch signaling via cell-cell interaction. DNER is expressed in human pancreatic islets, and the Notch signaling pathway plays a critical role in pancreatic development and β-cell fate. To follow-up on our GWAS findings, exons and exon-intron boundaries of *DNER* were sequenced in 180 Pima Indians to identify potentially functional novel variants. Additional tag SNPs (AF≥0.05, r^2≥0.8) were selected from the 1M GWAS to provide coverage in un-sequenced intronic and flanking regions (~25kb flanking each end of the gene). These 55 SNPs spanning *DNER* were genotyped in a population based sample of 3501 full-heritage Pima Indians for association analysis. Selected tag SNPs were further genotyped in a replication sample of 3723 Native Americans. Rs1861612, with a risk allele frequency of 0.64, was associated with type 2 diabetes in the full-heritage Pima Indian sample ($p=4.8\times10^{-5}$, OR=1.29, adjusted for age, sex, birth-year and family membership) and the association replicated in the separate Native American sample ($p=4.0\times10^{-4}$, OR=1.28, adjusted for age, sex, birth-year and heritage). Combining the full-heritage and replication samples (n=7224) provided the strongest evidence for association with diabetes for rs1861612 that achieved the statistical threshold typically required for genome-wide significance (adjusted $p=6.6\times10^{-8}$, OR=1.29). This SNP was also genotyped in 415 non-diabetic full-heritage Pima subjects who had been metabolically phenotyped for quantitative traits predicting T2D. The diabetes risk allele was associated with increased fasting plasma insulin levels (p=0.002, adjusted for age, sex, family membership and % body fat). Our data suggest that *DNER* may play a role in diabetes susceptibility in Native Americans.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1301W

Two multiple sclerosis risk variants from a large genome-wide association study are associated with disease progression in a population-based case cohort. *H. Quach[1], M.F. George[1], F.B.S. Briggs[1], P.P. Ramsay[1], B. Acuna[2], L. Shen[2], E. Mowry[3], A. Bernstein[2], C. Schaefer[2], L.F. Barcellos[3].* 1) Sch Pub Hlth, Univ California, Berkeley, Berkeley, CA; 2) Division of Research, Kaiser Permanente, Oakland, CA; 3) Johns Hopkins Hospital, Baltimore, MD.

Multiple sclerosis (MS) is an autoimmune disease affecting the central nervous system with genetic and environmental hypothesized predictors of disease progression. MS progression is heterogeneous, with varying manifestations in cognitive and physical systems. We investigated the effect of 52 risk loci from the largest genome-wide association study (GWAS) in MS (IMSGC and WTCCC, Nature 2011) and the combined weighted genetic risk score (wGRS) on three clinical phenotypes: cognitive status, MS severity score (MSSS), and age of onset in 1008 white MS patients identified from a large MS registry within Kaiser Permanente Northern California. We calculated wGRS by combining the weighted odds ratios from the 52 risk loci. The weight for each variant was the natural log of the odds ratio for each allele determined through GWAS (N≈30,000, Nature 2011). All samples were genotyped using the Illumina 660K platform. Cognitive score for each case was determined using a validated telephone interview cognitive status assessment tool. MSSS was calculated for each case using the expanded disability status score (EDSS) at study entry and disease duration. Age of onset was defined as the age at which first neurologic symptoms manifested. We investigated each of the 52 loci that comprise the wGRS with all clinical outcomes. Results show rs2303759, located in the promoter region of *DKKL1*, was significantly associated with cognitive score (β=-0.71; p=0.001). After adjusting for age, gender, history of smoking, history of infectious mononucleosis, disease course, use of disease modifying therapy, disease duration, and *HLA-DRB1*15:01* status, this association remained significant (β=-0.69; p=0.001). Rs2119704, located in the promoter region of *GALC*, was associated with MSSS when comparing the extreme phenotypes (<2.5 vs. ≥7.5; OR=0.42; p=0.001). After adjusting for the above covariates, results did not change. Both associations persisted after correction for multiple tests. The wGRS based on MS risk loci was not associated with cognitive status, MSSS, or age of onset. In our population of white non-Hispanic MS cases living in Northern California, two individual risk variants from the largest MS GWAS were more frequent among cases with a less severe disease presentation, in both cognitive and physical systems. This is among the first studies to investigate the effect of wGRS and individual established MS risk loci on important MS phenotypes.

1302T

Unraveling Phenotype Heterogeneity in Prostate Cancer Susceptibility in Finland Utilizing Covariate-Based Analysis. *C.D. Cropp[1], C.L. Simpson[1], T. Wahlfors[3,4], A. George[1,5], M.P.S. Jones[2], U. Harper[2], D. Ponciano-Jackson[2], T. Tammela[6], J. Schleutker[3,4], J.E. Bailey-Wilson[1].* 1) Inherited Disease Research Branch, National Human Genome Research Institute/ National Institutes of Health, Baltimore, MD; 2) Genomics Core/Genome Technology Branch, National Human Genome Research Institute, National Institutes of Health, Rockville, Maryland; 3) Institute of Biomedical Technology/BioMediTech, University of Tampere and Fimlab Laboratories, Tampere, Finland; 4) Department of Medical Biochemistry and Genetics, University of Turku, Turku, Finland; 5) Fox Chase Cancer Center, Philadelphia, Pennsylvania; 6) Department of Urology, Tampere University Hospital, University of Tampere, Tampere, Finland.

Prostate cancer is the most common male cancer in developed countries. Previously, we reported a genome-wide linkage scan in 69 Finnish Hereditary Prostate Cancer (HPC) families, which replicated the *HPC9* locus on 17q21-q22 and identified a locus on 2q37. We used ordered subset analysis (OSA) to detect other loci linked to HPC in subsets of families to identify and detect other loci linked to HPC incorporating age of onset as a trait-related covariate to address genetic heterogeneity and strengthen the linkage findings previously reported. The overall mean age of onset across the families was 66.2±8.8 years while the range of individual onset ages ranged from 46 to 98 years. Although the highest OSA LOD score with a ΔLOD (p=0.02) was 2.876 on 15q26.2-q26.3 in a subset of 40 families ascending by age at onset, no other ΔLOD scores were significant after permutation testing. Since OSA uses a single covariate value per family, we used mean age of onset for each pedigree. To better capture the effect of age on the linkage signal, we used LODPAL to perform a linkage analysis in affected relative pairs, while adjusting for the age of each individual family member as a single covariate. Preliminary results revealed strong evidence of linkage to HPC on chromosome 15q was (LOD=4.9, 132cM) and 8q (LOD=3.1, 157cM). Permutations are ongoing to determine empirical p-values for these LOD scores.

1303F

A search for genetic variants predisposing to radiation induced meningioma. *R. Bruchim[1], B. Oberman[1], R. Milgram[1], I. Novikov[2], S. Sadetzki[1,3].* 1) The Cancer & Radiation Epidemiology Unit, The Gertner Institute, Chaim Sheba Medical Center, Tel-Hashomer, 52621 Israel; 2) The Biostatistics Unit, The Gertner Institute, Chaim Sheba Medical Center, Tel-Hashomer, 52621 Israel; 3) Sackler Faculty of Medicine, Tel-Aviv University, Tel-Aviv, Israel.

Ionizing radiation is the only environmental exposure that has been shown to be a causative risk factor for meninigioma formation. However, only a fraction of previously exposed individuals develop the tumor. This supports the idea that an interaction between radiation and genetic susceptibility conferred by low penetrance genes could converge to facilitate tumor formation. The aim of this study was to test interaction between ionizing radiation and selected candidate genes in the development of radiation induced meningioma. Methods: The study population was taken from the Tinea Capitis (TC) cohort which includes a group of 10,842 individuals treated during the 1950's with radiation therapy for TC and 2 matched population and sibling non-irradiated comparison groups. This unique population has been followed for approximately 50 years for radiation sequelae in our institute. The project was designed as a 4 group nested case control study balanced for irradiation, and included meningioma patients (n=154), healthy controls irradiated for TC in childhood (n=189), sporadic meningioma patients (n=79) and healthy non irradiated controls (n=149). Data was collected by personal interview and blood samples were collected for genotyping of selected SNPs in 2 candidate genes involved in the DNA repair pathway: ATM and BRIP1. Results: One SNP in the ATM gene, rs1800057, showed indication for an association with borderline significance for meningioma formation (OR= 1.60, CI 0.93–2.75; p=0.09). The presence of the G allele in this SNP increased the risk for meningioma among the non-irradiated population by 5 fold (p=0.0063). Of 13 haplotypes constructed, 3 were associated with decreased meningioma risk in irradiated individuals and with increased meningioma risk in non-irradiated individuals. The BRIP1 gene analysis included 23 intragenic SNPs and yielded an association with meningioma for 3 of them (rs4986763, rs11871785, rs1157720). Although the results were not significant after adjustment for multiple comparisons, these SNPs are all located at the 5' of the BRIP1 gene within 72 KB. In addition, 2 haplotypes out of 24 were found to be associated with disease development and indicated a possible interaction between irradiation and meningioma. These results support the role of these genes in the formation of meningioma, and suggest a possible modification effect of the risk by irradiation status.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1304W

Variants in one-carbon metabolism and blood folate, homocysteine and B12 deficiency in a population-based study. *K. S. Crider[1], D. R. Maneval[2], N. F. Dowling[1], L. B. Bailey[3], G. Kaudwell[2], L. Hao[4], Z. Li[4], R. J. Berry[1].* 1) National Center on Birth Defects and Developmental Disabilites, CDC, Atlanta, GA; 2) University of Florida, Ganiesville, FL; 3) University of Gerogia, Athenes, GA; 4) Peking University, National Center of Maternal and Infant Health, Bejing, China.

One-carbon metabolism is critical to basic biological processes including DNA and protein methylation, as well as DNA replication. Genetic variation in this pathway has been associated with changes in disease risk, both for individuals (e.g. cancer, heart disease, mental health) and pregnancy outcomes (e.g. birth defect risk) although some of the reported associations are inconsistent across populations. In order to establish the impact of genetic variation on one-carbon metabolism we examined the association of 18 one-carbon metabolism polymorphisms with blood biomarkers—plasma and RBC folate, homocysteine and B12 concentrations—in a population-based study of reproductive age women from Northern China (n = 1019). This population has a high prevalence of the *MTHFR* 677 C to T polymorphism (rs1801133 TT frequency 36%) and is known to be at high risk for neural tube defect-affected pregnancies. Although each of the one-carbon biomarkers included in this study are significantly correlated with each other (r2 range 0.11–0.33; inverse for homocysteine range −0.25 to −0.28; all p <0.001) each has the potential to be differentially impacted by genetic variation and co-factors in the one-carbon pathway. Folate deficiency (plasma folate < 7>7;mol/L) was associated with two polymorphisms in *MTHFR* (rs1801133, rs535107 both chi square p < 0.001). Red blood cell folate deficiency (<317 >7;mol/L) was associated with variation in *MTHFR* (rs1801133 p<0.001) as well as *TCN2* (rs1801198p<0.01). B12 deficiency (<148 >7;mol/L) was associated with variation in *MTHFR* (rs1801133p<0.01) and marginally with TCN2 (rs1801198 p = 0.05). High levels of homocysteine (>10.4 >7;mol/L) were associated both *MTHFR* variants (rs1801133, rs535107; both p<0.001) as well as *CBS* (rs6586282; p<0.03). Individuals with 3 or more (vs. 2 or fewer) homozygous non-ancestral genotypes were significantly more likely to have high homocysteine (OR 3.7 95% CI 2.5–5.3) and have plasma folate deficiency (OR 1.6 95% CI 1.2–2.2). The *MTHFR* 677 C to T polymorphism critically affects each of the analyzed one-carbon metabolism biomarkers, which is not surprising given its unique unidirectional role in one-carbon metabolism; however, other variants had differential impacts on the clinical biomarkers assessed and should be taken into consideration when modeling one-carbon metabolism and disease risk.

1305T

Disease associated genotype and allele sharing among human populations from throughout the world. *M.M. DeAngelis[1], M.A. Morrison[1], D.J. Morgan[1], K. Mayne[1], R. Robinson[1], G. Silvestri[2], D.A. Schaumberg[3], E.E. Tsironi[4], I.K. Kim[5], J. Ramke[6], K.H. Park[7], L.A. Farrer[8].* 1) Assoc Prof, Ophthalmology & Visual Sci, Univ Utah, Salt Lake City, UT, USA; 2) Centre for Vision and Vascular Science, Queen's University, Belfast, UK; 3) Division of Preventive Medicine, Brigham and Women's Hospital and the Department of Ophthalmology, Harvard Medical School, Boston, MA, USA; 4) Department of Ophthalmology, University of Thessaly School of Medicine, Larissa, Greece; 5) Retina Service, Harvard Medical School, Massachusetts Eye and Ear, Boston, MA, USA; 6) The Fred Hollows Foundation New Zealand, Private Bag 99909, Newmarket, Auckland 1023, New Zealand and Population Health Eye Research Network, Auckland, New Zealand; 7) Department of Ophthalmology, Seoul National University Bundang Hospital, Seoul, Korea; 8) Departments of Medicine (Section of Biomedical Genetics), Ophthalmology and Biostatistics, Neurology, Epidemiology, Boston University Schools of Medicine and Public Health, Boston, MA, USA.

Age-related macular degeneration (AMD) is the most common cause of legal blindness in the United States. Significant work in AMD genetics has established alleles and haplotypes having large influences on AMD risk. While these findings have begun to provide insights into AMD etiology, there are still many questions about AMD risk that cannot be explained by the known AMD-related genes. For this reason, we genotyped 17 previously reported AMD-associated SNPs to examine the allele frequencies in populations of different ethnicities. Six SNPs from the CFH region (1q32), six SNPs from ARMS2/HTRA1 (10q26), three SNPs from ROBO1 (3p12) and two RORA (15q22) SNPs were genotyped in four populations from throughout the world that are well characterized and phenotyped for AMD status: a family-based cohort from New England (n=657), a case-control cohort from Central Greece (n=436), a case-control cohort recruited from Seoul National University Bundang Hospital (n=818), and a population-based cohort from Timor-Leste (n=535), a population with low documented frequency of AMD. Genotype and allele frequencies were calculated for each cohort. For normal subjects (average age > 55 years), genotype and allele frequencies were similar between all ethnicities for the ROBO1 and RORA SNPs, SNPs previously shown to have a modest effect on AMD risk. Frequencies were also similar at 10q26, the region most significantly associated overall with AMD risk, for the SNPs rs10490924, rs10664316, rs2672598, rs1049331 and rs2293870. Interestingly, the risk allele, A, of the HTRA1 promoter SNP rs11200638 was found at a higher frequency in the Timorese (63.5%) than in the other groups examined (23–40%). Of the six CFH region SNPs examined (rs800292, rs16840422, rs1061170, rs12144939, rs2284664 and rs12066959), 3 of the 5 SNPs previously shown to have a protective effect on AMD were found at a higher frequency in the Timor cohort. Specifically, rs800292, rs2284664, and rs12066959 were shown to have nearly three times the frequency in the Timorese than in the Korean or Caucasian populations. Additionally, the most commonly associated AMD-risk SNP, CFH rs1061170 (Y402H), was also seen at a much lower frequency in the Korean and Timor populations than in the Caucasian populations (C ~7% vs. ~40%, respectively). The low prevalence of AMD in the Timor population (2 subjects out of 535) may be due to the enrichment of protective alleles in this population.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

† Trainee Award Finalist

1306F

Hidden heritability and risk prediction based on genome-wide association studies. *N. Chatterjee[1], B. Wheeler[2], J. Sampson[1], P. Hartge[1], S. Chanock[1], J. Park[1].* 1) National Institute of Health, Rockville, MD, USA; 2) Information management system, Rockville, MD.

Known discoveries from genome-wide association studies have limited predictive ability for individual traits, but recent estimates of "hidden heritability" suggest that in the future performance of predictive models can be potentially enhanced by incorporation of a large number of SNPs each with individually small effects. We use a novel theoretical model, discoveries from the largest genome-wide association studies and recent estimates of hidden heritability to project the predictive performance of polygenic models for ten complex traits as a function of the number and distribution of effect sizes for the underlying susceptibility SNPs, the sample size of the training dataset and the balance of true and false positives associated with the SNP selection criterion. We project, for example, that while 45% of the total variance of adult height has been attributed to common variants, a predictive model built based on as many as one million people may only explain 33.4% of variance of the trait in an independent sample. For rare highly familial conditions, such as Type 1 diabetes and Crohn's disease, risk models including family history and optimal polygenic scores built based on current GWAS can identify a large proportion (e.g 80–90%) of cases by targeting a small group of high-risk individuals (e.g subjects with top 20% risk). In contrast, for more common conditions with modest familial components, such as Type 2 diabetes (T2D), coronary heart disease (CAD) and prostate cancer (PrCA), risk models built based on GWAS with current or foreseeable sample sizes (e.g triple in size) can miss large proportion (>50%) of cases by targeting a small group of high-risk individuals. For these common disease, the proportion of the population that can be identified to have 2-fold or higher risk than an average person in the population ranged between 1.1% (CAD) and 7.0% (PrCA) for polygenic models built based on current GWAS. If the sample size for future studies could be tripled, these proportions could range between 6.1% (CAD) and 18.8% (T2D). Our analyses suggest that the predictive utility of polygenic models depends not only on heritability, but also on achievable sample sizes, effect-size distribution and information on other risk-factors, including family history.

1307W

Obesity susceptibility loci and associations across the pediatric body mass index distribution. *S.F.A. Grant[1,2], H. Hakonarson[1,2], T.R. Rebbeck[3], J.A. Mitchell[3].* 1) Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, PA; 2) Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA; 3) Center for Genetics and Complex Traits, Department of Biostatistics and Epidemiology, University of Pennsylvania, Philadelphia, PA.

Most studies investigating childhood obesity have leveraged mean BMI z-score to define the trait. We aimed to determine if previously identified obesity susceptibility loci are associated uniformly with pediatric BMI across the BMI distribution. Children and adolescents were recruited from the Children's Hospital of Philadelphia (n=7,225). In this population *FTO* (rs3751812), *MC4R* (rs12970134), *TMEM18* (rs2867125), *BDNF* (rs6265), *TNNI3K* (rs1514175), *NRXN3* (rs10146997), *SEC16B* (rs10913469), and *GNPDA2* (rs13130484) have been previously associated with BMI using linear and logistic regression. In the present study these associations were re-assessed using quantile regression. BMI z-score was modeled as the dependent variable, and genotype risk score (sum of risk alleles carried at the 8 loci) was modeled as the independent variable. Each additional increase in genotype risk score was associated with an increase in BMI z-score at the 5th, 15th, 25th, 50th, 75th, 85th and 95th BMI z-score percentiles by 0.04 (+/−0.02, P=0.08), 0.07 (+/−0.01, P=9.58 × 10^{-7}), 0.07 (+/−0.01, P=1.10 × 10^{-8}), 0.09 (+/−0.01, P=3.13 × 10–22), 0.11 (+/−0.01, P=1.35 × 10^{-25}), 0.11 (+/−0.01, P=1.98 × 10^{-20}), and 0.06 (+/−0.01, P=2.44 × 10^{-6}), respectively. Each additional increase in genotype risk score was associated with an increase in mean BMI z-score by 0.08 (+/−0.01, P=4.27 × 10^{-20}). In conclusion, obesity risk alleles were more strongly associated with increases in pediatric BMI z-score at the upper tail compared to the lower tail of the distribution. As such, previous studies may have underestimated the strength of the association between these obesity-susceptibility loci and pediatric BMI in the context of obesity.

1308T

Large-scale genome-wide association meta-analysis using imputation from 2188-haplotype 1000 Genomes reference panel identifies five novel susceptibility loci for BMI and additional novel sex-specific loci for BMI and WHR. *M. Horikoshi[1,2], R. Mägi[3], I. Surakka[4,5], S. Wiltshire[1,2], A. Sarin[4,5], A. Mahajan[1], L. Marullo[1,6], T. Ferreira[1], S. Hägg[7], JS. Ried[8], T. Winkler[9], G. Thorleifsson[10], N. Tsernikova[3], T. Esko[3,11], C. Willenborg[12,13], CP. Nelson[14,15], M. Beekman[16,17], SM. Willems[18], AP. Morris[1], CM. Lindgren[1], MI. McCarthy[1,2,19], S. Ripatti[4,5,20], I. Prokopenko[1,2] on behalf of ENGAGE Consortium.* 1) Wellcome Trust Centre for Human Genetics, university of Oxford, Oxford, UK; 2) Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford, UK; 3) Estonian Genome Center, University of Tartu, Tartu, Estonia; 4) Institute for Molecular Medicine Finland FIMM, University of Helsinki, Helsinki, Finland; 5) National Institute for Health and Welfare, Helsinki, Finland; 6) Department of Evolutionary Biology, Genetic Section, University of Ferrara, Ferrara, Italy; 7) Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden; 8) Institute of Genetic Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany; 9) Public Health and Gender Studies, Institute of Epidemiology and Preventive Medicine, Regensburg University Medical Center, Regensburg, Germany; 10) deCODE Genetics, Reykjavik, Iceland; 11) The Institute of Molecular and Cellular Biology of the University of Tartu, Tartu, Estonia; 12) Universität zu Lübeck, Med. Klinik II, Lübeck, Germany; 13) Deutsches Zentrum für Herz-Kreislauf-Forschung e. V. (DZHK), Lübeck, Germany; 14) Department of Cardiovascular Sciences, University of Leicester, Leicester, UK; 15) National Institute for Health Research (NIHR) Leicester Cardiovascular Disease Biomedical Research Unit, Glenfield Hospital, Leicester, UK; 16) Department of Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands; 17) Netherlands Consortium for Healthy Ageing, Leiden, The Netherlands; 18) Department of Genetic Epidemiology, Erasmus MC, Rotterdam, The Netherlands; 19) Oxford NIHR Biomedical Research Centre, Churchill Hospital, Oxford, UK; 20) Wellcome Trust Sanger Institute, Hinxton, UK.

Genome-wide association studies (GWAS) have succeeded in detecting numerous loci associated with BMI and waist-hip-ratio (WHR). To find additional loci and to identify potential causal variants by fine-mapping already known loci, we performed imputation using the 1000 Genomes panel with 2188 haplotypes (June 2011 release) in 18 European cohorts with GWA data from ENGAGE, followed by fixed-effects inverse variance meta-analysis in 78,131 (44,301) individuals for BMI (and WHR adjusted for BMI) respectively. Imputation with the larger 1000G reference data set revealed several features not observed in earlier rounds of GWA meta-analysis with HapMap2 imputation. These include: (a) genome-wide significant associations for BMI at 19 loci, of which 5 were novel including *GALNT10* (p=7.8×10^{-10}), *ZHX2* (p=4.5×10^{-8}), *AKAP6* (p=7.7×10^{-9}) and *PAX2-HIF1AN* (p=2.0×10^{-8}); (b) lead SNPs with stronger associations and larger effect sizes than previous signal, e.g. *RSPO3* for WHR (new β=0.12, p=4.4×10^{-14}, minor allele frequency (MAF)=0.08 vs old β=0.039, p=1.0×10^{-8}, MAF=0.49) and at *FAIM2*, *MAP2K5* and *NRXN3* for BMI; (c) novel associations with low MAF variants in the vicinity of 14 known BMI loci, with larger effect sizes than previous GWAS hits, for which conditional analyses will help to define whether these represent independent signals or are driving the GWA signal (e.g. *BDNF* β=0.21, p=9.0×10^{-4}, MAF=0.018 vs lead β=0.048, p=5.1×10^{-11}, MAF=0.18, CEU r^2=0.001); (d) four additional previously unreported association signals from sex-differentiated analysis including *TMEM56* for BMI (female-specific, p=1.1×10^{-8}, MAF=0.08, $P_{sex\text{-}het}$=0.0050) and *COL11A1* for WHR (female-specific, p=4.5×10^{-8}, MAF=0.19, $P_{sex\text{-}het}$=2.5×10^{-4}); (e) evidence of substantial overlap in trait associations: of 20 loci significantly associated with either BMI or WHR, 8 were associated with HDL cholesterol and 5 with triglyceride levels within parallel ENGAGE analyses. Our results highlight the potential of using existing GWAS genotyping data, supplemented by imputation from the high-density reference panel of the 1000 Genomes Project for the identification of associations with anthropometric traits.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1309F

A Two-Stage Genome-Wide Association Study to Identify Single Nucleotide Polymorphisms Associated with Development of Erectile Dysfunction Following Radiotherapy for Prostate Cancer. *S. Kerns[1,4], R. Stock[1], N. Stone[1,2], M. Buckstein[1], Y. Shao[3], C. Campbell[4], L. Rath[1], R. Hixson[5], J. Cesaretti[5], M. Terk[5], H. Ostrer[4], B. Rosenstein[1,6,7], *Collaboration developed under the framework of the Radiogenomics Consortium.* 1) Department of Radiation Oncology, Mount Sinai School of Medicine, New York, NY; 2) Department of Urology, Mount Sinai School of Medicine, New York, NY; 3) Division of Biostatistics, New York University School of Medicine, New York, NY; 4) Departments of Pathology and Genetics, Albert Einstein College of Medicine, Bronx, NY; 5) Florida Radiation Oncology Group, Jacksonville, FL, United States; 6) Department of Radiation Oncology, New York University School of Medicine, New York, NY; 7) Departments of Dermatology and Preventive Medicine, Mount Sinai School of Medicine, New York, NY.

Purpose: To identify single nucleotide polymorphisms (SNPs) predictive of erectile dysfunction (ED) among prostate cancer patients treated with radiotherapy. Methods and Materials: A two-stage genome-wide association study (GWAS) was performed to identify SNPs associated with development of ED following radiation therapy. Patients were split randomly into a stage I discovery cohort (132 cases, 103 controls) and a stage II replication cohort (128 cases, 102 controls). The discovery cohort was genotyped using Affymetrix 6.0 genome-wide arrays and, after QC filtering, association between 614,453 SNPs and radiation proctitis was assessed using logistic regression controlling for non-genetic risk factors and the first five principle components from principle components analysis (to control for potential population stratification). The 940 top ranking SNPs selected from the discovery cohort were genotyped in the replication cohort using Illumina iSelect custom SNP arrays, and association with ED was similarly assessed using multivariable logistic regression. Fisher's method was used to combine p-values from the discovery and replication studies. Results: 12 SNPs identified in the discovery cohort and validated in the replication cohort were associated with development of ED following radiotherapy (Fisher combined p-values 2.1×10^{-5} to 6.2×10^{-4}). In a multivariable model including non-genetic risk factors, the odds ratios for these SNPs ranged from 1.6 to 5.6 in the pooled cohort. There was a striking relationship between the cumulative number of SNP risk alleles an individual possessed and ED status (Sommers' D p-value = 1.7×10^{-29}). A one-allele increase in cumulative SNP score increased the odds for developing ED by a factor of 2.2 (p-value = 2.1×10^{-19}). The cumulative SNP score model had a sensitivity of 84% and specificity of 75% for prediction of developing ED at the radiotherapy planning stage. Conclusions: This GWAS identified a set of SNPs that have utility for predicting which men are likely to develop ED following radiotherapy. Notably, these 12 SNPs lie in or near genes involved in erectile function or other normal cellular functions (adhesion and signaling) rather than DNA damage repair, the focus of earlier candidate gene studies.

1310W

Genetic variation and vitamin D sufficiency in the U.S. population (NHANES III). *C. O. S. Neal, J. M. Jackson, K. S. Crider.* Pediatric Genetics Team/NCBDDD, Centers for Diseases Control and Prevention, Atlanta, GA.

Studies have linked vitamin D (VitD) insufficiency to adverse outcomes including skeletal, cardiovascular, autoimmune and metabolic disease. Over 20% of the US population are VitD insufficient (serum 25-hydroxyvitamin D below 20ng/mL - Institute Of Medicine guidelines) and significant variation in VitD sufficiency exists between subpopulations. The National Health and Nutrition Examination Surveys (NHANES) are population based studies designed to assess the health and nutritional status of US residents using interviews and physical examinations. NHANES III (1988–1994) contains single nucleotide polymorphism (SNP) data for 7,159 participants, age 12 years and older. Among NHANES III participants, VitD insufficiency was present in 12% of non-Hispanic whites (NHW), nearly 63% of non-Hispanic blacks (NHB) and 27% of Mexican Americans. Differences in VitD intake did not sufficiently account for the variation in serum levels of VitD. VitD metabolism involves multiple organ systems and myriad enzyme and transcriptional components, any of which might affect serum VitD levels. The Cytochrome p450 (CYP) family of approximately 60 genes is known to function in the synthesis and metabolism of various lipid soluble cellular molecules and chemicals, including VitD. Among genotyped NHANES III participants we explored genetic variation in CYP (33 SNPs in 9 genes) and VitD receptor genes (2 SNPs) in addition to covariates linked to sufficiency in previous studies such as season of sample collection (SSC), body-mass index (BMI), sex, supplementation habit, income, and age for associations with VitD sufficiency. We used chi square tests and multiple logistic regression in stratified [by racial/ethnic (RE) group] and single stratum models to assess associations with VitD sufficiency. Significant associations were found between VitD sufficiency and SSC, BMI, sex and age across RE strata. Several CYP SNPs were associated with VitD sufficiency in overall models. CYP2A6 (rs1801272) was significantly associated ($p<0.05$) with VitD sufficiency in multiple RE groups in both crude and adjusted models [odds ratio (95% confidence interval)=1.9(1.03–3.53) in NHW and 4.24(1.7–10.6) in NHB]. Although this is the first report of CYP2A6 association with VitD sufficiency, there is biological plausibility because of its wide range of potential metabolic targets. The role of the CYP family of genes and CYP2A6 in particular in VitD metabolism and sufficiency warrants further investigation.

1311T

The Genome-Wide Association Study of Phenotypic Robustness in Human-a Canalization Study. *R. Li[1], T.D. Spector[2], J.B. Richards[1,2].* 1) Lady Davis Institute, Jewish General Hospital, Montreal, Quebec, Canada; 2) Twin Research and Genetic Epidemiology, King's College London, London, United Kingdom.

Human phenotypes exhibit significant levels of variation among individuals. Recently, genome-wide association studies (GWAS) have been performed to identify the genetic loci that control these variations. Despite the success of GWAS to identify a set of genetic loci, it has become widely recognized that these loci cannot fully explain the variance of a phenotype. One possible explaination is that there is phenotypic stochastic noise which has not been taken into account in typical GWASs. Normally, only one measurement of phenotype from each individual has been used in the typical GWAS which can introduce significant stochasticity, since many phenotypes, such as glucose show high intra-individual variability. However, if a phenotype is measured several times in the same individual, there is also considerable variation in the intra-individual stability of the phenotypic measurement, such that in some people the phenotype is relatively stable while in other people the phenotype is relatively variable. In other words, people show different levels of phenotypic robustness (canalization). The genetic control on phenotypic robustness has been reported in plants, yeast and mice. However, there is still scarce data in human. We hypothesized that genetic components controlling phenotype robustness may exist in humans. Since typical GWASs focus on the genetic variation impact upon average phenotypes, the findings from genetic control on phenotype stochastic noise would improve our understanding on the genotype-phenotype relationship. In TwinsUK cohort, we observed different levels of phenotypic robustness on glucose across multiple measures on each individual. We undertook GWA analysis on repeated glucose levels in 2,594 individuals from TwinsUK cohort, and we observed a trend of association on several genetic loci across the genome ($P < 5 \times 10e{-}6$). The identified loci are at chromosome 1q43 (rs12568231, *PLD5*), choromosome 6p21 (rs16871837, *SPATS1*), chromosome 6q24.2 (rs722755, *PHACTR2*), chromosome 7q22.1 (rs12539614, *EMID2*), chromosome 13q14 (rs7339165, *GTF2F2*), and chromosome 18q11.2 (rs17593321, *IMPA2*). While these loci are being replicated currently this studies identifies potential genetic determinants of canalization in humans.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1312F

A Genome- and Phenome-Wide Association Study to Identify Genetic Variants Influencing Platelet Count and Volume and their Pleiotropic Effects. *K. Shameer[1], J.C Denny[2], K. Ding[1], D.R. Crosslin[3], C.G Chute[4], P. Peissig[5], J. Pacheco[6], R. Li[7], M.de Andrade[4], M.D. Ritchie[8], D.R. Masys[9], R.L. Chisholm[6], E.B. Larson[10], C.A. McCarty[11], D.M. Roden[12], G.P. Jarvik[13], I.J. Kullo[1].* 1) Division of Cardiovascular Diseases, Mayo Clinic, Rochester, MN. 55905, USA; 2) Departments of Medicine and Biomedical Informatics, Vanderbilt University, Nashville, TN 37232, USA; 3) Department of Biostatistics, University of Washington, Seattle, WA 98195, USA; 4) Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, MN; 5) Biomedical Informatics Research Center, Marshfield Clinic, Marshfield, WI, 54449, USA; 6) Feinberg School of Medicine, Northwestern University, Chicago, IL 60611, USA; 7) Office of Population Genomics, National Human Genome Research Institute, 5635 Fishers Lane, Suite 3058, MSC 9307, Bethesda, MD, 20892, USA; 8) Center for Systems Genomics, Pennsylvania State University, Eberly College of Science, The Huck Institutes of the Life Sciences, 512 Wartik Laboratory, University Park, PA 16802 USA; 9) Department of Biomedical Informatics, Vanderbilt University School of Medicine, Room 416 Eskind Medical Library, Nashville, TN, 37232, USA; 10) Group Health Research Institute, 1730 Minor Avenue, Suite 1600, Seattle, WA, 98101, USA; 11) Center for Human Genetics, Marshfield Clinic Research Foundation, 1000 North Oak Avenue (MLR), Marshfield, WI, 54449, USA; 12) Center for Genetic Medicine, Northwestern University, Chicago, IL 60611, USA; 13) Department of Genome Sciences, University of Washington, 3720 15th Ave NE, Seattle WA 98195, USA.

Platelets are enucleated cell fragments derived from megakaryocytes that play key roles in hemostasis and in the pathogenesis of atherothrombosis and cancer. We conducted an electronic medical record (EMR)-based study to: 1) identify the variants that influence inter-individual variation in the number of circulating platelets (PLT) and mean platelet volume (MPV) by performing a genome-wide association study (GWAS) and 2) assess pleiotropic effects of such variants by performing a phenome-wide association study (PheWAS) with a wide range of EMR derived phenotypes. The study was performed in 13,582 participants of European ancestry in the electronic MEdical Records and GEnomic (eMERGE) network. We identified five chromosome regions associated with PLT and eight associated with MPV respectively ($P < 5 \times 10^{-8}$). In addition, we replicated 20 SNPs influencing PLT (out of 56 SNPs (α: 0.05/56 = 9E−04)) and 22 SNPs influencing (out of 29 SNPs (α: 0.05/29 = 2E-03)) MPV from among 69 SNPs identified in a recent meta-analysis of PLT and MPV. Genes in these regions influence thrombopoiesis and encode kinases (*AK3, DSTYK, PIK3CG, TAOK1*), membrane proteins (*BET1L, COPZ1, REEP3, TMCC2, TMEM81, TMEM120B*), proteins involved in cellular trafficking (*DNM3, NRBF2, REEP3*), transcription factors (*JMJD1C, MYB, NFE2, NRBF2, RBBP5*), proteasome complex subunit (*PSMD13*), proteins involved in signal transduction pathways (*ARHGEF3, SH2B3*), proteins involved in megakaryocyte development and platelet production (*AK3, CABLES1, EHD3, JMJD1C, MYB, NFE2, SH2B3*) and hemostasis (*SH2B3, JMJD1C, MYB, EHD3, PIK3CG, F2R, SIRPA, PTPN11, CABLES1, AK3, THPO, NFE2*). PheWAS analysis using a single-SNP Bonferroni correction (0.05/1368 = 3.6E−5) revealed that several variants in these genes have pleiotropic effects and are associated with diverse phenotypes including autoimmune diseases (inflammatory spondylopathies; P = 5.7E−08), hematologic disorders(B12-deficiency anemia; P = 6.1E−05), hypothyroidism(P = 1.7E−06) and myocardial infarction (P = 1.7E−04). Our results demonstrate the utility of EMR-based GWAS and PheWAS to identify genetic variants influencing traits of medical relevance and the pleiotropic effects of these variants.

1313W

Heterogeneity in Polygenic Contribution to Ischaemic Stroke subtypes by Age at Onset. *M. Traylor[1], HS. Markus[1], S. Bevan[1], M. Dichgans[2], R. Malik[2], H. Segal[3], PM. Rothwell[3], CL. Sudlow[4], CM. Lewis[5], Wellcome Trust Case Control Consortium (WTCCC2).* 1) Research Centre for Stroke & Dementia, St. George's University of London, London, United Kingdom; 2) Institute for Stroke and Dementia Research, Klinikum der Universität München, Ludwig-Maximilians-Universität, Munich, Germany; 3) Stroke Prevention Research Unit, University Department of Clinical Neurology, Oxford, UK; 4) Division of Clinical Neurosciences, University of Edinburgh, Edinburgh, UK; 5) Medical & Molecular Genetics, King's College London, UK.

Stroke is one of the three most common causes of death, the second most common cause of dementia and the major cause of adult chronic disability. It is a complex phenotype involving many pathological processes, and is generally classified into three subphenotypes: cardioembolic (CE), large vessel (LVD) and small vessel disease (SVD). Genome-wide association analysis of ischaemic stroke has identified a handful of variants in CE and LVD subtypes. We investigated the relationship between the polygenic contribution to disease and age-at-onset of stroke using the WTCCC2 ischaemic stroke cohort of 3,548 cases (790 CE, 844 LVD, 580 SVD) and 5,972 controls. We used the GCTA software to perform “sliding window” analyses of heritability within subsets of cases ranked by age-at-onset. We also explored the variance of heritability estimates within the subsets by bootstrap simulations, testing whether variance in age-at-onset-restricted subsets was equal to that in all cases, using Levene's test. Our heritability estimates for the subtypes using the GCTA software showed that a moderate proportion of variance can be explained by the common variants on GWAS chips for the CE (32.6%) and LVD (40.3%) subtypes, but a smaller proportion of the variance (16.1%) is explained for SVD. When considering age at onset, the proportion of variance explained was greatly reduced for subsets of older individuals in the CE subtype (16.5%), consistent with a stronger environmental contribution to disease for older individuals. Conversely, in SVD we found greatly reduced heritability estimates for younger subsets of the cases (5.3%). This may indicate a Mendelian inheritance for a subset of young onset SVD stroke cases. We found strong evidence in all subtypes of greater variance in the age-at-onset subsets of cases (compared to bootstrap samples from all cases), indicating significantly reduced heterogeneity in the polygenic component to disease when cases are stratified by age-at-onset. This was observed for all subtypes (p<10-3), with considerably stronger evidence in SVD (p<10-6). The results indicate significant heterogeneity in polygenic contribution to the ischaemic stroke phenotypes by age at onset, particularly for SVD. They may be used to guide case inclusion for further GWAS analysis and aid selection for exome sequencing.

1314T

Strategies for developing prediction models with increased discriminatory accuracy from genome-wide association studies. *J. Wu, R.M. Pfeiffer, M.H. Gail.* Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, MD.

Recent genome-wide association studies (GWAS) have identified hundreds of single nucleotide polymorphisms (SNPs) associated with complex human diseases. However, in most diseases, these risk loci have limited predictive value. Larger GWAS and improved technology will lead to the discovery of additional risk-associated SNPs, but it is likely they will have smaller odds ratios than those previously detected. It is possible that predictive value can be improved by including many such SNPs in models. Here we investigate various aspects of model building to improve prediction, as measured by the area under the receiving operating characteristic curve (AUC). In particular, we investigate: (1) how the training data should be allocated between SNP selection and estimation of genetic effect sizes; (2) how SNP selection based on p-value thresholding compares with selection based on ranking p-values, and how liberal one should be in setting thresholds or the number of top-ranked SNPs; (3) whether multivariate estimation of SNP effects is preferred to univariate estimation in the presence of linkage disequilibrium (LD); (4) whether design guidelines developed in the absence of LD also apply in the presence of LD; and (5) whether our findings for AUC hold for other figures of merit, such as the probability of correct classification. We used realistic estimates of the distributions of genetic effects and allele frequencies based on GWAS data, and we included analyses in the presence of LD by re-sampling chromosomes from GWAS controls. Broadly speaking, we found that much of the training data should be used for SNP selection, that univariate (marginal) estimation performs as well or better than multivariate estimation, and that one should not include too many SNPs in the prediction model.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1315F

Concurrent Modeling of Multiple Phenotypes and Genotypes in Family Data: Implications for Cleft Gene Discovery. *M. Govil[1], N. Mukhopadhyay[1], R.M. Silva[2], A.R. Vieira[1], S.M. Weinberg[1], K. Neiswanger[1], M.L. Marazita[1].* 1) Center for Craniofacial and Dental Genetics, Department of Oral Biology, School of Dental Medicine, University of Pittsburgh, PA; 2) Department of Endodontics, School of Dentistry, University of Texas Health Science Center at Houston.

Non-syndromic clefting, one of the most common birth defects worldwide, has been demonstrated to be a heterogeneous trait with a multifactorial etiology. Substantial work has been done in the past several years to further characterize this birth anomaly by exploring sub-clinical and extended phenotypes. Genetic studies accumulating evidence for the role of genes in cleft etiology have primarily incorporated this additional phenotypic information in analyses by applying three different approaches: creating potentially homogenous subsets of the data for categorical values of the phenotype(s), recoding the affection status of individuals based on their cumulative phenotypic information, or analyzing each phenotype individually. Here, we include multiple phenotypes concurrently as predictors while regressing them against individual SNP genotypes in a candidate gene inverse regression analysis, as implemented in the R package MultiPhen (O'Reilly et. al., 2011), utilizing data from 79 families collected as a part of the Pittsburgh Orofacial Cleft Project. In the current study, 43 probands with clefts of the lip and/or palate have been selected such that each individual has complete information for three additional phenotypes — lip muscle discontinuities (binary, 21 affected), velopharyngeal insufficiency (binary, 33 affected), and handedness preference (measured quantitatively by a laterality quotient, mean = 59.5). Further, all 43 subjects have been genotyped for a candidate gene panel with 1040 SNPs. Preliminary results from this analysis provide potential evidence for association (p-value = 0.0007) with SNPs in the SKI gene (1p36.33), a component in the TGFB1 signaling pathway. To provide a contrast, we added 78 unaffected individuals, one randomly selected from each family. The top hit (p-value = 0.0004) was an association with GABABR2 (9q22.33), a gene associated with smoking behavior, where maternal smoking is known to be a risk factor for clefts. These initial results indicate the utility of the method to not only model potential inter-relationships between the phenotypes but to also provide evidence for genes which may be involved in complex pathways influencing multiple phenotypes. We continue to expand and refine the approach to incorporate existing family data from multiple populations, genotyped on whole genome linkage and GWAS panels, thereby further increasing the power of this method. Support: DE018085, DE016148.

1316W

A Phenome-Wide Association Study (PheWAS) using multiple National Health and Nutrition Examination Surveys (NHANES) to identify pleiotropy. *M.A. Hall[1], A. Verma[1], K.D. Brown-Gentry[2], R. Goodloe[2], J. Boston[2], S. Wilson[2], B. McClellan[2], C. Sutcliffe[2], H.H. Dilks[2,3], N.B. Gillani[2], H. Jin[2], P. Mayo[2], M. Allen[2], N. Schnetz-Boutaud[2], S.A. Pendergrass[1], D.C. Crawford[2,3], M.D. Ritchie[1].* 1) Center for Systems Genomics, Department of Biochemistry and Molecular Biology, The Huck Institutes of the Life Sciences, The Pennsylvania State University, University Park, PA, USA; 2) Center for Human Genetics Research, Vanderbilt University, Nashville, TN, USA; 3) Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, TN.

Phenome-wide association studies (PheWAS) are a novel approach for discovering the complex mechanisms involved in human disease, through testing SNPs for association with a large and extensive set of phenotypes. The Epidemiological Architecture for Genes Linked to Environment (EAGLE) project uses the diverse genotypic and phenotypic data across race-ethnicity available from the National Health and Nutrition Examination Surveys (NHANES) conducted by the Centers for Disease Control and Prevention (CDC). Three NHANES surveys were used for our PheWAS: NHANES III phase 2 (1991–1994) and continuous NHANES (1999–2000 and 2001–2002). Comprehensive unadjusted tests of association were performed for 575 SNPs with 1009 phenotypes, stratified by race-ethnicity. We found 232 PheWAS results showing replication across the two surveys for the same SNP, phenotype-class, and race-ethnicity with $p < 0.01$ (allele freq > 0.01, sample size range n: 200–3974). Of these PheWAS associations, 20 replicated reported SNP-phenotype associations, 44 were related to reported associations, and 168 were novel associations. An intriguing novel result was for intronic FADS1 rs174547, previously associated with HDL cholesterol levels, associated here with total iron binding capacity (TIBC) in both non-Hispanic whites (W) and Mexican-Americans (MA). The p-values for ln(TIBC) in combined NHANES datasets were $p=1.91\times10\text{-}7$ in W ($\beta = 0.016$; n=6,368) and p=0.00021 in MA ($\beta = 0.015$; n=3,685). A total of 36 SNPs were associated with multiple phenotypes, indicative of pleiotropy. We found the missense SNP rs1260326 (within GCKR), associated across multiple classes in multiple race-ethnicities. Pleiotropic associations for rs1260326 include: for non-Hispanic blacks, the phenotype class of cell blood count (eosinophil %: $p=5.5\times10\text{-}4$; red blood cell count: p=0.007); for MA, the phenotype classes of cardiac/lipids (triglycerides: $p=4.05\times10\text{-}7$; QT interval: p=0.008), measures of nutrition (calories: p=0.009; thiamin: p=0.007), and diabetes (glucose: p=0.004; insulin: p=0.003); and for W, the phenotypes classes of diabetes (mother had diabetes: p=0.004; glucose: p=0.001) and serum vitamin levels (vitamin E: $p=9.14\times10\text{-}6$; vitamin B12: p=0.006). Previous associations for this SNP in W include phenotypes of platelet count, triglycerides, and glucose. These results demonstrate the utility of phenome-wide association studies for generating hypotheses about pleiotropic interactions for future research.

1317T

Genetic variants in pigmentation genes, skin color, and risk of skin cancer in Japanese. *T. Suzuki[1], Y. Abe[1], J. Yoshizawa[1], Y. Hozumi[1], T. Nakamura[2], G. Tamiya[2].* 1) Dept Dermatology, Yamagata Univ Sch Med, Yamagata, Japan; 2) Advanced Molecular Epidemiology Research Institute, Yamagata Univ Sch Med, Yamagata, Japan.

Melanin pigmentation plays an important role in shielding the body from ultraviolet (UV) radiation and may serve as a scavenger for reactive oxygen species. More than 150 genes have been implicated in determining in mice, and include transcription factors, membrane and structural proteins, enzymes, and several kinds of receptors and their legands, most of which have human orthologues. Although many molecular mechanisms involved in melanin pigmentation are being determined, relatively little is understood about the genetic component responsible for variations in skin color within or between human populations. First, in order to reveal their genetic contribution to skin color, we examined the association of pigmentation-related genes variants and variations in the melanin index in members of the general Japanese population whose skin color was objectively measured by reflectometry. The multiple regression showed that *OCA2* A481T rs74653330 (p = 6.18e-8) and, *OCA2* H615R rs1800414 (p = 5.72e-6) were strongly associated with the mean of the melanin index in the female population. Three variants (*SLC45A2* T500P rs11568737 p = 0.048, *OCA2* T387M p = 0.015, *TYR* D125Y rs13312741 p = 0.022) were also significantly associated with melanin index. However, no significant associations were found between age and melanin index for variants of *MC1R*. Second, we evaluated the associations of the pigmentation-related genes variants and the risk of skin cancer. The statistical analysis revealed that only *OCA2* H615R was associated with the risk of all skin cancers, especially malignant melanoma. We could not find any statistical significance in the associations of other variants, including *OCA2* A481T, or melanin index with the risk of skin cancer. This is the first report on the association between the genetic variants in pigmentation genes and the risk of skin cancer in East Asian population.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1318F

Validity of the Mendelian Randomization Approach: do the principles really hold? *M. Taylor, NJ. Timpson, JP. Kemp, B. St. Pourcain, DM. Evans, SM. Ring, DA. Lawlor, G. Davey Smith.* School of Social and Community Medicine, Bristol University, Oakfield House, Oakfield Grove, Bristol, BS8 2BN United Kingdom.

A clear focus of epidemiology is to attempt to establish causality. Mendelian randomization is an established approach, which uses genetic variants as instrumental variables to assess causality of associations between potentially modifiable risk factors and disease related outcomes. A key assumption of this approach is that genetic variants are independent of environmental characteristics and of other genetic variants that would result in a pathway between the variant being used as an instrumental variable to the disease outcome, other than through the risk factor of interest. One previous study found that 45% of pairwise associations between 96 unrelated non-genetic traits were statistically significant at the 1% level, whereas pairwise the expected 1% of pairwise associations between 23 independent SNPs and between these 23 SNPs and the 96 non-genetic traits were significant at 1%. The aim of this study was to further examine the Mendelian randomization assumption of independence by repeating the earlier study in a different dataset and with a much larger number of genetic variants. Patterns of association within a large collection of non-genetic variants between these and all single nucleotide polymorphisms with established genomewide association study (GWAS) results (Hindorff LA et al., www.genome.gov/gwasstudies) and with these genetic variants were assessed. Amongst 7260 possible pairwise associations between 121 phenotypic measures 1281 (18%) were statistically significant at a 1% level. From total of 13552 possible pairwise associations between 112 independent loci (pruned down from 4985 genetic variants) and phenotypic measures 167 (1%) were statistically significant at the 1% level. These results provide additional evidence supporting the assumption that genetic variants are largely unrelated to the large number of non-genetic characteristics or to each other. By contrast non-genetic characteristics are correlated with a wide-range of other non-genetic characteristics even when established associations have been removed from the characteristics tested. Mendelian randomization studies are therefore much less likely than conventional observational association approaches to be biased by residual confounding.

1319W

Polygenic modeling of healthy aging. *N.E. Wineinger[1], G. Atzmon[2], N. Barzilai[2], N.J. Schork[1].* 1) Scripps Translational Science Institute, La Jolla, CA; 2) Albert Einstein College of Medicine of Yeshiva University, Bronx, NY.

Past studies have shown that the manifestation of most common, heritable human diseases cannot be entirely attributed to one, or even a handful of segregating genetic variants. More recent statistical approaches which leverage numerous genetic markers, including entire genomes, have shown promising results. However, these methods have only been explored in a small subset of diseases and their clinical utility remains unclear. Healthy aging - that is, extended longevity and the absence of disease - is, perhaps, the ideal polygenic trait. One might expect that the genetic components contributing to this trait include the absence of risk variants associated with human diseases, the presence of beneficial variants specific to longevity, or some combination of both. We examined this in the Wellderly study, a cohort consisting of 397 healthy, elderly subjects and 386 matched controls. Wellderly cases had a median age of 87 and had remained free from major diseases and the need for long-term medications over the course of life. We found that 50.9% of the phenotypic variation could be explained by genetic variants genotyped on the Illumina 1M array and imputed from 1000 Genomes haplotypes, and that 18.4% of the genetic variation was due to variants recognized in the Catalog of Published Genome-Wide Association Studies or the Human Genome Mutation Database. We have begun attempts to replicate these findings in an advanced aging cohort from a dissimilar population (Ashkenazi Jews). Our preliminary results estimate 11.3% of the phenotypic variation can be explained by genotyped variants in this cohort, but we expect this figure to rise with the inclusion of imputed markers. We have explored alternative whole-genome approaches which allow for the estimation of individual marker effects and assessment of prediction within each cohort via cross-validation and between cohorts. Our findings highlight the utility of polygenic modeling, and the challenges when forming inference across populations.

1320T

Extended haplotype association study in Crohn's disease identifies a novel Ashkenazi Jewish-specific missense mutation in the *NF-kB* pathway gene, *HEATR3*. *K.Y. Hui[1], W. Zhang[1], A. Gusev[2], N. Warner[3], G. Nuñez[3], I. Pe'er[2], I. Peter[4], J.H. Cho[1].* 1) Departments of Internal Medicine and Genetics, Yale University School of Medicine, New Haven, CT; 2) Department of Computer Science, Columbia University, New York, NY; 3) Department of Pathology, University of Michigan School of Medicine, Ann Arbor, MI; 4) Department of Genetics and Genomic Sciences, Mount Sinai School of Medicine, New York, NY.

The Ashkenazi Jewish population has a unique genetic history, characterized by population bottlenecks, expansions, and endogamy, resulting in LD blocks of greater length and stronger correlation, compared to non-Jewish European ancestry populations. We present a study to uncover novel extended haplotypes associated to Crohn's disease and discover constituent genetic variants contributing to the several-fold higher prevalence of the disease in the Ashkenazim. Haplotype association has long played an important role in Crohn's disease in the Ashkenazi Jewish population, most notably at *NOD2*, in which three causal, uncommon, and conditionally independent variants (Arg702Trp, Gly908Arg, Leu1007fsinsC) reside on a shared background haplotype tagged by the common SNP rs2076756. In case-control cohorts of 828 Ashkenazi and 896 non-Jewish individuals, we detected shared haplotypes of length greater than 3 Mb, using the program GERMLINE; these showed significantly greater genome-wide association to Crohn's disease in the Ashkenazi Jewish population compared to a non-Jewish population (145 Jewish haplotypes with P-value < 10^{-3}; no significant non-Jewish haplotypes). Two haplotype regions, one each on chromosomes 16 and 21, conferred increased disease risk within established Crohn's disease loci. To detect novel variants within these regions, we performed exome sequencing of 55 Ashkenazi Jewish individuals, including carriers of these two haplotypes. In follow-up genotyping of chromosome 21 variants, we observed Ashkenazi Jewish-specific nominal association at R755C in *TRPM2*. Genotyping data within the chromosome 16 region showed genome-wide significance at rs9922362 of *BRD7* and R642S of *HEATR3*. We performed siRNA knockdown and over-expression studies of *HEATR3* demonstrating a positive role in *NOD2*-mediated *NF-κB* signaling. In logistical regression, the *BRD7* signal showed conditional dependence with only the rare causal variants in *NOD2*, but not with rs2076756, which elaborates this region as a key illustration of synthetic association's role in disease genetics. Notably, it illustrates conditional dependence of the cluster of *NOD2* functional polymorphisms on a secondary common background SNP across a span of 0.4 Mb.

1321F

Association analysis of CYP2A6 polymorphism in SIDS. *Y. Inaoka[1], M. Nakatome[2], F. Satoh[1], I. Hasegawa[1], M.Q. Fujita[3], M. Osawa[1].* 1) Department of Forensic Medicine, Tokai University School of Medicine, Isehara, Kanagawa 259-1193, Japan; 2) Department of Legal Medicine, Fujita Health University School of Medicine, Aichi 470-1192, Japan; 3) Department of Legal Medicine (Forensic Medicine), Keio University School of Medicine, Shinanomachi 35, Shinjuku-ku, Tokyo 160-8582, Japan.

The exposure to tobacco smoke is the major risk factor for the occurrence of sudden infant death syndrome (SIDS). A variety of components are present in the smoke, but nicotine is of particular concern in relation to SIDS because of penetrance in the placenta and effects on fetal development. The majority of intake nicotine is metabolized to the inactive forms, cotinine and 3'-hyroxycotinine by CYP2A6 in the liver. A large interindividual difference in the metabolism is evident, which is mainly attributed to CYP2A6 polymorphism. Three major alleles of *1A, *1B and *4, are present in the East Asia populations. In particular, CYP2A6*4 is the deficient allele with deletion of two exons, and clearly affects the nicotine metabolism. In this study, we performed an association study of CYP2A6 polymorphism with SIDS. A total of 71 specimens from SIDS subjects, diagnosed after full autopsy and subsequent examinations, were available. For the control, 225 samples of DNA from healthy adult individuals who had no relatedness were used. The present study has been approved by the institutional ethic committee. The genotype was determined by PCR-RFLP method to the DNA samples of SIDS and the control group. Allele frequencies of *1A, *1B and *4A were 0.436, 0.364 and 0.200, respectively. In the Japanese individuals, genotype frequencies were consistent with those expected under Hardy-Weinberg equilibrium ($\chi 2 = 4.40$, $P = 0.90$). Relative to the allele frequency of CYP2A6*1A, the odds ratio (OR) of *1B and *4A in the SIDS group was 1.20 (95% confidence intervals; 0.79–1.80) and 0.71 (95% CI; 0.41–1.25) respectively. In the homozygote-based analysis to *1A/*1A, the OR of *1B/*1B and *4A/*4A in healthy controls and SIDS samples were 1.50 (95% CI; 0.69–3.24) and 0.55 (95% CI; 0.11–2.78), respectively. These results suggest that CYP2A6 genetic polymorphism is not significant determinants affecting SIDS. We are planning to examine the effect onto sudden death of adult subjects.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1322W

Do severity of early lung disease and meconium ileus in cystic fibrosis have common genetic contributors? *W. Li[1, 2], D. Su[2], T. Chiang[2], X. Li[2], MR. Miller[2], K. Keenan[3], H. Corvol[4], FA. Wright[5], S. Blackman[6], ML. Drumm[7], GR. Cutting[6, 8], MR. Knowles[9], PR. Durie[3, 10], JM. Rommens[11, 12], L. Sun[1, 13], LJ. Strug[2, 1]*. 1) Biostatistics Division, Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada; 2) Program in Child Health Evaluative Sciences, Hospital for Sick Children, Toronto, Ontario, Canada; 3) Program in Physiology and Experimental Medicine, Research Institute, Hospital for Sick Children, Toronto, Ontario, Canada; 4) AP-HP, Hospital Trousseau, Pediatric Pulmonary Department, Paris, France; 5) Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA; 6) Department of Pediatrics, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA; 7) Departments of Pediatrics and Genetics, Case Western Reserve University, Cleveland, Ohio, USA; 8) McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA; 9) Cystic Fibrosis-Pulmonary Research and Treatment Center, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA; 10) Department of Pediatrics, University of Toronto, Toronto, Ontario, Canada; 11) Program in Genetics and Genome Biology, Hospital for Sick Children, Toronto, Ontario, Canada; 12) Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada; 13) Department of Statistics, University of Toronto, Toronto, Ontario, Canada.

Motivation We have established that common variation in multiple constituents of the apical plasma membrane are associated with meconium ileus (MI) in cystic fibrosis (CF) patients, as revealed by SNPs in *SLC26A9*, *SLC9A3*, and *SLC6A14* (Sun et al. 2012). MI is an intestinal obstruction seen in ~15% of CF patients at birth, and we sought to determine whether these 3 genes are pleiotropic for other CF traits, notably early CF lung disease, as all three genes are expressed in lung epithelia. Pleiotropy is a phenomenon whereby a single gene influences multiple phenotypic traits. Although conceptually straightforward, it is difficult to define and analyze pleiotropy statistically because, for example, the alternative hypothesis of interest is that the variant is associated with *all* phenotypes under study, in contrast to the conventional definition of *at least one*. **Methods** We used a cross-sectional measure of lung disease severity (SaKnorm, Taylor et al. 2011) that averages Forced Expiratory Volume in one second (FEV1) over a 3-year period of time, adjusting for age, sex and mortality. Unrelated individuals and sibling pairs from the North American CF Gene Modifier Consortium, for whom MI status was known and whose average age at FEV1 measurement was ≤ 18, were included for analysis (n=1,609). Four SNPs in *SLC26A9*, one in *SLC9A3* and three in *SLC6A14*, were analyzed for pleiotropy. To detect the pleiotropic effects of a SNP, we tested for its association with SaKnorm after accounting for its significant effect on MI. Specifically, we fit two nested mixed-effects proportional odds models at each SNP for which the ordinal response variable is the number of minor alleles of the SNP (0, 1 or 2), with a random intercept accounting for sibling correlations. The full model included both SaKnorm and MI, as well as top principal components and a recruitment site covariate; the reduced model omitted SaKnorm. We then used a likelihood ratio test to evaluate the statistical significance of the pleiotropic effects at the SNP. **Results** rs17563161 in *SLC9A3* had $p = 0.0005$ and rs3788766 in *SLC6A14* had $p = 0.02$. No SNPs in *SLC26A9* provided $p < 0.05$. **Conclusion** Evidence of pleiotropic effects for *SLC9A3* and *SLC6A14* were observed for two early CF traits, the increased risk of MI and decreased lung function in pediatric patients. The proportional odds and nested mixed modeling approach can assess evidence for pleiotropy among categorical and quantitative phenotypes in related samples.

1323T

Influence of Host Genetics and Environment on Nasal Carriage of Staphylococcus aureus in Danish Middle-Aged and Elderly Twins. *PS. Andersen[1], JK. Pedersen[2], P. Fode[1], RL. Skov[1], VG. Fowler JR[3], M. Stegger[1], K. Christensen[2]*. 1) Dept of Public Health Microbiology, Statens Serum Inst, Copenhagen, Denmark; 2) Danish Aging Research Center, Epidemiology, University of Southern Denmark, Odense, Denmark; 3) Department of Medicine, Duke University Medical Center, Durham, North Carolina, USA.

Background. Nasal carriage is a major risk factor for Staphylococcus aureus infection. Approximately, one-quarter of adults carry S. aureus. However, the role of host genetics on S. aureus nasal carriage is unknown. Methods. Nasal swabs were obtained from a national cohort of middle-aged and elderly Danish twins. Subjects colonized with S. aureus were identified by growth on selective plates and spa typing. A second sample was obtained from twins initially concordant for carriage. Twins found to again be colonized with S. aureus were defined as persistent carriers. Results. The prevalence of S. aureus carriage among 617 twin pairs (monozygotic/dizygotic pairs: 112/505) was 26.3% (95%-CI 32.2–29.9%). The concordance rate for carriage did not differ significantly between pairs of monozygotic (37.5%; 95%-CI 22.3–53.8), same sex (24.2%; 95%-CI 15.4–34.5%), and opposite sex (21.4%; 95% CI-12.0–33.4%) dizygotic twins. Despite shared childhood only 1/ 617 pairs was concordant with respect to lineage. Although heritability increased for S. aureus and lineage persistency, no significant heritability was detected. Conclusion. In this study, host genetic factors exhibited only a modest influence on the S. aureus carrier state of middle-aged and elderly individuals.

1324F

Polygenes and Estimated Heritability of Prostate Cancer in an African American Sample Using GWAS Data. *J. He, G. Chen, B. Henderson, C. Haiman, D. Stram, Genome-Wide Association Studies of Prostate Cancer in African Americans.* Preventive Medicine, University of Southern California, Los Angeles, CA.

Purpose: There is a gap between the phenotypic variance explained by genome-wide association study (GWAS) hits and those estimated from classical heritability methods. Here we utilize a more comprehensive evaluation of genetic variation trying to find the “missing” heritability in GWAS data. Methods: We adopted score analysis (Purcell et al) and linear mixed effects model (Yang et al) to estimate narrow sense heritability of prostate cancer using all SNPs in an African American Prostate Cancer (AAPC) GWAS simultaneously. In score analysis, two independent populations were used as discovery and target samples; each locus is tested in the discovery sample, and then the variation across nominally associated loci is summarized into quantitative scores that manifest a combined effect of many SNPs, and the scores are related to disease in the target sample. Linear mixed model is an alternative approach to fit all the GWAS SNPs where the effects of the SNPs are treated statistically as random, and the variance explained by all the SNPs together is estimated. In order to determine the sensitivity of both methods to long-range hidden relatedness, we simulated phenotypes and genotypes that have low pairwise correlation to see whether small correlations between individuals have a major effect on the estimate. Results: In the score analysis, both the statistical significance and phenotypic variance explained increased as a greater number of SNPs were considered; on a liability scale 29.5% of variance of AAPC can be explained by considering 873,644 SNPs simultaneously in linear mixed model. However, we concluded that the estimates from both methods may be biased upward due to unmeasured causal alleles and confounding effects of distant relatedness.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1325W

Using identity-by-descent information to detect *de novo* and recent mutations in population-based exome-sequencing studies. *N. Solovieff[1,2,3], M. Fromer[1,3,4], D. Ruderfer[4], J. Moran[3], K. Chambert[3], C. Hultman[5], P. Sullivan[6], H. Williams[7], E. Rees[7], P. Gormley[8], A. Palotie[8], G. Kirov[7], M. Owen[7], M. O'Donovan[7], P. Sklar[4], S. McCarroll[3], S. Purcell[1,2,3,4].* 1) Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA; 2) Department of Psychiatry, Harvard Medical School, Boston, MA; 3) Stanley Center for Psychiatric Research, the Broad Institute of Harvard and MIT, Cambridge, MA; 4) Division of Psychiatric Genomics, Mount Sinai School of Medicine, New York, NY; 5) Department of Medical Epidemiology, Karolinska Institutet, Stockholm, Sweden; 6) Departments of Genetics, Psychiatry, and Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, NC; 7) MRC Centre for Neuropsychiatric Genetics & Genomics, Cardiff University School of Medicine, Cardiff; 8) Wellcome Trust Sanger Institute, Hinxton, UK.

De novo copy number variation and point mutations have been implicated in a number of psychiatric disorders. To detect *de novo* variation, a trio design is typically used in which the proband is sequenced with both parents, although trios are more difficult to collect. We developed a method that utilizes identity by descent (IBD) information from exome sequencing in large population-based samples of unrelated samples to identify mutations that are more likely to be relatively recent, “private” mutations. These prioritized mutations will be enriched for *de novo* events, as well as mutations arising within the past several generations. For each singleton mutation in a population-based sequencing study, we search for other subjects who share the surrounding region IBD, by use of a hidden Markov model. If we observe data consistent with the index subject's maternal and paternal haplotypes each being shared by at least one other person (none of whom also share the index's singleton mutation), we can conclude that the mutation must post-date their most recent common ancestor. By focusing only on sharing of multi-megabase segments (indicative of more recent co-ancestry), we can effectively reduce the large number of rare variants each individual carries to a shortlist of truly rare alleles, more likely to contain large effect disease alleles. *De novo* variants represent a special case of this class. To evaluate this approach, we applied the method to a set of Bulgarian schizophrenia trios from an ongoing exome sequencing study. For validated *de novo* mutations, we asked whether we could flag these mutations without using any parental data, but instead using genome-wide IBD data from ~2000 unrelated, Bulgarian individuals. We were able to recover approximately 10% of the validated *de novo* mutations: this proportion will increase with larger sample sizes. For each proband, our approach flagged a small fraction of the hundreds of neutral, novel singleton variants each case carries. We posit that this fraction will be enriched for private mutations of large effect on disease as well as *de novo* variants and we are currently testing this hypothesis in a large cohort of ~12,000 unrelated Swedish samples (schizophrenia cases and controls), of which 5,000 have exome sequence data. We predict that singletons flagged by our approach will be more likely to be novel, loss-of-function and enriched in cases, compared to the much larger class of all singleton variants.

1326T

Mitochondrial DNA (mt DNA) Polymorphism 13704 (C/T) might Induce Hematologic Malignancy Predisposition in Turkish Population. *N. Duzkale, A. Tatar.* Medical Genetics, Ataturk University Medical Faculty, Erzurum, Erzurum, Turkey.

Hematologic malignancies are a group of heterogeneous neoplasia which consist of myeloid and lymphoid stem cell mass in the blood and bone marrow. Mitochondrial genome has been under investigation for various cancer types including hematologic malignancies to determine whether common polymorphisms of mtDNA can be considered as new moleculer markers of DNA damage and individual sensitivity to carcinogenic agents. It has been shown that mtDNA polymorhism can not only predict the risk of predisposition to cancer development but also tumor progression, prognosis and response to treatment. In this study, we evaluated mitochondrial (mtDNA) polymorphisms to understand the role they may have in hematologic cancer predisposition and/or pathogenesis. We have investigated 80 patients diagnosed with hematologic malignancy (30 CML, 18 CLL, 20 AML, 12 ALL) and 80 healthy controls assessing DNA samples from their peripheral blood white blood cells. Thirteen single nucleotide polymorphisms (SNPs) encompassing 6 mitochondrial regions that were evaluated were as follows: cytochrome b-15315 (C/T), NADH dehydrogenase subunit 5- 13704 (C/T), 12S rRNA -740 (G/A), 12S rRNA -680 (T/C), cytochrome c oxidase I- 7319 (T/C), cytochrome c oxidase I- 7444 (G/A), cytochrome c oxidase II- 7660 (G/A), cytochrome c oxidase II- 7975 (A/G), tRNA lysine- 8310 (T/C), cytochrome c oxidase II- 8014 (A/G, A/T), cytochrome c oxidase II- 8252 (C/G, C/T), cytochrome c oxidase II-8113 (C/A) ve cytochrome c oxidase II- 8152 (G/A). Polymorphisms were identified by the polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) analysis. Chi-square test was used to compare the study groups for their association with polymorphisms and the statistical significant level was accepted as $p<0.05$. One single SNP identified in the ND5-13704 region, 13704 (C/T), in 12 patients (3 CML, 1 CLL, 6 AML, 2 ALL) and one healthy control was statistically significant when compared to control group ($p=0.001$). We hypothesize that mtDNA 13704 (C/T) SNP can play a role in carcinogenesis process by creating susceptibility to cancer development. Further studies to elucidate the mechanism by which it does that, such as increased production of reactive oxygen species which damage the mitochondrial genome, is warranted.

1327F

Significant association of *LPL* polymorphisms with HDL- cholesterol level in Tongan adults. *I. Naka[1], R. Kimura[2], T. Inaoka[3], Y. Matsumura[4], J. Ohashi[1].* 1) Faculty of Medicine, University of Tsukuba, Ibaraki, Japan; 2) Transdisciplinary Research Organization for Subtropics and Island Studies, University of the Ryukyus, Okinawa, Japan; 3) Department of Environmental Sociology, Faculty of Agriculture, Saga University, Saga, Japan; 4) Faculty of Health and Nutrition, Bunkyo University, Kanagawa, Japan.

Lipoprotein lipase gene (*LPL*) is located at 8p22 and encodes lipoprotein lipase, which is primarily expressed in fatty (adipose) tissue and in muscle. The main function of lipoprotein lipase is to hydrolyze triglycerides in circulating chylomicrons and in very low density lipoproteins into free fatty acids and glycerol. Lipoprotein lipase also promotes the cellular uptake of chylomicron remnants, cholesterol-rich lipoproteins, and free fatty acids. Several single nucleotide polymorphisms (SNPs), including S447X (rs328), of LPL have been shown to be associated with levels of lipid fractions, hypertension, and coronary disease. Recently, genome wide association study of metabolic syndrome reported the significant associations of rs13702 at the 3' UTR of LPL with HDL-cholesterol and triglyceride levels. In this study the associations of these LPL SNPs, S447X and rs13702, with total-cholesterol, HDL-cholesterol, LDL-cholesterol, and triglyceride levels were evaluated in 127 adult subjects living in Tonga. A multiple regression analysis adjusted for age, sex, and body mass index (BMI) revealed that a copy of 477X significantly increased serum HDL-cholesterol by 8.6 mg/dl (*P*-value = 0.0151) and a copy of rs13702-T allele significantly decreased serum HDL-cholesterol by 2.9 mg/dl (*P*-value = 0.0363). No other significant associations were observed. Although rs328 and rs13702 were in linkage disequilibrium each other (D'= 1 and r^2 = 0.09), these SNPs were simultaneously included in regression analysis as independent variables. The results revealed that rs17302-T allele was significantly associated with the decrease in serum HDL-cholesterol (*P*-value = 0.025), but 477X did not show the significant association. Our results suggest that rs17302 affects the level of serum HDL-cholesterol independent of 477X and BMI in Tongan adults.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1328W

Personalized Risk Prediction for Prostate Cancer According to Specific Family History. *L. Cannon-Albright[1], F. Albright[2], W. Lowrance[3], R. Stephenson[3].* 1) Genetic Epidemiology, University of Utah School of Med., Salt Lake City, UT; 2) Pharmacotherapy Outcomes Research Center, Department of Pharmacotherapy, College of Pharmacy, University of Utah, Salt Lake City, UT; 3) Department of Surgery, Division of Urology, Huntsman Cancer Institute, University of Utah, Salt Lake City, UT.

Objective: Family history of prostate cancer (PC) is a key risk factor for the disease and may inform screening strategies. The impact on risk of PC by the number and degree of affected relatives is uncertain. We provide risk estimates for PC based on specific family history data obtained from a statewide cancer registry linked to detailed genealogy data. Methods: We analyzed the Utah Population Data Base (UPDB), a computerized genealogical resource linked to a statewide SEER cancer registry. Over 1 million individuals were analyzed. All males with specific patterns of PC family history were identified and the observed number of prostate cancer cases was compared to the expected number of prostate cancer cases using internal cohort-specific rates to estimate relative risk (RR). We considered family history out to 3rd degree relatives, number of relatives affected, paternal versus maternal family history, and age at diagnosis. Results: We analyzed a total of 611,799 males in the UPDB, of which 17,207 male subjects were diagnosed with histologically confirmed prostate cancer. Risk for prostate cancer was significantly elevated with each additional first-degree relative (FDR) affected, ranging from RR=1.85*, [95% CI: 1.80–1.89] for at least 1 FDR to RR=4.93*, [95% CI: 4.11–5.87] for at least 4 FDRs affected. In the absence of FDR family history, increased risk for PC was only observed in the presence of at least 2 second-degree relatives (SDR) affected (RR=1.18*, [95% CI: 1.12–1.25]). In the absence of positive FDR and SDR family history, there was no increased risk for PC for up to 5 or more third-degree relatives affected (RR=1.06, [95% CI: 0.92–1.21]). Earlier age at diagnosis of the affected relative was associated with higher estimates of PC risk, with RR= 4.06*, [95% CI: 3.27–4.98] for at least one FDR diagnosed before age 50 years and RR = 2.14*, [95% CI: 2.05–2.23] for at least one FDR diagnosed between age 60–69 years. We did not detect a significant risk difference based on whether the family history of PC was paternal or maternal. Conclusions: This study provides unbiased, population-based estimates of PC risk based on a man's specific PC family history constellation. The results of this study can be used to provide more informative and personalized risk estimates for PC in an individual clinical setting and on a population basis. These data should help guide PC screening policies targeting those at highest risk of the disease.

1329T

Estimating disease risk from environmental and genetic factors to motivate behavioural changes. *C.M. Lewis, J.M. Yarnall, G.H.M. Goddard, D.J.M. Crouch.* Med & Molec Genetics, King's College London, London, United Kingdom.

The identification of genetic variants predisposing to disease has raised the prospect of predicting individual-level disease risk. However, prospective studies have indicated that genetic profiling with SNPs of modest effect size provides little information of value. Here, we develop a mathematical model for joint modelling of genetic and environmental risk factors for disease to calculate a relative risk. An empiric confidence interval (CI) for the relative risk is simulated from sources of variation (genotype frequencies, odds ratios for environmental risk factors, coefficient of variation for disease prevalence). Using these CIs allows the population to be classified into risk groups, where genetic-environmental profiles of low, elevated and high risk have a level of risk that is statistically different from individuals of average risk. Applying this model to common, complex disorders using REGENT (open source software written in R) shows the following. 1) The addition of many SNPs for risk prediction is counterproductive with low risk SNPs having a low signal-noise ratio, producing wider CIs and reducing discrimination into risk categories. 2) For SNP alleles and environmental risk factors of similar risk level and frequency, the genetic component contributes more information to the risk model. 3) CIs and risk classifications become stable with estimates extracted from case-control studies with total sample sizes of >= 20,000. 4) Application of the model to common complex diseases with risk factor estimates from published studies enables a reasonable proportion of the population to be classified as at elevated or high risk, for example 19% in age-related macular degeneration (AMD), 17% for colorectal cancer, 18% for Crohn's disease. These individuals have a risk that is statistically elevated above an average member of the population and may be targeted for behavioural modification programmes or regular screening. The model also enables users to identify subgroups of the population, or specific individuals, where environmental changes (e.g. smoking cessation, BMI reduction) could reclassify their risk from high to elevated, or elevated to average, thus motivating adherence to healthy lifestyle. Simulating individuals representative of the population shows that, for example, 6.75% of obese individuals (BMI >=30 kg/m^2) classified as high or elevated risk for AMD would be reclassified to lower levels of risk if they reduced their BMI.

1330F

Improving celiac disease risk prediction by testing non-HLA variants additional to HLA variants. *J. Romanos[1,2], A. Rosen[3,4], V. Kumar[1], G. Trynka[1], L. Franke[1], A. Szperl[1], J. Gutierrez-Achury[1], C.C. van Diemen[1], R. Kanninga[1], S.A. Jankipersadsing[1], A.K. Steck[5], G. Eisenbarth[5], D.A. van Heel[6], B. Cukrowska[7], V. Bruno[8], M.C. Mazzilli[9], C. Nunez[10], J.R. Bilbao[11], M.L. Mearin[12], D. Barisani[13], M. Rewers[5], J. Norris[14], A. Ivarsson[3], H.M. Boezen[15], E. Liu[5], C. Wijmenga[1], PreventCD group.* 1) Department of Genetics, University of Groningen, University Medical Centre Groningen, Groningen, the Netherlands; 2) Lebanese American University, School of Medicine, Beirut, Lebanon; 3) Department of Public Health and Clinical Medicine, Epidemiology and Global Health, Umeå University, Umeå, Sweden; 4) Department of Medical Biosciences, Medical and Clinical Genetics, Umeå University, Umeå, Sweden; 5) Barbara Davis Centre for Childhood Diabetes, University of Colorado Denver, Aurora, Colorado, USA; 6) Institute of Cell and Molecular Science, Barts and the London School of Medicine and Dentistry, London, UK; 7) Department of Pathology, Children's Memorial Health Institute, Warsaw, Poland; 8) European Laboratory for Food-Induced Disease, University of Naples Federico II, Naples, Italy; 9) Department of Molecular Medicine, Sapienza University of Rome, Rome, Italy; 10) Clinical Immunology Department, Hospital Clínico San Carlos, Instituto de Investigación Sanitaria San Carlos IdISSC, Madrid, Spain; 11) Immunogenetics Research Laboratory, Hospital de Cruces, Barakaldo 48903 Bizkaia, Spain; 12) Department of Paediatrics, Leiden University Medical Centre, Leiden, the Netherlands; 13) Department of Experimental Medicine, Faculty of Medicine University of Milano-Bicocca, Monza, Italy; 14) Epidemiology Department, Colorado School of Public Health, Aurora, USA; 15) Department of Epidemiology, University of Groningen, University Medical Centre Groningen, Groningen, the Netherlands.

Background: At present, the majority of coeliac disease (CD) patients are not being properly diagnosed and therefore remain untreated, leading to a greater risk of developing CD-associated complications. Identifying high-risk individuals could help prioritize those to target for closer follow-up. The major genetic risk for CD, HLA-DQ2 and DQ8, is already used in clinical practice to exclude the disease. However, approximately 40% of the population carry these alleles and never develop CD. We recently identified 57 non-HLA single nucleotide polymorphisms (SNPs) that contribute to CD development. Here, we explore if CD risk prediction can be improved by adding non-HLA susceptible variants to the common HLA testing, and we assess how well our risk models can be transferred to other cohorts. Methods: We developed an average weighted genetic risk score (GRS) with 10, 26 and 57 SNPs in 2,675 cases and 2,815 controls and assessed the improvement in risk prediction provided by the non-HLA SNPs. Moreover, we assessed the transferability of the genetic risk model with 26 non-HLA variants to a nested case-control population (n=1,709) and a prospective cohort (n=1,245) and then tested how well this model predicted CD outcome for 985 independent individuals. Findings: Adding 57 non-HLA variants to HLA testing showed a statistically significant improvement compared to scores from models based on HLA only, HLA+10 SNPs, and HLA+26 SNPs. The area under the ROC curve reached 0·854 and 11·1% of individuals were reclassified to a more accurate risk group. Moreover, we show that the risk model with HLA+26 SNPs is useful in independent populations, mainly for HLA-DQ2- and DQ8-positive individuals. Interpretation: Predicting risk with 57 additional non-HLA variants improved the identification of potential CD patients. This demonstrates a possible role for combined HLA and non-HLA genetic testing in diagnostic work for CD.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1331W

Bayesian polygenic prediction of myocardial infarction risk and lipid levels. *B.J. Vilhjalmsson[1], R. Do[2], E.A. Stahl[3], B. Pasaniuc[1], S. Pollack[1], N. Zaitlen[1], J. Yang[4], M.E. Goddard[4], P.M. Visscher[5], P. Kraft[1], N. Patterson[2], S. Kathiresan[2], A.L. Price[1].* 1) Harvard School of Public Health, Boston, MA, USA; 2) Broad Institute, Cambridge, MA, USA; 3) Harvard Medical School, Boston, MA, USA; 4) Queensland Institute of Medical Research, Brisbane, Queensland, Australia; 5) The University of Melbourne, Parkville, Victoria, Australia.

Despite the success of GWAS, associated loci generally explain only a small fraction of total heritability and provide poor predictive accuracy. Although polygenic scores using less stringent P-value thresholds have been shown to yield greater predictive accuracy, we show here that under realistic genetic architectures the standard thresholding approach can be improved. Assuming that only summary statistics are available from the training data set, our approach is to estimate posterior mean effect sizes while explicitly modeling the prior mixture distribution of underlying causal effect sizes, the impact of linkage disequilibrium (LD) on indirect associations at non-causal markers, and the effects of sampling noise. We adjust estimated effect sizes for local LD as estimated from an external reference sample (Yang *et al.* 2012), and further impose a nonlinear Bayesian shrink, generalizing the standard thresholding approach that applies a shrinkage factor of 0 or 1 to each marker.

In simulations of unlinked markers with causal effect sizes drawn from a Gaussian mixture or exponential mixture distribution with $h^2=0.5$, we observed a relative increase in prediction r^2 of up to 20% for the Bayesian scheme, as compared to the standard thresholding approach with optimal threshold. When simulating phenotypes using real genotypes from the Wellcome-Trust Case Control Consortium, we observed a similar improvement in prediction r^2 when accounting for LD.

We applied our method to predict myocardial infarction (MI) phenotypes in 6,000 test samples using summary statistics from 80,000 training samples, obtaining a prediction r^2 of 3%, which is lower than predicted by theory. Results may be limited by discordance between MI training samples and early-stage MI test samples, as our method produced superior results for height (see Speliotes *et al.* abstract). Our study demonstrates the efficient use of genetic data for prediction of clinically important diseases and traits when only summary statistics are available. As training sample size increases, the prediction accuracy will further increase.

1332T

Testing for the Presence of Liability Models. *C. Herold[1, 2], T. Vaitsiakhovich[3], T. Becker[2, 3].* 1) Harvard School of Public Health, Boston, USA; 2) German Center for Neurodegenerative Diseases, Bonn, Germany; 3) Institute for Medical Biometry, Informatics and Epidemiology, Bonn, Germany.

It has recently been suggested that so-called liability models may define the architecture of the genetics of complex diseases [Zuk et al., 2012]. For individuals with less than k risk alleles a moderate baseline penetrance f0 <=5% applies, whereas for individuals with an allele load above the threshold the penetrance f>k can be as high as 50%–80%. Such models in general do not lead to a type of epistasis that is detectable by pair-wise testing but exhibit interaction of all higher orders. For a phenotype with m confirmed SNPs from the GWAS catalogue, it is an important question whether these SNPs follow a threshold model. To this purpose, we present a simple 1 d.f. regression test for the presence of a liability threshold. The test is based on the evaluation of the significance of a simple indicator parameter and allows the analysis of covariates and quantitative trait data. We conducted an extensive simulation for the threshold models with $n \in \{5, 10, 20, 50\}$ and liability threshold corresponding to a critical portion of 70%–85% causal alleles. Our models lead to marginal effects that are compatible with typical GWAS effect sizes and pair-wise interaction effects that typically are not detectable. However, with moderate sample size (3000 individuals) the power to detect the presence of a threshold model is effectively 100%, given that all SNPs of the threshold model are known. For the more realistic situation, in which the number of already detected SNPs m is smaller than the number n of SNPs in the unknown model, we can show that for m > 2/3 n there is decent power to proof the involvement in a threshold model. We present results to Alzheimer's and Type II Diabetes GWAS data sets.

1333F

Alcohol and aldehyde dehydrogenase polymorphisms and blood pressure elevation in Japanese over 20 years. *M. Isomura, T. Wang, Y. Yoshida, T. Nabika.* Functional Pathology, Shimane University, Shimane, Shimane, Japan.

Purpose: Effects of the genetic polymorphisms in the alcohol dehydrogenase-2 (ADH1B) and aldehyde dehydrogenase-2 (ALDH2) on the blood pressure (BP) were investigated in a 20-year follow-up study involving Japanese rural residents. Methods: 570 individuals attending both health examinations held in 1987 and 2006 were participated in this study. BP and demographic factors of interest were recorded on site at both time. Genomic DNAs of participants were extracted form blood and genotypes were determined by the TaqMan method. Multivariate regression analyses were performed to examine association between BP and genetic polymorphisms of those genes. Results: In the 20-year follow-up study, systolic and diastolic BP at baseline and follow-up were comparable among the ADH2 and the ALDH2 genotypes. Multivariate quantitative analyses showed that homozygote of ALDH2*2 exhibit significantly lower SBP and DBP than carrier of ALDH2*1. By contrast, ADH2 genotype did not influenced blood presser. Genetic variation in the both genes apparently influenced the drinking behavior that the number of the drinkers was significantly reduced in the homzygote of ALDH2*2. Genetic variations in either gene confer the risk of blood presser change in 20-yerars. Conclusion: These observations clarified the conflicting results regarding the effects of the ALDH2 on BP and exemplified gene-gene and gene-environment interaction influencing BP.

1334W

The *PPP6R3/LRP5* locus influences lean mass in children of different ethnic background and highlights pleiotropic effects on muscle-bone interactions. *M.C. Medina Gomez[1,2,3], D.H.M. Heppe[2], K. Estrada[1], J. Van Meurs[1], A. Hofman[3], Y.H. Hsu[4], D. Karasik[4], V.W.V. Jaddoe[2], M.C. Zillikens[1], A.G. Uitterlinden[1,2,3], F. Rivadeneira[1,2,3], GEFOS consortium.* 1) Internal Medicine, Erasmus MC University, Rotterdam, The Netherlands; 2) The Generation R Study Group, Erasmus Medical Center, Rotterdam, The Netherlands; 3) Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands; 4) Hebrew SeniorLife and Harvard Medical School, Boston, MA.

Aim: Lean and bone mass have considerably high phenotypic and genetic correlations with a shared heritability estimate ranging between 30–40% in adults. A genome-wide association study (GWAS) on total body lean mass and a bivariate GWAS on lean mass & BMD were ran in a cohort of children to identify genes with pleiotropic effects on muscle mass and peak bone mass attainment. **Methods:** Subjects are part of the Generation R study, a prospective multiethnic birth cohort in Rotterdam, The Netherlands; we included 4,096 children (mean age=6.2,SD=0.50 years) with total body DXA measurements (GE-Lunar iDXA) and genomewide genotyping (Illumina 660K). The univariate and bivariate GWAS were adjusted for age, sex, height, fat percent and 20 genomic principal components using bivariate PLINK. A $P<5\times10^{-8}$ was considered genome-wide significant (GWS). **Results:** Genomic inflation factors were close to unity indicating adequate correction for stratification. In the univariate analysis we identified a GWS association with lean mass ($\beta=0.13$, $P=2.9\times10^{-8}$) for a SNP mapping to 11q13.2, in the *PPP6R3/LRP5* locus. The SNP explained 0.8% of the variation in lean mass and was nominally significantly associated with BMD ($\beta=0.10$, P=7.6×10-5) explaining 0.4% of BMD variation. The association with lean mass was reduced after additional correction for bone mineral content ($\beta=0.08$, P=0.001), explaining 0.2% of the phenotypic variance. In the bivariate GWAS this SNP was also associated at GWS level ($P=4.8\times10^{-8}$) showing positive correlations of the bivariate trait with both lean mass (0.96) and BMD (0.68). **Conclusion:** While *LRP5* (an ubiquitously expressed gene member of the Wnt signaling pathway) is known to play a key role in bone mechanosensing with GWAS showing association with BMD and fracture in elderly adults; we showed that genetic variation in the *PPP6R3/LRP5* locus exerts pleitropic effects on muscle mass and peak bone mass acquisition of children. Given the high regional LD it is difficult to establish from which gene the GWAS signal is arising. *PPP6R3* is a gene of unknown function ubiquitously expressed across tissues among others in bone and muscle. Replication in additional children cohorts is underway while the exact same SNP has been found associated at genome significant level in a bivariate GWAS of bone strength and lean mass in two large consortia of adult individuals. Such pleiotropic effects on muscle mass and BMD observed in children are still evident later in life.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1335T

Examining the Interaction Effect of SNPs Associated with Urinary Bladder Cancer. *H. Schwender[1,2], S. Selinski[3], T. Stöcker[1,2], M. Blaszkewicz[3], R. Marchan[3], K. Ickstadt[1], K. Golka[3], J. Hengstler[3].* 1) Faculty of Statistics, TU Dortmund University, Dortmund, Germany; 2) Heinrich-Heine University Duesseldorf, Duesseldorf, Germany; 3) Leibniz Research Center for Working Environment and Human Factors (IfADo), Dortmund, Germany.

Genetic association studies aim to detect single nucleotide polymorphisms (SNPs) and combinations of SNPs associated with a disease, where it is hypothesized that the effect of individual SNPs on the disease risk increases when they are considered in combination. In genome-wide association studies concerned with urinary bladder cancer (UBC), e.g., a handful of SNPs has been detected to be associated with UBC (e.g., Kiemeney et al., 2008; Rothman et al., 2010) and was confirmed in other independent cohorts, e.g., in the case-control data collected by the Leibniz Research Center for Working Environment and Human Factors (IfADo), consisting of 1,595 Caucasian UBC cases and 1,760 Caucasian controls. It, however, has not been investigated yet if the high-risk alleles of these SNPs interact with one another, leading to an even higher disease risk. We therefore considered the most influential SNPs (rs1014971, rs9642880, rs710521, rs8102137, rs11892031, rs1495741) as well as GSTM1 (Glutathione S-Transferase M1), a detoxifying phase II metabolizing enzyme for which it is well established that a deletion variant is associated with an increased UBC risk (e.g., Bell et al., 1993), and investigated whether combinations of these polymorphisms increase the individual UBC risks. Moreover, we compared the resulting UBC risks with the ones of cigarette smoking and occupational exposure. In our presentation, we summarize the results of this examination. We found that for combinations of up to three polymorphisms the UBC risk increases with the number of polymorphisms, where the most important three-way interactions substantially differed between smokers and non-smokers. All the important interactions showed stable results, as a bootstrap analysis revealed. Nonetheless, the most influential three-way interaction still had a UBC risk smaller than the one of smoking and comparable to occupational carcinogen exposure risks. References: Bell, D.A., et al. (1993). Genetic risk and carcinogen exposure: a common inherited defect of the carcinogen-metabolism gene glutathione S-transferase M1 (GSTM1) that increases susceptibility to bladder cancer. J. Natl. Cancer. Inst. 85, 1159–1164. Kiemeney, L.A., et al. (2008). Sequence variant on 8q24 confers susceptibility to urinary bladder cancer. Nat. Genet. 40, 1307–1312. Rothman, N., et al. (2010). A multi-stage genome-wide association study of bladder cancer identifies multiple susceptibility loci. Nat. Genet. 42, 978–984.

1336F

Multivariate modeling for a more powerful genetic association analysis. *S. Eyheramendy[1], C. Meza[2].* 1) Statistics, Pontificia Univ Catolic, Santiago, Chile; 2) CIMFAV, Universidad de Valparaiso, Valparaiso, Chile.

Hundreds of genome-wide association studies (GWAS) have been performed leading to the identification of many genetic variants associated with common and complex diseases. Usually, these studies are performed on several phenotypes, but the association analysis is performed separately at each one. Statistical correlations have been observed between many traits and also pleiotropy in not uncommon among these studies. Our purpose in this work is to show that genetic association studies can benefit by jointly modeling correlated pheno- types. We developed a multivariate model and software that jointly assess the association between a genetic variant and two correlated phenotypes. We compare our approach with the univariate standard methodology and show that an increased power is achieved to detect genetic associations. Our methods is also able to detect pleiotropic loci. We show our results on real data from two large German cohorts.

1337W

Random-effects model for massive meta analysis in Genome wide association study with structured populations. *E. Kang[1], N. Furlotte[1], B. Han[2,3,4], J. Joo[5], R. Davis[6], A. Lusis[6,7], E. Eskin[1,7].* 1) Department of Computer Science, University of California, Los Angeles, CA, USA; 2) Division of Genetics, Brigham & Women's Hospital, Harvard Medical School, Boston, MA, USA; 3) Partners HealthCare Center for Personalized Genetic Medicine, Boston, MA, USA; 4) Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA, USA; 5) Interdepartmental Program in Bioinformatics, University of California Los Angeles, CA, USA; 6) Department of Medicine, University of California, Los Angeles, CA, USA; 7) Department of Human Genetics, University of California, Los Angeles, CA, USA.

Meta-analysis is an effective approach to achieve higher power in genome-wide association studies (GWAS) by combining results from multiple different studies when there is not enough power to detect the very small effect size of the SNP to target phenotype. Previously we proposed optimal meta-analysis method based on mixed model for structured populations such as mouse model organism to achieve higher power while correcting population structure which cause a large number of spurious association and inflates the significance of true associations. One promising main application of this proposed meta-analysis method is applying it to a large number of publicly available data set such as mouse phenome database. These publicly available large number of data sets are very promising sources to discover causal SNPs with small effect size through combining these studies using meta-analysis approach. However one possible challenge to performing meta-analysis with these data set is the heterogeneity in the effect size of the causal SNP to the target phenotype due to the various different conditions such as diets and sex when the data is collected. Previous studies shows that the heterogeneity in the effect size can result in the lost of statistical power in the association study when fixed effects model is applied. In this paper, we propose the random-effects model which accounts for the heterogeneity in the effect size between studies in the meta-analysis procedure, which extends our mixed model based meta-analysis approach. In the simulation experiments, we show that the new method achieves higher statistical power than previously proposed fixed effects model when there is heterogeneity, while correcting the population structure in the association procedure under the condition that there exists a population structure in the data. We show the practical utility of the proposed method by applying it to 19 HDL studies containing 4000 individuals, some of which are from publicly available mouse phenome database and 3 different atherosclerosis studies.

1338T

Comparison of rare variant tests. *J. Lihm[1,5], E. J. Yoon[1], H. Wu[1], T. Harris[1], C. Wong[1], J. Ceris[1], J. Weisburd[1], N. R. Mendell[1], W. Kim[2], K. Ahn[3], D. Gordon[4], S. J. Finch[1].* 1) Department of Applied Mathematics and Statistics, Stony Brook University, Stony Brook, NY; 2) Department of Mathematics and Statistics, University of South Florida, Tampa, FL; 3) Division of Biostatistics, Department of Public Health Sciences, Penn State College of Medicine, Hershey, PA; 4) Department of Genetics, Rutgers University, Piscataway, NJ; 5) Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

We study the null rejection rate and power of two classes of statistical procedures proposed to detect rare variants. The first class includes two multivariate tests using Hotelling's T-squared test without collapsing and with collapsing into a number of indicator variables within a gene. Collapsing all variants into a single indicator variable is studied as well. The second class of methods studied sums allele frequencies across the variants. We implement the weighted-sum method of Madsen and Browning which weights each variant by its estimated variance as well as a fixed and variable threshold methods that sum the frequencies of variants with minor allele frequencies less than a threshold. The fixed threshold methods using 1% and 5% and the collapsing method have rejection rates for non-causal genes closest to the nominal 0.05 level. Although the power of detecting causal genes for each method varies across the genes, the sum statistic of minor allele frequencies less than 5% fixed threshold consistently shows the greatest power for causal genes.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1339F

Artificial Intelligence Analysis of Prostate Cancer Susceptibility Using a Pareto-Optimized Computational Evolution System. *J.H. Moore[1], D.P. Hill[1], A. Sulovari[1], L. Kidd[2].* 1) Department of Genetics, Institute for Quantitative Biomedical Sciences, The Geisel School of Medicine, Dartmouth College, Lebanon, NH; 2) University of Louisville, KY.

Given infinite time, humans would progress through modeling complex data in a manner that is dependent on prior knowledge of their domain, computer science and statistics as well as their prior experience working with other data. For example, a human modeler interested in identifying genetic risk factors for type II diabetes might start by examining insulin metabolism genes. The goal of the present study is make extensions and enhancements to a computational evolution system (CES) that has the ultimate objective of tinkering with data as a human would (i.e. artificial intelligence). This is accomplished by providing flexibility in the model-building process and a meta-layer that learns how to generate better models. The key to the CES system is the ability to identify and exploit expert knowledge from biological databases or prior analytical results. Our prior results have demonstrated that CES is capable of efficiently navigating these large and rugged fitness landscapes toward the discovery of biologically meaningful genetic models of disease predisposition. Further, we have shown that the efficacy of CES is improved dramatically when the system is provided with statistical expert knowledge, derived from a family of machine learning techniques known as Relief, or biological expert knowledge, derived from sources such as protein-protein interaction databases. The goal of the present study was to apply CES to the genetic analysis of prostate cancer aggressiveness in a large sample of European Americans. We introduce here the use of Pareto-optimization to help address overfitting in the learning system. We further introduce a post-processing step that uses hierarchical cluster analysis to generate expert knowledge from the landscape of best models and their predictions across patients. We find that the combination of Pareto-optimization and post-processing of results greatly improves the genetic analysis of prostate cancer. This works takes a step closer to the development of genetic analysis software that approaches modeling as a human would.

1340W

A simple Bayesian method for modeling effect heterogenity across genetic studies. *C.C.A. Spencer[1], P.J. Donnelly[1,2], M. Pirinen[1].* 1) Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom; 2) Department of Statistics, University of Oxford, Oxford, United Kingdom.

There is now a wealth of successful genome-wide association studies for hundreds of phenotypes. An exciting opportunity is to compare and contrast observations across studies and to combine studies to discover new associations. Most standard statistical methods assume a genetic variant to have the same effect on risk across studies. As an increasing amount of genetic data becomes available from different populations and/or related phenotypes it is beneficial to relax the fixed effects assumption for the purpose of identifying new associations. Moreover, it is often of interest to compare statistical models directly in order to quantify the evidence for similarities and differences in effect between studies. We introduce a novel Bayesian framework to answer questions like:

- Is there evidence for association in all studies?
- What is the combined evidence assuming the effect to be similar?
- Is the effect the same across studies?
- Is there an effect in only one study?
- How similar are the effects across studies?

An appealing property of the approach is its computational simplicity which makes it widely accessible to the genetics research community. We illustrate the utility of the proposed methodology with two applications: a study into genetic determinants of malaria susceptibility across three African populations, and combining existing studies of auto-immune disease.

1341T

Effectively identifying and characterizing eQTLs from multiple tissues using a meta-analytic approach. *J. Sul[1], B. Han[1], C. Ye[2], T. Choi[3], E. Eskin[1,4].* 1) Computer Sci, Univ California, Los Angeles, Los Angeles, CA; 2) Broad Institute of MIT and Harvard, Cambridge, Massachusetts, 02142, USA; 3) Predictive Biology, Inc., San Diego, CA; 4) Department of Human Genetics, University of California, Los Angeles, California 90095, USA.

Gene expression data, in conjunction with information on genetic variants, have enabled studies to identify expression quantitative trait loci (eQTLs) or locations in the genome that affect gene expression. Moreover, technological developments and cost decreases enable studies to collect expression data in multiple tissues. One advantage of multi-tissue datasets is that studies can combine results from different tissues to identify eQTLs more accurately than examining each tissue separately. The idea of aggregating results of multiple tissues is closely related to the idea of meta-analysis which aggregates results of multiple genome-wide association studies. In principal, meta-analysis methods can be used to combine results from multiple tissues. However, a key challenge is that genetic variants may have different effects in different tissues, a phenomenon known as heterogeneity. In this paper, we propose a novel approach leveraging a recently developed meta-analysis method that incorporates heterogeneity. We first show by simulations that our method is more powerful than a meta-analysis method that does not consider heterogeneity in detecting eQTLs from multiple tissues when heterogeneity exists. We then apply our method to multi-tissue dataset from mouse. This dataset contains expression from many different tissues and different numbers of samples for tissues allowing us to observe power to detect eQTLs under various scenarios. We analyze four tissues from 50 samples per tissue and ten tissues from 22 samples. We demonstrate that our method detects 95% and 23% more eQTLs than a traditional eQTL method or the meta-analysis method that does not consider heterogeneity, respectively, in the four tissues. In the ten tissues, we detect 520% and 11% more eQTLs. Additionally, our method provides an interpretation framework that accurately predicts whether an eQTL has an effect in a specific tissue. Our method computes the posterior probability of the presence of an effect building on recent work in interpretation of meta-analysis. Simulations show that this probability predicts effects more accurately than an approach using p-values. Applying the framework to the mouse dataset, we try to detect eQTLs having effects in all tissues since they can help us understand tissue-independent regulatory mechanisms. We show that our approach detects 221 and 45 more such eQTLs than the approach using p-values in the four and ten tissues, respectively.

1342F

PEDF gene Polymorphism (p.Met72Thr) Analysis in Diabetic Retinopathy Cases. *S. Passan, V. Vanita.* Department of Human Genetics, Guru Nanak Dev University, Amritsar, Punjab, India.

Purpose: To study the association of Pigment Epithelium Derived Factor (PEDF) gene polymorphism (p.Met72Thr) in type 1 and type 2 diabetic retinopathy cases. Material and Methods: The present case control association study included 777 type 2 diabetes patients and 42 type 1 diabetes patients. All the patients were unrelated and of Indian origin. These patients were recruited from Dr. Daljit Singh Eye Hospital, Amritsar. 507 diabetic patients had different types of retinopathies. Of these 128 had Back-ground Diabetic Retinopathy (BDR), 299 patients were suffering from Proliferative Diabetic Retinopathy (PDR), 80 patients had Pre Proliferative Diabetic Retinopathy (PPDR), and 312 diabetic patients with no sign of retinopathy were collected as controls. The genotyping was carried out by bi-directional sequence analysis of the amplified products. Statistical analysis was performed using SPSS for windows (Version. 16.0). One way ANOVA test was applied to find any association for various clinical parameters (age, age of onset of diabetes, duration of diabetes, basal metabolic index (BMI), systolic and diastolic blood pressure) with diabetic retinopathy. Genotype distributions and allele frequencies were evaluated using statistical calculator (StatPac V. 3. 0). MedCalc (V 9. 3. 9. 0) was used to determine the odds ratios (OR) and 95% confidence intervals (CI). Results: The present study, revealed statistical significant differences between age, age of onset, duration of diabetes, and BMI when DR patients (BDR, PPDR, and PDR) were compared with controls. For *PEDF* p.Met72Thr polymorphism the allele frequencies (p=0.9266) and genotype frequencies (p=0.5373) in DR patients showed no statistical difference when compared with controls. However, OR analysis revealed a significant relationship with (TC) heterozygote genotype of the *PEDF* polymorphism in DR cases in comparison to controls (OR=1.32). Conclusions: To our Knowledge, the present study is the first report of case control association study of *PEDF* polymorphism (p.Met72Thr) with diabetic retinopathy in North Indian type 1 and type 2 diabetic patients. Present findings suggest no association of PEDF gene polymorphism (p.Met72Thr) with DR. However, TC heterozygote genotype showed 1.3 times increased risk of diabetic retinopathy indicating that TC genotype of *PEDF* is associated with increased risk of diabetic retinopathy.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1343W

Chromosome 10 gene may be associated to response to Leishmania antigens. *L. Pereira, J. Pescarini, R. Ferreira, L.M. Camargo, H. Krieger, L.M. Garrido.* Dept Parasitology, Univ Sao Paulo, São Paulo, Brazil.

Leishmania is a protozoan that causes an endemic disease at the Amazonian region, affecting about 2,000 individuals every year, with a mortality ratio close to 10%. In Brazil, there are two forms of leishmaniasis: cutaneous and visceral, both transmitted by an insect of the subfamily Phlebotominae. The test of Montenegro (MN), is an immunological technique of delayed skin hypersensitivity with specificity between 86% around 100%. It is performed by inoculating a small amount of leishmania antigen in the arm and after 48 hours it is possible to observe a papule. This papule is measured and if its size is equal or larger than 5mm, the individual is assumed to be positive to leishmania infection. The sample is composed by two subsamples one of 80 individuals from Monte Negro (10° 15' S, 63° 18' W) collected in the late 90s and 34 individuals from the same locality collected in the late 2000. This last sample was ascertained in order to verify the association of some genetic markers and Montenegro phenotypes. Two Taqman (rs11251056 and rs11251062) from chromosome 10 (Applied Biosystems) SNP genotyping probes were chosen based on suggestions of association disclosed by genomic scanning with the Genechip Affymetrix 250 Nsp I. Joint-analysis uses statistical techniques to combine into one summary measurement the results of independent studies aimed at a single issue. The meta-analysis, using small samples such as these are relevant since it reveals important characteristics of complex traits such as the response to infection and may show possible relationship to infection and therefore display mechanisms of susceptibility/resistance. The meta-analysis disclosed an association for rs11251062 (p=0.03) with MN reaction while the analysis for SNP rs11251056 showed a borderline significance value (p= 0.07). These two SNPs are located near lincRNA which is an important regulatory factor and shows differential expression patterns in a wide variety of cells and tissues (LI, et al., 2012; NIE, et al., 2012). This regulatory region showed promising results indicative associations with our phenotype, which can be re-evaluated in future samples and thus actually confirm this strong suggestion (CNPq).

1344T

Genetic associations with essential amino acids in infants with patent ductus arteriosus. *K.K. Ryckman[1], J.M. Dagle[1], O.A. Shchelochkov[1], S.L. Berberich[2], J.C. Murray[1].* 1) Dept Pediatrics, Univ Iowa, Iowa City, IA; 2) State Hygienic Laboratory, Univ Iowa, Iowa City, IA.

Patent ductus arteriosus (PDA) is a significant complication of prematurity where the ductus arteriosus fails to close normally after birth. If left untreated the infant can experience difficulty breathing, poor growth and eventually heart failure. The risk for PDA is directly correlated with gestational age (GA), where the earlier an infant is born the more risk there is for PDA; however not all infants born early experience a PDA. In addition to the GA risk for PDA we have previously demonstrated an associations with alleles of the genes transcription factor AP-2 beta (TFAP2B) and tumor necrosis factor receptor-associated factor 1 (TRAF1). We have also observed that infants with PDA on day of life 5–7 have higher levels (compared to infants without PDA adjusted for GA) of essential amino acids leucine (P=1.0×10-3), methionine (P=1.2×10-3), phenylalanine (P=4.7×10-4) and valine (P= 3.2×10-4) measured from newborn screening samples at 24 hours of age. To explore these associations further we examined 20 SNPs in 8 candidate genes for PDA (including TFAP2B and TRAF1) for association with ten amino acid concentrations captured from the Iowa Neonatal Metabolic Screening Program on 199 infants. Infants included in the study were born less than 32 weeks GA and not on total parenteral nutrition at the time of newborn screening. Non-parametric Kruskal-Wallis tests were used to test for associations between each SNP and amino acid levels in all infants, infants with a PDA and infants without a PDA. The strongest associations were with rs7267595 in the jagged 1 gene (JAG1) and alanine (P=0.03), leucine (P= 3.5×10-3), methionine (P=4.7×10-3) and valine (5.1×10-3) in all infants. When examining associations further the association was only present in the absence of a PDA. Jag1 knockout mice die in early postnatal period secondary to PDA. In addition, rare mutations in JAG1 are associated with Alagille syndrome, a rare genetic disorder of the liver, heart (including PDA) and kidney. Preterm infants have higher essential amino acid levels likely due to liver and kidney immaturity. Genetic associations with JAG1 and the essential amino acids may represent a genetic predisposition to the liver dysfunction observed in extremely preterm infants.

1345F

Low frequency SNPs are associated with congenital and severe clinical phenotypes in an Electronic Medical Record. *J.D. Mosley[1,2], D.M. Roden[1,2], J.D. Denny[1,3].* 1) Department of Medicine, Vanderbilt University, Nashville, TN; 2) Clinical Pharmacology, Vanderbilt University, Nashville, TN; 3) Biomedical Informatics, Vanderbilt University, Nashville, TN.

The human genome harbors numerous SNPs with low minor allele frequencies (MAF). Their low frequency has been attributed to both recent rapid population expansion and negative selection pressure due to the deleterious effects of the resultant gene products. We hypothesized that subjects who were homozygous for the minor alleles of relatively rare SNPs would have an increased burden of human disease. To test this hypothesis, we analyzed a set of 8,659 Vanderbilt University Medical Center patients (6012 whites, 1743 blacks, 911 other races) who had been previously genotyped in GWAS studies. Phenotype data for these subjects were generated from ICD-9, CPT and physician-maintained problem lists contained in the electronic medical record (EMR). 4220 SNPs with a MAF less than 0.10 in both whites and blacks were tested against 1670 phenotypes representing individual or aggregate clinical problems. The median number of homozygotes for the minor allele of a SNP was 20 subjects (s.d. 14). Bonferroni-corrected p-values were obtained using exact multivariable logistic regression analyses comparing minor-allele homozygotes to the common-allele homozygotes and the number of significantly associated ($p<0.05$) phenotypes was tallied for each SNP. 224 SNPs were significantly associated with 1 to 4 phenotypes after Bonferroni adjustment. Among SNPs causing missense mutations, higher polyphen scores were associated with more phenotypic associations (p=0.04). A review of the SNPs with multiple phenotype associations demonstrated a number of diverse biologically plausible genotype-phenotype associations. For instance, 6 of 12 subjects homozygous for a BRCA1 missense variant (MAF=0.04) were diagnosed with leukemia (p=0.00001). The associations fell across a number of domains including developmental conditions (renal agenesis, coarctation of the aorta, congenital heart disease), acquired conditions (ulcerative colitis, osteomyelitis), tumors (leukemia, pheochromocytoma) as well possible syndromic phenotypes (such as adrenal and thyroid cancers). In summary, our analytical approach supports the idea of leveraging EMR data to identify clinical phenotypes associated with rare minor alleles in a diagnosis-agnostic fashion.

1346W

Rare variant testing for meta-analysis and survival analysis. *H. Chen[1], T. Lumley[2], J. Dupuis[1, 3], J. Brody[4], A. Morrison[5], B. Cornes[6, 7], D.J. Lybarger[3], B. Davis[5], C. Sitlani[4], D. Siscovick[4], J.B. Meigs[6, 7], L.A. Cupples[1, 3].* 1) Department of Biostatistics, Boston University School of Public Health, Boston, MA; 2) Department of Statistics, University of Auckland, Auckland, New Zealand; 3) The National Heart, Lung and Blood Institute's Framingham Heart Study, Framingham, MA; 4) Cardiovascular Health Research Unit, Departments of Medicine and Epidemiology, University of Washington, Seattle, WA; 5) Human Genetics Center, School of Public Health, The University of Texas Health Science Center at Houston, Houston, TX; 6) General Medicine Division, Massachusetts General Hospital, Boston, MA; 7) Department of Medicine, Harvard Medical School, Boston, MA.

The Sequence Kernel Association Test (SKAT) (Wu et al 2011, Am. J. Hum. Gen) is an omnidirectional test for association between a set of possibly-rare genetic variants and a binary or continuous phenotype. We propose two extensions to the test that increase its applicability in the context of a consortium of cohort studies. The first is an approach to meta-analysis for settings where individual data cannot be shared, and the second is a version for survival phenotypes. Both extensions are based on the characterization of the SKAT test as a weighted sum of squared z-statistics for association between phenotype and individual variants. The meta-analysis requires each participating study to share the cohort-specific minor allele frequencies, genotype covariance matrix and the z-statistics for score tests of each variant in a generalized linear model. The score tests may optionally be adjusted for population structure or other factors, and may involve sampling weights to adjust for phenotype-based sampling. The score test statistics are then meta-analyzed with weights based on the statistical information provided by each study (or simply on sample size), and combined in a weighted sum to give the overall SKAT test statistic. The non-standard sampling distribution for the test statistic is estimated using the eigenvalues of a sample-size-weighted average of the per-study genotype covariance matrices. With no adjustment variables the meta-analysis test statistic is identical to the SKAT test statistic using individual data and adjusting for study; with adjustment variables the agreement is still very close. Generalizing the test to survival phenotypes using the score test in a Cox model (logrank test) does not work. The logrank test is anticonservative for rare variants, and leads to a seriously anticonservative overall test. Replacing the score z-statistic with the signed square root of the log partial likelihood ratio test statistic gives a test with accurate control of Type I error. Meta-analysis of this test involves weighted meta-analysis of the signed likelihood root test statistics in exactly the same way that score z-statistics statistics are meta-analyzed for linear and binary phenotypes. To illustrate the method we will present the meta-analysis SKAT results for associations with fasting insulin in the CHARGE-S targeted resequencing study.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1347T

Selection of sequence variants for quantitative traits using penalized regression: using LASSO, LARS and Elastic Net in the Tiled Regression framework. *Y. Kim[1], B. Suktitipat[1,2], A.J.M Sorant[1], A.F. Wilson[1].* 1) IDRB/ Genometrics Sec, NHGRI/NIH, Baltimore, MD; 2) Biochemistry Dept, Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand.

Selecting the set of independent genetic variants that contribute to the variation of quantitative traits can be challenging because of the multicollinearity caused by highly correlated sequence variants either in linkage or gametic disequilibrium; and the use of regression methods that ignore multicollinearity can be problematic. Penalized regression methods are known to be robust to the multicollinearity of variables — shrinking the coefficients of less informative variables close to or equal to zero until the regression model is optimized. In this study we investigate the performance of penalized regression methods and compare these results to those from stepwise regression as currently implemented in the Tiled Regression Analysis Program (TRAP v1.0). We compared the performance of three penalized regression methods: LASSO, Elastic Net (EN) and LARS, to results from stepwise regression (STEP). We simulated 4000 individuals with 300,000 variants based on HapMap data, and 100 replicates of quantitative traits from seven independent causal loci with the same effect size but with various locus-heritability (h^2) ranging from 0.00045 to 0.01 based on an additive genetic model. We examined the detection rates (% per 100 replicates) of each causal locus, precision (true positive/true positive + false positive) and sensitivity (true positive/ true positive + false negative) of all causal loci in the final subset under two scenarios: including and excluding causal variants. In the presence of causal variants, LASSO, EN, and STEP showed high detection rates (76–100%) for causal variants with locus-specific $h^2 > 0.005$, but LARS and STEP were higher than LASSO and EN for detecting causal variants of $h^2 < 0.005$. However, LARS showed the least precision (0.004) and sensitivity (0.42) for the final subsets. After excluding causal variants in causal loci, the penalized regression methods showed slightly higher detection rates (1–28%) than STEP (0–15%) — doing a somewhat better job at detecting variants that were correlated to the excluded causal variants. Overall sensitivity and precision were decreased compared to the results including causal variants. LARS showed the least precision (0.002), but observed similar sensitivity (0.18) with LASSO (0.16) and EN (0.16). In conclusion, LASSO and EN may be an effect tool in detecting variants that are correlated with causal variants when the causal variants are not among the variants considered.

1348F

Identification of genome-wide SNPs that are informative for individual genetic heritage in the family investigation of nephropathy and diabetes. *R. C. Williams, FIND Research Group.* Phoenix Epidemiology and Clinical Research Branch, NIDDK, NIH, Phoenix, AZ.

Assessment of IGH is potentially important in genome-wide association studies (GWAS) but evaluation of an American Indian component is limited because there is little information on parental frequencies from American Indians typed on a commercially available platform. We therefore searched the 22 autosomes for SNPs informative for contrasts between three parental populations: European Heritage (Eu) (represented by HAPMAP, CEU), African Heritage (Af) (HAPMAP, YRI and LWK, combined and averaged allele frequencies), and American Indian heritage (AI) (represented by a GWAS on the Affymetrix 6.0 SNP array in 964 full heritage Pima Indians). From among 719,109 SNPs with valid genotypic data in all parental samples, 1390 SNPs were identified according to the following criteria: 1) a difference in allele frequency (Δ) in at least one pair of parental groups ≥ 0.5, 2) a distance $\geq$ 500kb between syntenic loci, 3) a sample size ≥ 100 for estimating the allele frequencies, and 4) no A/T or G/C polymorphisms. The mean distance between adjacent syntenic loci was 729,148bp (range: 500,002–12,126,238bp). The mean of the maximum Δ value was 0.708 (range: 0.500–0.986). Maximum likelihood methods were used to calculate IGH using these markers. C.C. Li's method for estimating the information and variance was used to quantify precision of the estimates. In the parental populations, where the expected mean is 1.000 for the respective IGH, the agreement was very close: Eu, N=174, HapMap CEU (Mean=0.993 ± std.= 0.008); Af, N=209, HapMap YRI (0.999±0.003); Af, N=110, HapMap LWK (0.981± 0.014); AI, Pima, N=964 (0.990± 0.047). The method was then applied to data from the FIND GWAS for four ethnic groups: 1) American Indians, N= 869 (mean Eu=0.046, ±std.=0.090; mean AI=0.945 ± 0.110; mean Af=0.009 ± 0.048); 2) African Americans, N=1385 (Eu=0.151 ± 0.103; AI=0.016 ± 0.026; Af=0.832 ± 0.109); 3) Mexican Americans, N=1451 (Eu= 0.473 ± 0.134; AI=0.451 ± 0.139; Af=0.076 ± 0.048); 4) European Americans, N= 826 (Eu=0.958 ± 0.087; AI= 0.015 ± 0.021; Af=0.026, ± 0.083). In the combined FIND samples, N=4531, the mean standard errors for IGH are small (Eu, 0.014; AI, 0.010; Af, 0.017). We have identified a set of SNPs, included on the Affymetrix array, that provide precise estimates of IGH from 3 parental groups. This will facilitate assessment of IGH in studies that have used this array; genotyping these markers, or a subset, may also be useful in other studies.

1349W

Testing Copy Number Variant/Trait Associations Detected using Manhattan Plots. *G.A. Satten[1], D. Ramachandran[2], J.G. Mulle[2], A.S. Allen[3], L.J.H. Bean[2], C. Maslen[4], S.L. Sherman[2], R.H. Reeves[5], M.E. Zwick[2].* 1) Centers for Disease Control and Prevention, Atlanta, GA; 2) Emory University, Atlanta GA; 3) Duke University, Durham NC; 4) Oregon Health Sciences University, Portland OR; 5) Johns Hopkins University, Baltimore MD.

Background: Finding association between copy number variants (CNVs) and disease is hampered by poor detection of CNVs using log-intensity ratio (LIR) data from genotype arrays. Breheny et al. (Plos One 2012; 7(4):e34262) suggested directly comparing LIR values between cases and controls, rather than first inferring copy number states, obtaining a p-value at each locus. CNV-trait association is then found by using CNV-detection software on the Manhattan plot of -log(p) values. We sought to improve upon the methods of Breheny et al. by developing a method to assign a formal p-value to associations found in this way. We applied our approach to a GWAS of children with Down syndrome (DS), who are at greatly increased risk for congenital heart defects including atrioventricular septal defects (AVSDs).

Methods: DNA from children with DS and AVSDs (cases, n=238) and with DS but without heart defects (controls, n=267) was genotyped using the Affymetrix 6.0 array. We fit a linear regression model for LIR at each locus, testing the effect of case status on LIR in the presence of genotype principal components to control for population stratification. We used our CNV BEAST software (www.duke.edu/~asallen/Software.html) to find runs of large -log(p) values, and found several candidate associations. We tested significance with a bootstrap hypothesis test. We first calculated the residuals from our regression model. We then resampled from these residuals to generate 1000 bootstrap replicate datasets by that preserve population structure but have no other difference between cases and controls. We resampled residuals by individual to preserve spatial correlations. Because the CNV BEAST gives a scalar score to each CNV, we can calculate a p-value by comparing the distribution of largest scores assigned to any CNV in each replicate to the score observed in the actual data.

Results: Although our largest signal seemed robust, comprising over 100 adjacent loci on Chromosome 1 between 149041794 and 149246018, we found a 7.1% chance of seeing such a large signal under the null hypothesis.

1350T

A novel collapsing method for rare copy number variants (CNVs). *J.P. Szatkiewicz[1], P.F. Sullivan[1,2], J.Y. Tzeng[3].* 1) Department of Genetics, University of North Carolina, Chapel Hill, NC, 27599; 2) Department of Psychiatry, University of North Carolina, Chapel Hill, NC, 27599; 3) Department of Statistics and Bioinformatics Research Center, North Carolina State University, Raleigh, NC, 27607.

CNVs play an important role in the etiology of multiple psychiatric disorders. Multiple recent studies found that the genome-wide burden of rare CNVs is greater in cases with schizophrenia, autism, or bipolar disorder than in controls. Several specific and likely the most pathogentic CNVs have been discovered; however, the individual effects of most CNVs on disease risk remain unknown due to modest marginal effect size or rarity of the CNV. Collapsing approaches could be important to evaluate multigenic models to study how CNVs impact risk for psychiatric disorders. In contrast to sequence variants that affect single nucleotides, CNVs vary in size, type, dosage, and sequence level details of gene disruption. Because of its multifaceted nature, CNV analysis is more challenging and the most important consideration is allelic heterogeneity. For example, heterogeneous clusters of rare CNVs are often observed where CNVs with variable breakpoints and sizes disrupt genes in different ways. CNVs can also have heterogeneous effects by being a mixture of neutral, risk, and protective variants. Consequently, collapsing methods developed for rare sequence variants cannot be generalized directly to rare CNVs. Existing burden tests also do not fully explore CNV-specific challenges. Typically, rare CNVs are aggregated into discreet categories, such as deletion, duplication, exon, gene, or genome; and within each category the rates of events were compared between cases and controls, treating all events equivalently. This approach falls short in two counts: ignoring heterogeneity entails loss of power and testing one-category-at-a-time has suboptimal efficiency. We propose a novel collapsing method for CNVs that is robust to multiple types of heterogeneity. Our method is based on SimReg (Tzeng et al. PMID 21835306), a similarity collapsing approach, to collectively examine the effects of multiple CNV features (e.g. size, type, dosage, mixture effects) and weight CNVs by their frequencies and details of gene disruption. Multiple confounders can be simultaneously corrected. We conducted simulation to evaluate type I error and power in various scenarios and compared our method to the existing CNV burden test as implemented in PLINK. We then applied our method to multiple well-powered psychiatric GWAS studies where each included more than 10K subjects. Full results will be reported at the conference, but preliminary results suggest manifest improvement in performance.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1351F

A pseudo-score-based meta-analysis in genetic association studies: Application to Mayo Genome Consortia data. *E. Ryu[1], RE. Gullerud[1], PA. Decker[1], J. Pathak[1], PJ. Limburg[2], SJ. Bielinski[1].* 1) Health Sciences Research, Mayo Clinic, Rochester, MN; 2) Gastroenterology and Hepatology, Mayo Clinic, Rochester, MN.

Objective: The Mayo Genome Consortia (MayoGC) is a cohort of 6,307 Mayo Clinic patients who have electronic medical record (EMR) and genotype data. Among patients who have a record of colonoscopy, we identified 1048 cases with neoplastic colon polyps (adenoma, serrated, or cancer) and 2386 patients with normal colonoscopy results. We propose a new meta-analysis method that will be used to investigate less-common genetic variants that contribute to neoplastic colon polyps. **Method:** A pseudo-score confidence interval is obtained by inverting a test that uses a Pearson chi-square statistic to compare fitted values under two models: one under the null hypothesis that takes various fixed values, and the other model under the alternative hypothesis. For each study, midpoint and pseudo standard error will be obtained from the pseudo-score confidence interval. Borrowing an idea from inverse-variance weighting approach, the common parameter estimate and p-value will be calculated using a new meta-analysis method. As an example, SNP rs9532209 in chromosme 13 with MAF equal to 2.5% was randomly chosen for an association testing with being neoplastic. A simulation study was also performed to compare the new approach to two popular meta-analysis approaches (inverse-variance weighting method, and the weighted Z-score method) and a method based on profile likelihood. **Result:** P-values for testing an effect of rs9532209 on case status are 0.018, 0.020, 0.011, and 0.015 for inverse-variance weighting method, the weighted Z-score method, profile likelihood-based approach, and pseudo-score based approach, respectively. Common parameter estimates are -0.451, -0.486, and -0.457 for inverse weighting method, profile likelihood-based approach, and pseudo-score based approach, respectively. Assuming MAFs close to 0.05 for three studies and true odds ratio being 2, with sample sizes 400, 500, and 600, a simulation study showed that our new approach performed similar to profile likelihood based approach, and better than two popular existing methods. **Conclusion:** Our proposed method appears to perform better than existing popular meta-analysis methods for variants with MAF < 5%. Furthermore, this approach can be applied to the datasets that the popular methods do not work due to convergence failures in standard logistic regression. It is because pseudo-score-based method does not require the estimated variance of the effect size for each study.

1352W

Diacylglycerol kinase K variants impact hypospadias in a California study population. *S. L. Carmichael[1], N. Mohammed[2], C. Ma[1], D. Iovannisci[2], S. Choudhry[3], L. S. Baskin[3], J. S. Witte[4], G. M. Shaw[1], E. J. Lammer[2].* 1) Pediatrics, Stanford University School of Medicine, Stanford, CA; 2) Children's Hospital Oakland Research Institute, Oakland, CA; 3) Department of Urology, University of California, San Francisco, CA; 4) Department of Epidemiology and Biostatistics and Institute of Human Genetics, University of California, San Francisco, CA.

PURPOSE: Hypospadias occurs when the urethral meatus is on the underside of the penis. It is one of the most common birth defects, affecting 4–6 per 1000 male births. It can range from mild (coronal or glanular), moderate (penile shaft) or severe (at the peno-scrotal junction or perineal area). A recent genome-wide association study reported the novel finding that variants in diacylglycerol kinase kappa (DGKK) were associated with hypospadias. The discovery phase was based on pooled samples, and replication was restricted to two SNPs. Our objectives were to determine whether this finding could be replicated in a more racially-ethnically diverse study population of California births and to provide a more comprehensive investigation of variants. METHODS: We examined the association of 27 DGKK SNPs with hypospadias, relative to population-based non-malformed controls born in selected California counties from 1990–2003. Analyses included a maximum of 928 controls and 665 cases (91 mild, 336 moderate, 221 severe, 17 undetermined). Results for mild and moderate cases were similar so they were grouped together. RESULTS: For mild and moderate cases, odds ratios (OR) for 15 of the 27 SNPs had p-values <0.05; two of the ORs were <1, and the others ranged from 1.3 to 1.8. Among severe cases, ORs tended to be closer to one and none of the p-values were <0.05. Due to high LD across the SNPs, haplotype analyses were conducted, and two blocks were generated. These analyses identified a set of eight variants that was associated with a three- to four- fold increased risk, relative to the most common haplotype, regardless of severity of the phenotype (the OR was 4.1, p<10-4 for mild to moderate cases and 3.3, p=0.001 for severe cases). CONCLUSIONS: This study confirms that DGKK variants are associated with hypospadias. Further studies are needed to enable a more thorough investigation of DGKK variability and to delineate the mechanism by which DGKK contributes to urethral development.

1353T

GEEWAS: Family-based genome-wide association studies for non-normal responses. *C. Ekstrom.* Biostatistics, Univ Southern Denmark, Odense, Denmark.

Genome-wide association studies (GWAS) are widely used to identify genes influencing traits and diseases from data on unrelated individuals. However, the lack of success in explaining more than a fraction of the heritability has lead to increased focus on family-based GWAS studies - in partiular since family-based data are available back from linkage era. The linear mixed effect model has been the workhorse for GWAS of normally distributed traits for both unrelated and related individuals, but its generalizations to non-normal traits have been prohibitively slow.

We examine the use of generalized estimating equations (GEE) for family-based genome-wide association studies with unequal family structure. GEEs accommodate both normal and non-normal responses, provide robust standard errors against misspecification of the intra-family correlations, and are heteroscedasticity consistent, which is relevant when, for example, several rare variants are segregating in unidentified sub-populations. Correct modelling of the working correlation among individuals in a pedigree improves the efficiency but the kinship coefficient matrix used with linear mixed models needs to be adapted for use with GEEs.

Our implementation, GEEWAS, runs in R and simulations show that the GEEWAS implementation runs approximately 30–50 times faster than a corresponding linear mixed effect model and provides the same unbiased estimates. Our results show that correct modeling of the correlation matrix improves the efficiency of the estimates and approximately 10% smaller standard errors compared to an independent working model which in turn provides more powerful tests to detect genes associated with a qualitative or quantitative phenotype. The GEEWAS implementation handles families of varying structure and the simulations show that GEEWAS is able to identify genes influencing both binary and ordinal categorical scale responses that are used for example to classify mental disorders.

1354F

An multiple testing procedure for association studies appropriately incorporating admixture signals. *G. Gao, W. Chen.* Dept Biostatistics, Virginia Commonwealth Univ, Richmond, VA.

Admixed populations (such as African Americans) denote the populations formed by recent admixture of two or more ancestral populations. For gene mapping in admixed populations, admixture mapping tests use admixture linkage disequilibrium (LD) existing in large chromosomal segments inherited from an ancestral population and can only identify a causal variant in a large chromosomal region (several Mbs). To identify a causal variant in a small region (less than a few hundred Kbs), association tests that correct for local ancestry and therefore use background LD existing in the ancestral populations have been developed. However, these tests can have relatively low power. Recently, joint association tests that combine information from admixture mapping tests and association tests that correct for local ancestry have been proposed, but the joint methods may have inflated type I error rates when null single nucleotide polymorphisms (SNPs) are located in regions with admixture signals but far from causal variants (i.e., in admixture LD but not in background LD with the causal variants). A main reason for this is that these joint methods may use too much information from the admixture mapping at a test SNP. In this study, we adapt the generalized sequential Bonferroni procedure (GSB) to association studies to incorporate appropriate amount of information from admixture mapping into association tests that correct for ancestry. Our simulation studies indicate that the GSB procedure not only control family-wise error rates, but also can have improved power compared to the association tests that correct for local ancestry. We applied the GSB procedure to association studies in a data set from the Multi-Ethnic Study of Atherosclerosis project.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1355W

Correcting for population structure due to rare and common variants with an approximate Bayesian regression framework. *G.E. Hoffman[1], J.G. Mezey[1,2].* 1) Biological Statistics and Computational Biology, Cornell University, Ithaca, NY; 2) Department of Genetic Medicine, Weill Cornell Medical College, New York, NY.

With the growing interest in testing associations of rare variants, new problems of population structure are arising in exome and genome sequence datasets due to the more recent origin and more localized distribution of rare compared to common variants. Two widely used statistical approaches, namely including principal components as covariates in a regression or using a linear mixed model, have proven successful in correcting for structure due to common variants as such structure is well modeled by a linear gradient [1,2]. However, these methods are not able to account for the complexities unique to rare variants where population structure can be more localized due to spatially private alleles [3–5]. We propose a novel regression framework for testing associations of rare variants that incorporates nonlinear functions of principal components to explicitly model the localization of rare variants. Since only a small fraction of the principal components and nonlinear functions thereof will be biologically relevant, we use an approach akin to penalized splines to select a small subset of relevant basis vectors using a sparse Bayesian learning method for variable selection. Moreover, our method approximately integrates over parameter uncertainty using a variational Bayes technique. We consider linear and logistic regression models to test a single marker but our framework can incorporate collapsing or multiple-marker tests of association [6,7]. We conduct an extensive simulation study using simulated spatial population data and exomes sampled from the NHLBI Exome Sequencing Project to evaluate the conditions under which our novel method correctly accounts for population structure due to rare variants.

[1] Price et al., 2006. Nature Genetics 38: 904–9. [2] Kang, et al., 2010. Nature Genetics 42: 348–54. [3] Mathieson and McVean, 2012. Nature Genetics 44: 243–6. [4] Nelson, et al., 2012. Science In press [5] Tennessen et al., 2012. Science In press. [6] Wu, et al. 2011. American Journal of Human Genetics 89: 82–93 [7] Asimit and Zeggini. 2010. Annual Review of Genetics 44:293–308.

1356T

Unified association analysis of genes, regions or pathways containing multiple SNPs. *L.C. Lazzeroni, A. Ray.* Psychiatry and Behavioral Sciences, Stanford University, Stanford, CA.

Candidate gene and genome-wide association studies typically examine several genetic polymorphisms within the same gene, genomic region or biological pathway. This comprehensive strategy increases the likelihood that genetic associations existing in the population will be represented by one or more functional or tagging SNPs in the data. Unfortunately, the numerous SNP-specific test results produced by this strategy are often difficult to interpret, especially given the stringent standards imposed by multiple testing corrections. If the gene, region or pathway is the target of scientific inquiry, inconsistent results across SNPs can lead to uncertainty in reaching an overall conclusion. As a consequence, several authors have recently proposed various methods for conducting unified gene-wide association tests using data from multiple SNPs that can also be applied to regions or pathways. We will consider the desirable properties of such tests, which include: (1) power and maintenance of nominal significance levels in various settings; (2) interpretability of results and accuracy of effect sizes at both the gene and SNP levels; and (3) potential for use in extended analyses of gene-gene (G×G) and gene-environment (G×E) interactions. The relative importance of these criteria depend on the goals of a particular study. We will introduce a novel empirical Bayes method for tests and estimates of gene-wide association and describe new ways to improve the interpretability of an alternative method using principal components. Our comparisons of these and competing methods, such as haplotype analysis, will focus on interpretability and include application to analysis of G×G interactions.

1357F

Genome-wide association study on the variation of quantitative trait as a tool to identify important genetic variants for related complex diseases, an example of CHD risk in T2D patients. *Z. Liu[1,3], Q. Qi[2], F. Hu[2], L. Qi[2], L. Liang[1,3].* 1) Department of Epidemiology, Harvard School of Public Health, Boston, MA; 2) Department of Nutrition, Harvard School of Public Health, Boston, MA; 3) Department of Biostatistics, Harvard School of Public Health, Boston, MA.

Extreme levels of underlying quantitative traits may be important to the development of complex diseases. Genome-wide association studies for quantitative traits have been focusing on the comparison of the mean levels of the traits among individual with different genotypes. Genetic variants related to the increased variation of the trait also can lead to extreme level of the quantitative trait. In this study, we use mathematical derivation and simulation to show that novel genetic variants for complex disease may be identified by giving high priority to variants associated with different magnitude of the variation of related quantitative traits. We applied this approach and carried out a genome-wide association scan for the variation of HDL using the 700,000 SNPs on Affymetrix 6.0 array typed on 1461 T2D cases (505 with CHD and 745 without CHD, 66 stroke and 145 unknown) from a T2D study involving the Nurses' Health Study (NHS, n=660) and the Health Professionals Follow up Study (HPFS, n=801). After adjusting for age, BMI and eigenvectors obtained from genotype, we found a locus, SLC2A9, on chr4p16.1 significantly associated with the variation of HDL in both NHS and the HPFS samples (meta-analysis p-value=4.2×10^{-9}). The SLC2A9 locus we identified has been shown to be associated with uric acid level which is a risk factor for T2D and CHD and might imply important gene-environment interaction, gene-gene interaction or other mechanisms underlying the genetics of HDL. We further tested the association with CHD risk in T2D and found that the top SLC2A9 SNP associated to HDL variation was associated with CHD risk in diabetic men (OR [95%CI]: 0.52 [0.29–0.94], P=0.029) but not in diabetic women (1.24 [0.74–2.08], P=0.40). Among the top 50 SNPs that have significant interaction effect with the SLC2A9 locus on HDL level in diabetic men, we found a SNP at chr16q11.2 also has significant interaction with the SLC2A9 SNP on CHD risk (p-value for interaction effect on CHD risk = 0.006, pvalue for interaction effect on HDL level = 5.59×10^{-7}) in diabetic men. In summary, we show that genome-wide association study on the variation of quantitative trait provides a promising approach to identify new loci and improve the understanding of the genetic bases of complex diseases.

1358W

Power and sample size calculations for genetic association tests in the presence of pleiotropy. *D. Londono, D. Gordon.* Department of Genetics, Rutgers The State University of New Jersey, Piscataway, NJ.

Genetic pleiotropy is a common phenomenon in studies of complex diseases. Its effect on the ability to localize genes in association studies is not yet fully understood. In this work, we investigate what factors most significantly affect statistical power of some commonly used association tests, namely the Pearson chi-square genotype test, the Cochran-Armitage Trend Test and the Transmission Disequilibrium Test, when two different but clinically related diseases are associated with a common locus. We derive the non-centrality parameter (NCP) of these tests in the presence of pleiotropy. With the NCP formulae, sample size can be determined for a given power and significance level. We estimate empirical power and type-I error rates by simulating case-control and trio data using a factorial design for different settings of: (i) linkage disequilibrium; (ii) marker and disease allele frequencies; (iii) genotype relative risk; (iv) prevalence; and (v) the proportion of individuals in the observed sample expressing the second disease phenotype. Our results show that, for all tests, the heterozygote relative risk in the second disease is the most significant factor affecting difference in power between using only individuals affected with the first disease and using both types of affected individuals. Additionally, the heterozygote relative risk in the first disease and the frequency of the disease susceptibility allele are also important factors, although the significance of their effect is different across different tests. Also, we show under which circumstances power is gained/lost by adding individuals expressing the second disease when the two diseases have different inheritance models. Finally, our power results show high degree of correlation with a previously proposed decision rule for case-control data that determines whether two similar phenotypes should be combined to improve power.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1359T

Using regression residuals to map association between genetic variants and environmentally influenced longitudinal phenotypes. *A. Musolf[1], D. Londono[1], R. Wang[2], J. Brandon[3], J. A. Herring[4], C. A. Wise[3,5,6], H. Zou[7], M. Jin[7,8], L. Yu[1,9], S. J. Finch[2], T. C. Matise[1], D. Gordon[1]*. 1) Department of Genetics, Rutgers University, Piscataway, NJ; 2) Department of Applied Mathematics and Statistics, Stony Brook University, Stony Brook, NY; 3) Seay Center for Musculoskeletal Research, Texas Scottish Rite Hospital for Children, Dallas, TX; 4) Department of Orthopedic Surgery, Texas Scottish Rite Hospital for Children, Dallas, TX; 5) Department of Orthopaedic Surgery, University of Texas Southwestern Medical Center, Dallas, TX; 6) McDermott Center, University of Texas Southwestern Medical Center, Dallas, TX; 7) Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China; 8) Shanghai Bio Corp, Shanghai, China; 9) Center of Alcohol Studies, Rutgers University, Piscataway, NJ.

We have previously developed a method that tests for association between longitudinal phenotypes and genetic variants. This method uses growth mixture models to determine longitudinal trajectory curves. The method computes also the Bayesian posterior probability (BPP) of belonging to a specific curve as a quantitative phenotype in association analyses. The method was shown to maintain correct type I error rate. We observed high power for both single disease loci and multiple disease loci in numerous simulation studies. Here, we present two major extensions to this method. The first is a correction for population stratification. Our previous method showed inflated type I error in data sets that were known to be of mixed populations. Our correction involves a cluster analysis of the data set. The probability of being in a certain cluster was used as a predictor in a linear regression, with BPP as the response variable. Residuals were then used as a quantitative phenotype for association analysis. This correction reduced inflation while still maintaining high power. We conclude that using the residuals as a quantitative phenotype is more effective at detecting association because it allows for the correction of confounding factors without sacrificing power. Second, we tested our method with simulations involving environmental covariates. Our simulations had a factorial design, with factors including noise level (variance of trajectory curves), minor allele frequency of the disease variant, and a metric representing distance among trajectory curves. Various biological factors were also considered, such as inheritance modes (additive, dominant, and recessive). Single locus and multi-locus models were also considered, as were varying penetrance levels. Finally, data was simulated for all scenarios using equations from two distinct data sets. We report that throughout all simulations our method: 1) maintains the correct empirical type I error, and, 2) produces 95% power or higher for most simulations. We conclude that our method robustly detects allelic association with longitudinal phenotypes that are partially influenced by environmental variables. Overall, we conjecture that this method can be used in practice by fellow researchers that are interested in finding association between longitudinal phenotypes and genetic variants, even in the presence of confounders such as environment and population stratification.

1360F

Enabling mixed model association analysis for large case-control studies. *A. Price[1], N. Zaitlen[1], B. Vilhjalmsson[1], T. Hayeck[1], S. Pollack[1], J. Yang[2], G.B. Chen[2], M. Goddard[3], P. Visscher[2], N. Patterson[4]*. 1) Harvard Sch Pub Hlth, Boston, MA; 2) University of Queensland, Brisbane, Australia; 3) University of Melbourne, Melbourne, Australia; 4) Broad Institute, Cambridge, MA.

Mixed linear models (MLM) are an effective approach to eliminate confounding and increase power in association studies of quantitative traits. However, existing MLM methods are not applicable to large case-control studies, because case-control ascertainment (CCA) induces correlations between unlinked causal SNPs, leading to a loss in power. Our case-control simulations in unrelated individuals indicate that the power loss of MLM increases with the ratio of sample size (N) to effective number of independent SNPs (M), with liability-scale heritability explained by genotyped SNPs (h2g) and with lower disease prevalence (f). For N/M=1, h2g=0.5, we observed an average drop in test statistics at causal SNPs of -15% at f=0.01 (-36% at f=0.001) as compared to linear regression (LR), which is equivalent to MLM with genetic relationship matrix (GRM) = I. The observed-scale heritability explained by genotyped SNPs (h2g_obs) estimated by MLM was deflated relative to its correct value, but using the correct value did not improve the power of MLM. Analysis of WTCCC phenotypes (WTCCC 2007) using subsets of the genome to vary N/M led to MLM power loss consistent with simulations. The power loss caused by including SNPs in LD with the candidate SNP (due to CCA) in the GRM is analogous to the power loss of including the candidate SNP in the GRM (Listgarten 2012), but we excluded the candidate SNP from the GRM in our analyses. Our solution to the power loss of MLM is to estimate posterior mean liabilities (PML) of each sample conditional on case-control status and liability-scale phenotypic covariance h2g*GRM+(1-h2g)*I, to account for CCA-induced LD. h2g is estimated accurately by transforming (Lee 2011) h2g_obs from H-E regression. PML are estimated via MCMC with total running time O(N^3+ITER*N^2), with ITER<N via Rao-Blackwellization. The association between genotype and PML is then tested, accounting for genetic covariance based on GRM. This method (LTMLM) does not suffer a loss in power. For N/M=1, h2g=0.5, f=0.01, we observed an increase in test statistics of +4% vs. LR. The increase in power arises by implicitly conditioning on unknown causal variants. MLM and LTMLM both had correct false-positive rate in simulations with population stratification. N/M=1 corresponds to N=60,000, since effective M~60,000 in humans (Goddard 2009), but MLM suffers a proportionate loss in power for smaller N/M. We recommend using LTMLM instead of MLM for case-control GWAS with N>10,000.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1361W

Accounting for parent-of-origin effects detects association between 4q35 genetic variants and combined asthma-plus-rhinitis phenotype. *C. Sarnowski[1,2,3], G. Malerba[4], Q. Vincent[5], C. Laprise[6], K. Rohde[7], M. Moffatt[8], P. Margaritte-Jeannin[1,2,3], M-H. Dizier[1,2,3], P.F. Pignatti[4], W.O.C. Cookson[8], M. Lathrop[3], F. Demenais[1,2,3], E. Bouzigon[1,2,3], EGEA collaborative group.* 1) U946, INSERM, PARIS,75010, France; 2) Univ Paris Diderot, Sorbonne Paris Cité, Institut Universitaire d'Hématologie, France; 3) Fondation Jean Dausset-Centre d'Etude du Polymorphisme Humain (CEPH), France; 4) Section of Biology and Genetics, Department of Mother and Child and Biology-Genetics, University of Verona, Italy; 5) U550, INSERM, PARIS, France; 6) Université du Québec à Chicoutimi, Canada; 7) Max-Delbrück-Center for Molecular Medicine (MDC), Germany; 8) National Heart Lung Institute, Imperial College, UK.

A strong linkage signal was previously detected in the 4q35 region with the combined asthma-plus-rhinitis phenotype in 640 families from European ancestry (French (EGEA), British (MRCA) and Italian) when accounting for imprinting (LOD=3.14, $P=2.5\times10^{-5}$). We further investigated this region in 161 families (206 offspring) from the French EGEA study (Epidemiological study on the Genetics and Environment of Asthma) using a panel of 1300 SNPs (spanning 6Mb). Two different methods aiming to detect parent-of-origin and/or maternal genotype effects were used to test for association between these SNPs and asthma-plus-rhinitis phenotype: 1) the Monte-Carlo Pedigree Parent-Asymmetry-Test (MCPPAT) and 2) the Parent-of-origin-Likelihood ratio Test (PO-LRT). We identified 26 markers associated with asthma-plus-rhinitis ($P\leq0.005$) among which one reached the multiple testing-corrected threshold of $P\leq6.5\times10^{-5}$. Analyses conducted with imputed data (Hapmap2) strengthened the evidence for association with three genes. In order to replicate our findings, we conducted association analysis in 152 French Canadian families (Saguenay-Lac-Saint-Jean) under the same epigenetic model detected in the discovery set. The combination of EGEA and SLSJ results using a fixed-effect model evidenced association with two SNPs tagging two different genes ($P_{comb}=9\times10^{-5}$ and $P_{comb}=9\times10^{-4}$) under parent-of-origin effect model. Further linkage analyses performed in EGEA sibships stratified according to the genotypes at each of the two most significant SNPs showed that one SNP accounted for most of the linkage signal detected in the 4q35 region. Moreover, association analyses performed separately with each allergy-related phenotypes (asthma, rhinitis, atopy) suggested that the other SNP is more likely associated to atopy ($P=5\times10^{-5}$) than the combined asthma-plus-rhinitis phenotype. Further investigation is needed to confirm our findings and to better understand the role of these loci in asthma-plus-rhinitis and their relationships with respect to allergy. The combination of these results with expression and methylation data will help pinpointing towards the functional SNPs influencing asthma-plus-rhinitis and allergy phenotypes. This study highlights that taking into account complex mechanisms facilitates the identification of new genes. Funded: French Min Education & Research, AFSSET, ANR-CEBS, ANR-CEST, GABRIEL & Région Ile de France.

1362T

Evaluation of classical multivariate methods for gene-based tests of association with multiple related traits. *Y. Shu, R. Watanabe, J. Lewinger.* Preventive Med, Univ Southern California, Los Angeles, CA.

Many statistical methods have been proposed for testing the global effect of a group of SNPs in a gene or pathway on a single trait. The premise behind these methods is that, if many SNPs in a gene have small effects, combining the evidence across them may enhance the power to detect associations. Similarly, if a complex disease exhibits a syndromic phenotype that is not fully captured by a single binary or quantitative trait, power maybe enhanced by combining the evidence across multiple related traits. However, we are unaware of any methodology specifically proposed for testing the association of a group of SNPs with multiple related traits. Here we explore a range of classical dimension reduction techniques such as principal component analysis (PCA), partial least squares (PLS), and canonical correlation analysis (CCA), coupled with multivariate regression-based methods such as MANOVA, to detect genetic association between multiple (possibly correlated) SNPs and multiple related traits. We completed an initial series of simulations in which we evaluated the power of these methods under a range of scenarios covering differing degrees of correlation among the SNPs and the traits, and also a range from high sparsity scenarios (few SNPs associated with few traits) to low sparsity scenarios (many SNPs associated with many traits). We assessed statistical significance using permutation tests, which ensures the type I error is maintained at the target level and provides objective comparison across methods. We find that applying MANOVA on latent variables for both SNPs and traits obtained from dimension reduction yields higher power to detect a global effect than raw variables in most scenarios. For a scenario (200 SNPs, 40 traits, 300 individuals) with 8% of the SNPs and traits are associated with a common linear effect of 0.8, MANOVA using latent variables from PLS has highest (90%) power, MANOVA on raw variables has 63% power, and univariate analysis on raw variables only has 16% power. Latent variables from PLS and CCA, which are derived from both SNPs and traits jointly, did not result in a significant advantage compared with latent variables derived from PCA, which are derived from SNPs and traits separately. Knowing biologic pathways are likely to involve complex non-linear relationship between genetic variants and traits, in a second set of simulations we explore performance of nonlinear kernel-based methods and compare to the linear methods above.

1363F

A two-stage random forest approach to identify genetic variants using recombination hotspot information. *S. Szymczak[1], Q. Li[1], Y. Kim[2], A. Dasgupta[3], J.D. Malley[4], J.E. Bailey-Wilson[1].* 1) Statistical Genetics Section, Inherited Disease Research Branch, National Human Genome Research Institute, NIH, Baltimore, MD; 2) Genometrics Section, Inherited Disease Research Branch, National Human Genome Research Institute, National Institutes of Health, Baltimore, MD; 3) Clinical Sciences Section, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, MD; 4) Center for Computational Bioscience, Center for Information Technology, National Institutes of Health, Bethesda, MD.

In genome-wide association studies (GWAS) each single-nucleotide polymorphism (SNP) is usually analyzed separately for association with a disease. However, it is expected that multiple genetic loci jointly contribute to disease risk. Recently, several machine learning algorithms have been used to study many or even all genome-wide genetic variants simultaneously. One promising approach is random forest (RF), an ensemble method based on a large number of classification and regression trees trained on bootstrap samples. RF provides variable importance measures that can be used to select the most relevant SNPs. However, identification of true risk SNPs that are in strong linkage disequilibrium (LD) with non-risk SNPs is challenging. Large regions of strong LD can lower the ranking of all SNPs in a causal region since they can serve as proxies for each other in different trees.

We propose a two-stage approach that uses information about recombination hotspots in the genome. In a first step, a random forest is trained for each region between two hotspots. SNPs within these regions are more likely to be in strong LD. Predicted case-control status or estimated probability of case/control status based on each region's SNPs then replaces the group of SNPs as a predictor variable. These new predictor variables for all hot spot blocks are used in a genome-wide random forest analysis in the second stage.

We compare our approach with a random forest analysis using all SNPs simultaneously based on simulated GWAS data mimicking local LD patterns observed in European samples of the 1,000 genomes project. Genotype data for 500K SNPs in 5,000 cases and 5,000 controls are generated using the software GWAsimulator. We compare the false-positive errors under a null model with no true genetic effects. For a power analysis several variants with different minor allele frequencies and in regions with different local LD structure are modeled as true risk SNPs with independent effects on the disease status.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1364W

Effect size distribution for prioritizing results of association studies. *D. Zaykin, C-L. Kuo.* Dept Biostatistics, NIEHS/NIH, Res Triangle Park, NC.

In large scale association studies, signals of detectable magnitude are rare, while majority of variants have tiny or essentially zero effect size. Because of this sparseness of effects, the proportion of false positives among most significant findings is difficult to estimate. A crucial component of such calculation is an accurate representation of the effect size distribution among variants examined in a study. We develop efficient methods for characterizing the effect size distribution. Knowledge of this distribution allows estimation of the proportion true positives in a specified number of top hits and leads to more rigorous study designs.

1365T

Modeling multivariate and correlated data in genetic association analysis. *J. H. Zhao[1], F. Xue[2], Q. Tan[3], S. Li[4], J. A. Luan[1].* 1) MRC Epidemiology Unit and Institute of Metabolic Science, Addenbrooke's Hospital Box 285, Hills Road, Cambridge CB2 0QQ, United Kingdom; 2) Department of Epidemiology and Health Statistics, School of Public Health, Shandong University, Jinan, PR China; 3) Epidemiology and Biostatistics, University of Southern Denmark, Denmark; 4) Department of Epidemiology, School of Public Health and Tropical Medicine, Tulane University, New Orleans, USA.

Background: Considering the numerous limitations with available methods in genetic association analysis, we focus on aspects related to multivariate and correlated data. Methods: We described a latent variable model to account for both multiple single nucleotide polymorphisms (SNPs) and multiple phenotypes, as with a mixed model to account for correlation among related individuals. The models were applied to simulated and real data. Results: The combination of dimension reduction and latent construct as in partial least squares path modelling offers a viable option to regional association involving genetic variant(s) as with correlated phenotypes for a latent variable. When applied to the EPIC-Norfolk study, stronger evidence of association with obesity in the FTO region was demonstrated compared to its single-SNP counterpart. Moreover, a latent score has been derived from body mass index, waist and hip circumferences as with a predisposition score from multiple SNPs. As for the mixed model, the great appeals lies in its ability to account for various sources of variations whose feasibility was demonstrated with Genetic Analysis Workshop 17 simulated data while the utility of whole genome data was shown with the Framingham data in deriving relationships among related individuals comparable to those defined by family structure. Some computational issues were also resolved in the process. Conclusion: Our work suggests that both models have many desirable properties and are preferably combined in a vigorous framework for both genetic and environmental factors involved with different types of data such as unrelated individuals from the population and relatives in families. In addition, R (http://www.r-project.org) could facilitate this greatly. References Xue F, et al. A latent variable partial least squares path modeling approach to regional association and polygenic effect with applications to a human Obesity study. PLoS One 2012; 7(2): e31927 Zhao JH, Tan Q. Integrated analysis of genetic data with R. Hum Genomics 2006; 2(4): 258–265 Zhao JH, Luan JA. Mixed modelling with whole genome data. J Prob Stat (accepted).

1366F

A novel spatial mapping method identifies shared genetic effects across immune-mediated diseases. *C. Cotsapas[1,2], J. Kasvin-Felton[3].* 1) Neurology, Yale School of Medicine, New Haven, CT; 2) Genetics, Yale School of Medicine, New Haven, CT; 3) Yale Combined Program in Biological and Biomedical Sciences, Yale School of Medicine, New Haven, CT.

Genome-wide association studies (GWAS) have identified hundreds of risk loci for autoimmune and inflammatory diseases. We have previously shown that almost half these appear to be shared across diseases and that they encode sets of interacting genes, suggesting that the architecture of pathogenesis is shared. However, comparisons of individual markers cannot conclusively distinguish between true pleiotropy, where the same genetic variant influences risk to multiple traits, and independent disease-specific variants located in the same genomic region. We can thus not formally identify the extent of disease overlap or underlying pathogenic mechanisms without resolving this question. We address this deficit with a novel spatial mapping approach, comparing association signals across two diseases. Using publicly available summary statistics, our method compares the pattern of association across an interval to determine whether it is the same in the two diseases. We define a fixed window around each marker and compute the correlation of association statistics between the two diseases in the window. We then use a wavelet-based peak-detection algorithm to find regions of the genome where correlation coefficients deviate from expectation, indicating simultaneous association. We show using simulation that the method can distinguish between cases of true pleiotropy and co-localization of independent signals. We also show that, in real GWAS data for Crohn's disease and ulcerative colitis, we detect simultaneous association at 60 loci known to harbor variants to both diseases; simple meta-analysis of these GWAS data across the two diseases only identifies 11 loci. We also identify a further 80 regions not previously known to harbor variants for both diseases, which we suggest are novel risk loci. We are currently comparing publicly available GWAS data for these and several other autoimmune/inflammatory diseases, including multiple sclerosis and type 1 diabetes, with preliminary results suggesting similar levels of overlap. The number of loci discovered suggests our method has more statistical power to detect overlaps simultaneous associations than simple meta-analysis by virtue of considering intervals rather than single markers. Our results also suggest that genetic factors for immune-mediated disease are indeed shared to a large extent; we are currently pursuing several validation and interpretation strategies to aid in identifying shared pathogenic pathways.

1367W

Adjust local ancestry to detect rare variants in next-generation sequencing data. *X. Wang, X. Zhao.* Zilber School of Public Health, University of Wisconsin-Milwaukee, Shorewood, WI.

Population stratification is a well-known issue in genome wide association studies (GWAS) in which the local ancestry at a test single nucleotide polymorphism (SNP) may confound with the association signal and ignoring it can lead to spurious association. Practical evidences and a report of Genetic Analysis Workshop 17 (GAW17) show that population stratification is expected to be a more severe issue with rare variants than with common variants in next-generation sequencing studies. In this study, we demonstrate theoretically that adjustment for local ancestry at the test collective rare variants is sufficient to remove the spurious association regardless of the mechanism of population stratification whether due to local or global ancestry differences among study subjects. We further develop a novel method to detect collective rare variants while adjusting for local ancestry. This method is based on a conditional likelihood framework which models the distribution of the collective rare variants in a considered region given disease status and flanking marker genotypes. We conducted extensive simulation studies using population genetic models. The population-specific levels of rare variation and driven stratification in association tests are affected by the cumulative effects of different sources such as growth, migration and selection. The type I error rates of our test are under control and it performs similarly or more powerful than existing methods in association tests when there exist different scenarios of stratification. The key advantage of the new method is that it is robust and very powerful when different directions of association exist in the ancestral populations.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1368T

Detecting association for low-frequency variants in case-control studies. *C. Xing, CY. Lin.* McDermott Ctr, Univ Texas SW Med Ctr, Dallas, TX.

In testing association between genetic variants and a disease, there exist popular metrics such as odds ratio and Phi coefficient to measure the strength of association, and there exist popular tests such as Pearson's chi-squared test, Fisher's exact test, and the likelihood ratio test to test the strength of association. However, these measures and tests are not sensitive to detect association for low-frequency exposures. In this study, we propose a new statistic for the purpose of detecting low-frequency variants associated with the disease. The association test based on the new statistic is more sensitive than existing methods in detecting association for low-frequency variants. We are able to numerically prove that the new statistic offers higher power while maintains the same type I error rate as Pearson's chi-squared test in a balanced study design. Application of the new method in genetic studies of detecting association of rare genetic variants with diseases is shown.

1369F

The Empirical Assessment of Statistical Power of Rare Variant Association Methods. *K. Hao[1], H. Chen[1], H. Zhou[1], Z. Kan[2], H. Zheng[3], X. Liu[3], S. Li[4], T. Barber[4], Z. Gong[3], H. Gao[3], M. Willard[4], J. Xu[2], R. Hauptschein[2], P. Rejto[2], G. Wang[3], Q. Zhang[3], R. Cheng[1], K. Sung[5,6], Z. Peng[3], C. Zhang[1], Q. Zhang[3], R. Poon[5], S. Fan[5], J. Wang[3,7,8], J. Hardwick[1,9], C. Reinhard[4], Y. Li[3], J. Luk[5], M. Mao[2,9], H. Dai[1].* 1) Merck Research Laboratories, Boston, MA; 2) Pfizer Oncology, San Diego, CA; 3) Beijing Genomics Institute, Shenzhen, China; 4) Eli Lilly and Company, Indianapolis, IN; 5) Department of Surgery, University of Hong Kong, Hong Kong; 6) School of Computing Science, National University of Singapore, Singapore; 7) Department of Biology, University of Copenhagen, Copenhagen, Denmark; 8) The Novo Nordisk Foundation Center for Basic Metabolic Research, University of Copenhagen, Copenhagen, Denmark; 9) Asian Cancer Research Group, Inc., Wilmington, DE.

Backgrounds The role of rare genetic variation in the etiology of complex disease remains as an active research field. The next-generation sequencing technologies were powerful in identifying genetic variants responsible to Mendelian disorders, but had limited success in dissecting complex diseases. Several novel statistical methodologies have been recently proposed to assess the contribution of rare variation to complex disease etiology. So far, their statistical power was assessed mainly on simulated phenotypes. **Methods** We quantified the statistical power of popular rare variants association methods empirically using the eQTL framework. That is the relative statistical power is proportional to the number of eQTLs captured at fixed FDR. **Results** On N=100 human non-tumor liver tissues, we measured gene expression trait with Affymetrix HU133-Plus2 chip, and conducted whole genome DNA sequencing (WGS) at 30X coverage. Also, we measured the genotype using Illumina 650Y array. Results of conventional single marker tests with 1000G imputation was compared to those of rare variants association methods, including cohort allelic sum test (CAST), weighted sum test (WST) and SKAT. When only focusing on SNVs of MAF<5%, none of the rare variants association methods have meaningful statistical power. When including all SNVs, rare variants association methods offer comparable power to single marker test (ie, GWAS methods) **Conclusion** WGS has no advantage in term of statistical power over SNP array in identifying eQTLs. Gene-level tests offer higher power due to less multiple testing penalties. SNVs on exons, splicing site or UTRs were more likely to control gene expression levels.

1370W

Rare Variant Burden Tests and Meta-Analysis Using Summary Level Statistics. *D. Liu, G. Abecasis.* Dept Biostatistics, University of Michigan, Ann Arbor, MI.

Studies of the contribution of rare functional variation to complex traits could elucidate the biology of many complex traits and diseases. Since most coding variation is very rare, achieving sufficient power will often require examining the average effect of multiple variants in a gene (through “gene burden tests”) and combining information across many samples. Here, we develop a framework for combining information across many studies that allows for the calculation of gene burden test statistics using summary statistics from each component study. First, we show that gene burden test statistics including the Collapsing of Rare Variants Test, the Variable Threshold Collapsing of Rare Variants Test, and Sequence Kernel Association Tests that allow variants in the same gene to have opposite directions of effect can be constructed by combining a linkage disequilibrium matrix with carefully selected single marker summary statistics. Second, we show that when information on linkage disequilibrium patterns and summary statistics are combined across several samples, our method results in “gene burden test” statistics that are equivalent to those obtained by merging raw genotype data across studies. Finally, we provide examples of the utility of our method with examples from ongoing studies of the genetics of blood lipid levels and other traits. Software implementing our methods is freely available and should facilitate analysis of rare variation that combine data across multiple sequenced or genotyped samples.

1371T

A geometric framework for the evaluation of rare variant tests of association. *N. Tintle[1], K. Liu[2], S. Fast[3], M. Zawistowski[4].* 1) Math, Stat and CS, Dordt College, Sioux Center, IA; 2) Dept. of Biostatistics, Harvard University, Cambridge, MA; 3) Massachusetts Institute of Technology, Cambridge, MA; 4) Center for Statistical Genetics, University of Michigan, Ann Arbor, MI.

The tidal wave of next-generation sequencing (NGS) data has arrived, but more questions than answers exist about how to best analyze NGS data to investigate the potential contribution of rare genetic variants to human disease. Numerous rare variant association testing methods have been proposed which attempt to aggregate association signals across multiple variant sites in an effort to increase statistical power. While emerging simulation results suggest that some rare variant testing methods work better than others for particular genetic architectures, little concrete understanding of the tests is available. We propose a geometric framework which quickly classifies existing rare variant tests of association into two broad categories: length and joint tests. We then demonstrate how genetic architecture (relative risk distribution, allele frequency distribution and number of variants) directly relates to the behavior of length and joint tests. We go on to illustrate further implications of the geometric framework including the differential impact of variant weighting strategies, population stratification and genotype uncertainty on length and joint tests. We also describe how the geometric framework suggests a lack of diversity in current approaches for rare variant testing, along with numerous potential alternative rare variant association tests and how they will behave. The geometric framework articulates the connection between disease architecture and test behavior, providing a clear set of next steps for applied and theoretical researchers.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1372F

Subset-based Approach to Combined Genetic Association Analysis of Heterogeneous Phenotypes. *S. Bhattacharjee[1,2], S.M. Boca[2], P. Rajaraman[2], K.B. Jacobs[3], W.A. Wheeler[4], B.S. Melin[5], P. Hartge[2], M. Yeager[3], C.C. Chung[6], S.J. Chanock[6], N. Chatterjee[2], GliomaScan Consortium.* 1) National Institute of Biomedical Genomics, Kalyani, West Bengal, India; 2) Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, MD; 3) Core Genotyping Facility, National Cancer Institute, Gaithersburg, MD; 4) Information Management Services, Rockville, MD; 5) Department of Radiation Sciences, Oncology, Umeå University, Umeå, Sweden; 6) Laboratory of Translational Genomics, DCEG, National Cancer Institute, Bethesda, MD.

Combining GWAS of related traits can provide promising directions for the discovery of loci with small but common pleiotropic effects. Here, we study an agnostic, subset-based approach for conducting meta-analysis of heterogeneous genetic association studies. It involves exploring the subset space to identify the strongest association signal and then appropriately evaluating the statistical significance of the signal accounting for the multiple testing incurred due to subset search. We consider pooled case-control studies of heterogeneous traits as well as analysis of heterogeneous disease subtypes of a single case-control study. To avoid computationally intensive resampling-based procedures, we develop fast analytic approximations for the p-values of our proposed statistics. We demonstrate using simulations and real data applications that our methods retain substantial advantage in terms of power and interpretation over standard meta-analysis approaches. Further, we study extensions of subset-based methods to analyze SNP data across a gene or region. These methods can take into account the linkage disequilibrium structure among SNPs and detect overall gene-level pleiotropic association with multiple related traits.

Subset-based association analysis of some known cancer GWAS hits with 6 Cancers

SNP	Primary (Known)	Genes (Chr)	P-value	Secondary (Risk)	Secondary (Protective)
rs401861	Basal Cell	CLPTM1L (5)	5.99E-08	Pancreas	Bladder, Lung
rs2660753	Prostate	LOC285232 (3)	3.17E-03		Breast, Kidney
rs10993994	Prostate	MSMB (10)	8.62E-03	Bladder, Kidney	
rs6983267	Prostate, Colorectal	POU5F1B (8)	1.08E-02	Breast	Bladder, Kidney, Lung, Pancreas

1373W

Genome wide significance thresholds for SNPs ascertained by sequencing. *T. Blackwell[1], A.K. Manning[2], GoT2D Investigators.* 1) Center for Statistical Genetics, Univ Michigan, Ann Arbor, MI; 2) Broad Institute, Cambridge MA.

Resequencing projects including 1000 Genomes have greatly expanded the number of SNPs and the range of allele frequencies available to genome wide association studies. Previously, the HapMap catalogue constituted the universe of SNPs one could hope to test for association, and one could set a p-value threshold for genome wide significance that was invariant to a study's sample size. Most newly identified SNPs have minor allele frequencies (MAF) well below those of HapMap sites, and linkage disequilibrium (LD) generally decreases with MAF. Low MAF also introduces a dependence on the study's overall sample size, since a minimum number of observations of the minor allele are needed before it is possible to reach genome wide significance using a case control test for a single variant.

We quantify both effects using the June 2011 whole genome data freeze from GoT2D as an example. Permuting case control status gives a valid 5% experiment wide significance threshold which accounts for LD. With 1510 cases and controls combined, a conventional 5×10^{-8} genome wide threshold is appropriate for testing all 4.3 M SNPs with 0.11 MAF and above in this study. A threshold of 2×10^{-8} is required for testing the 8.5 M SNPs down to 0.01 MAF. If the study sample size were expanded tenfold to 15,100 cases and controls, without ascertainment of additional variants, extrapolation using an analysis from Moskvina and Schmidt (2008) shows that a threshold of 1×10^{-8} would suffice for testing 12.5 M SNPs down to 0.001 MAF.

One can measure the redundant information due to LD using the ratio of the actual number of SNPs tested to the number of independent tests that would produce the same empirical threshold. The sites above 0.11 MAF show an average fold redundancy 4.3. This decreases linearly with log allele frequency to an average of 1.5 fold redundancy among the SNPs at 0.001 MAF. These results are restricted to single variant tests for case control outcomes and refer to an ethnically homogeneous sample from northern Europe. Added ethnic diversity will almost certainly increase the total number of SNPs at each MAF and reduce the average redundancy. This argues for using a study-specific threshold for genome wide significance. Gene level `burden' tests would show quite different behavior, since the number of tests is invariant to the number of SNPs and to a study's sample size.

1374T

On the analysis of imputed genotypes in family-based association studies. *A. Cobat[1,2], A. Alcais[3], E. Schurr[1,2].* 1) McGill Centre for the Study of Host Resistance, Montreal, QC, Canada; 2) Human Genetics Dept, McGill University, Montreal, QC, Canada; 3) Laboratory of Human Genetics of Infectious Diseases, INSERM U980, Paris, France.

Genotype imputation is an efficient statistical approach for inferring genotypes at ungenotyped variants. However, because of the probabilistic nature of imputed SNPs specific techniques are needed for their analysis to account for genotypic uncertainty. One of the most popular approach relies on the expected allele count E1=2P11+P12, also called allele dosage, to test for association1. In population-based designs, testing of association between the phenotype and expected allele count (instead of genotype) is straightforward using standard linear or logistic regression frameworks. By contrast, equivalent methods for family-based studies are missing. To close this gap, we developed a new analytical approach which is based on the methodology proposed for the analysis of copy-number variants2 to perform association testing of imputed SNPs in the family-based design (FBAT-dosage). We further extended our method to combine the family-based statistic with a paired test statistic for founders when parental phenotypes are available (FBAT-dosage_c). Simulation studies under the null hypotheses of “no linkage, no association” and “linkage, no association” showed that the FBAT-dosage and FBAT-dosage_c methods provide very consistent type I errors, whatever the sibship size and the level of certainty with which the marker under study was imputed. An alternative to our approach is the Best-guess genotype approach which does not allow testing for association between the phenotype and SNPs imputed with low certainty. Hence using FBAT-dosage will provide substantial increase in power for most scenarios. For SNPs imputed with certainty (or observed SNPs), the FBAT-dosage statistic is equal to the FBAT statistic for genotypes using the empirical variance3. Finally, the FBAT-dosage method and its extension have been implemented in C++ language and are suitable for genome-wide imputation analysis. 1. Marchini J, and Howie B. Nat Rev Genet. 2010 Jul;11(7):499–511. 2. Ionita-Laza I et al. Genet Epidemiol. 2008 Apr;32(3):273–84. 3. Lake S, et al. Am J Hum Gen. 2000. 67:1515–1525.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1375F

Sex-specific and X-chromosome Association Studies of Venous Thromboembolism (VTE). *M. de Andrade[1], S.M. Armasu[1], L.L. Chan[1], J.A. Heit[2].* 1) Div Biomed Statistics & Informatics, Mayo Clinic, Rochester, MN; 2) Div Cardiovascular Diseases, Mayo Clinic, Rochester, MN.

Background: VTE is a complex disease (h^2~0.6), and likely results from multigenic action as well as environmental exposures. **Objective:** To identify autosomal and X-chromosome VTE-susceptibility genes, both overall and by sex. **Methods:** Genome-wide scan (Illumina 660W Quad [530,865 SNPs]) genotypes from 1270 non-Hispanic adults of European ancestry with objectively-diagnosed VTE and 1302 controls (frequency-matched on case age, sex, race, MI/stroke status) were imputed to about 9.2 million SNPs with BEAGLE using EUR 1000G. Genome-wide association analyses were performed on autosomal and X-chromosome SNPs, both overall and by sex, using PLINK. **Results:** For the autosomal females analysis, after adjusting for age, state of residence and stroke/MI status, 3 SNPs on chromosome 1q24.2 (including *F5* rs6025 [Factor V Leiden], odds ratio [OR]=3.2, p=2.5E-9) and 38 on chromosome 9q34.2 (including *ABO* rs2519093, OR=1.7, p=3.7E-9; and rs495828, OR=1.7, p=3.5E-9) exceeded genome-wide significance. Similar results were observed for males (*F5* rs6025, OR=4.1, p=3.0E-12; *ABO* rs2519093, OR=1.9, p=1.4E-10; and rs495828, OR=1.7, p=2.7E-8). For the X-chromosome overall analyses, the most significant SNPs were near 4 pseudo-genes and *FAM46D* (rs12688059, OR=1.48, p=7.07E-5); for females only, near *FAM47A* (chrX:34138682, OR=0.048,p=2.54E-5) and *TEX11* (chrX:69992070, OR=2.81, p=6.20E-5); and for males only, near *SCML2* (rs58926087, OR=0.74, p=8.14E-5). **Conclusion:** Autosomal and X-chromosome VTE-susceptibility genes vary, overall and by sex, consistent with the hypothesized multigenic action.

1376W

Correcting for the effect of cryptic relatedness and population structure among cohorts in meta-analysis of GWAS. *T. Feng, N. Morris, X. Zhu.* Epidemiology & Biostat, Case Western Reserve Univ, Cleveland, OH.

Genome-wide association study (GWAS) is a main approach to detect genetic variants underlying complex traits, but is often conducted through meta-analysis of the summary statistics from multiple studies and assumes that the samples from different studies are independent. In practice, genome control procedure is often applied at both individual study level and meta-analyzed result in order to control the cryptic relatedness and population structure within and between studies. However, this procedure may fail to control type I error and reduce statistical power. Here, we present a new meta-analysis method to correct the effect of cryptic relatedness and population structure by estimating the correlation structure among cohorts using the GWAS summary statistics. We conducted a simulation study using the Welcome Trust Case control Consortium data (WTCCC). Our results showed that the proposed meta-analysis method can control the type I error and preserve statistical power well but the genome control procedure failed to control type I error. We will also present the results in the BP GWAS in Family Blood Pressure Program using the proposed method.

1377T

Genome-wide association study (GWAS)-identified genetic variants for lipid traits are associated with gallstone disease in the diverse Third National Health and Nutrition Examination Survey. *R. Goodloe[1], K. Brown-Gentry[1], N. Gillani[1], H. Jin[1], P. Mayo[1], M. Allen[1], B. McClellan[1], J. Boston[1], C. Sutcliffe[1], N. Schnetz-Boutaud[1], H. Dilks[1,2], D. Crawford[1,2].* 1) Center for Human Genetics Research, Vanderbilt University, Nashville, TN; 2) Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, TN.

Gallstone disease (GD) is one of the most common digestive disorders, affecting more than 30 million Americans. Previous twin studies suggest a heritability of 25% for gallstone formation. To date, one genome-wide association study (GWAS) has been performed in a population of European-descent. Several candidate gene studies have been performed in various populations, but most have been inconclusive. Given that gallstones consist of up to 80% cholesterol, we hypothesized that common genetic variants associated with high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), and triglycerides (TG) would also be associated with gallstone risk. To test this hypothesis, we as the Population Architecture using Genomics and Epidemiology (PAGE) study and Epidemiologic Architecture for Genes Linked to Environment (EAGLE) performed tests of association between 49 GWAS-identified lipid trait SNPs and GD in the Third National Health and Nutrition Examination Survey (NHANES III; n= 7,159). NHANES III is a population-based cross-sectional study of Americans ascertained regardless of health status by the National Center for Health Statistics at the Centers for Disease Control and Prevention. NHANES III is diverse with non-Hispanic whites (NHW; n=2631), non-Hispanic blacks (NHB; n=2,108), Mexican Americans (MA; n=2,073), and Asians (A; n=348). GD was defined by “yes” to the question “Has the doctor told you had gallstones?” or a positive ultrasound (defined as presence of scar tissue (surgical removal) or demonstration of gallstones present in the gallbladder). A total of 379 NHW, 137 NHB, 227 MA, and 28 A cases and 2251 NHW, 1971 NHB, 1846 MA, and 320 A controls were identified. Tests of association were performed using logistic regression, assuming an additive genetic model stratified by race/ethnicity and adjusted sex, age, and body mass index. Most associations were observed among NHB, with 2 HDL-C, 4 LDL-C, and 1 TG-identified SNPs associated with GD at p<0.05. Among all other groups, we identified at least one significant finding for each race-ethnicity (4 total associations: 2 NHW; 1 MA; 1 A). The most significant finding was rs6756629, previously associated with LDL-C, in NHW (odds ratio=1.518; 95% confidence interval= (1.137–2.026); p=0.0046; coded-allele frequency (CAF) =0.07). Overall, we demonstrate that lipid-identified genetic variants are associated with GD in this US representative population.

1378F

A unified framework for testing for genetic associations integrating environmental exposures. *S.S. Han, P.S. Rosenberg, N. Chatterjee.* National Cancer Institute, Bethesda, MD.

Recent genome-wide association studies (GWAS) have discovered over 6,000 susceptibility loci for more than 125 complex human traits. Despite this success, the common genetic variants detected through GWAS can explain only a small fraction of the heritability of these traits; this has led to challenges in dissecting the full genetic architecture of complex diseases. One major barrier to the current genetic association study using case-control data is the lack of statistical methods with full consideration of environmental exposure effects. A standard approach for incorporating exposures is to include them as covariates to a logistic regression. A major limitation of this approach is that it doesn't allow heterogeneous genetic effects across exposure levels—which is commonly observed. One existing approach that could potentially deal with this limitation is to introduce interaction terms to a logistic regression, and to jointly test for association and interaction. A drawback of this approach is that it inevitably increases degrees of freedom by testing additional interaction parameters, leading to loss of power. In addition, this method also often leads to association findings that lack biological interpretations in which the directions of genetic effects are reversed by exposure levels. Another fundamental barrier to genetic association studies is the lack of statistical methods capable of employing an appropriate disease risk model based on evidence from data. Current methods rely on a specific risk model, a multiplicative risk model via logistic regression, which is not based on evidence from data or biological explanations, but on statistical convention and convenience. In this study, we propose a novel framework for testing genetic associations that integrates the effects of environmental exposures; a class of score tests is constructed by imposing weights on genotypes by exposure levels, allowing heterogeneous genetic effects by exposures, and also leading to the detection of genetic associations whose directions are invariant with respect to exposure levels. The proposed method unifies a class of underlying disease risk models by maximizing suitable score statistics over a range of disease risk models including multiplicative, supra-multiplicative, additive and liability threshold risk models. We illustrate our method by applying it to National Cancer Institute lung cancer GWAS data in consideration of gene by smoking interactions.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1379W

Efficient association analysis for groups of genetic markers that avoids confounding by genetic structure. *D. Heckerman, C. Lippert, J. Listgarten.* Microsoft Res, Los Angeles, CA.

Approaches for testing groups of variants for association with complex traits are becoming critical. Examples of groups typically include a set of rare or common variants within a gene, but could also be variants within a pathway or any other set. These tests are important for aggregation of weak signal within a group, allow interplay among variants to be captured, and also reduce the problem of multiple hypothesis testing. Unfortunately, these approaches do not address confounding by, for example, family relatedness and population structure, a problem that is becoming more important as larger data sets are used to increase power. We introduce a new approach for group tests that can handle confounding, based on Bayesian linear regression, which is equivalent to the linear mixed model. The approach uses two sets of covariates (equivalently, two random effects), one to capture the group association signal and one to capture confounding. We also introduce a computational speedup for the two-random-effects model that makes this approach feasible even for extremely large cohorts, whereas it otherwise would not be. Application of our approach to richly structured GAW14 data, comprising over eight ethnicities and many related family members, demonstrates that our method successfully corrects for population structure, and application of our method to WTCCC Crohn's disease and hypertension data demonstrates that our method recovers genes not recoverable by univariate analysis, while still correcting for confounding structure.

1380T

A non-threshold region-specific method for detecting rare variants in common diseases. *A.R. Hsieh, C.C. Chang, C.S.J. Fann.* Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan.

To date, rare variants (with MAF < 1% or 5 %) have a proven role in some complex diseases. Owing to the low frequency and large number of rare variants, many statistical methods proposed for the detection of common variants associated with diseases are not suitable and have low power in the case of rare variants. Furthermore, some limitations exist in previous developing collapsing-based methods for testing rare variants. Accordingly, we developed a region-specific method that do not use the threshold for defining rare variants and take the directions of effects into account. Our region-specific method used the concept of weighting variants according to their minor allele frequencies and odds ratios (OR) to combine effects of common and rare variants on disease occurrence into a single score, and provided a test statistic in assessing the significance of the score. To evaluate the performance of our method, we simulated extensively under different disease models and effect sizes according to Basu and Pan (2011). We found that the power of our method increased as the effect sizes increased. The power of our method (86.22% on average) was very good in dominant and additive models. The type I error of our method was controlled well in spite of the simultaneous variations. Moreover, we also compared the performance of detecting associated signals between our proposed method, and Combined Multivariate and Collapsing (CMC) proposed by Li and Leal (2008). Results from our method showed a 20% increase in power comparing with CMC (71.85% vs 53.18%) for recessive models. For other models, the two methods were compatible. We applied our method to a publicly available dataset of Diabetic Nephropathy disease downloaded from the database of Genotypes and Phenotypes (http://www.ncbi.nlm.nih.gov/gap). We confirmed previous findings. Furthermore, we identified the association between Diabetic Nephropathy disease and COL23A1 (collagen, type XXIII, alpha 1) which was originally implicated by bioinformatics studies using other diabetic related data sets. The COL23A1 gene included 194 SNPs and 9 of them were rare SNPs. Although these SNPs were not significant (p-value <10 -5) by using the Trend test, they were detected by our method. In this study, we showed that using rare variant analysis is possible to assist researchers in gaining more insight into identifying genetic risk factors for complex diseases.

1381F

Joint admixture and association test on African American inflammatory bowel disease identifies novel significant loci. *C. Huang[1,2], S. Kugathasan[3], D.P. McGovern[4], J.H. Cho[5], R.H. Duerr[6], J.D. Rioux[7,8,9], M.S. Silverberg[10,11], T. Dassopoulos[12], L.W. Datta[1], L.W.H. Kao[2], S.R. Brant[1,2].* 1) Lyn P. and Harvey M. Meyerhoff Inflammatory Bowel Disease Center, Dept of Medicine, Johns Hopkins University School of Medicine, Baltimore, MD; 2) Dept of Epidemiology, Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD; 3) Dept of Pediatrics, Emory University, Emory, GA; 4) Inflammatory Bowel and Immunology Research Institute, and Medical Genetics Institute, Cedars-Sinai Medical Center, Los Angeles, CA; 5) Dept of Medicine and Genetics, Yale University, New Haven, CT; 6) Dept of Medicine, and Dept of Human Genetics, University of Pittsburgh, Pittsburgh, PA; 7) Depts of Medicine, Surgery, Public Health Sciences, Immunology, and Molecular and Medical Genetics, University of Toronto, Toronto, Canada; 8) Samuel Lunenfeld Research Institute and Mount Sinai Hospital, Toronto, Canada; 9) Toronto General Hospital Research Institute, Toronto, Ontario, Canada; 10) Université de Montréal and the Montreal Heart Institute, Research Center, Montreal, Quebec, Canada; 11) Broad Institute of MIT and Harvard, Cambridge, MA; 12) Gastroenterology Division, Washington University School of Medicine, St. Louis, MO.

Inflammatory bowel disease (IBD) is a disorder of immune dysregulation causing chronic idiopathic inflammation of the gastrointestinal tract. More than 160 loci have been recently confirmed in white populations using the “Immunochip,” a custom Illumina array designed to fine map and deep replicate GWAS signals for 12 autoimmune and inflammatory diseases including IBD. The array also contains ancestry informative SNPs located throughout the genome. We initiated evaluation of IBD loci using the Immunochip in African Americans (AA), a recent admixed population with 80% West African and 20% European global ancestry. Information from local admixture mapping is rarely incorporated into genomewide evaluations to increase power. We examined a joint admixture and association test described by Shriner et al (2011) that uses admixture posterior probabilities as prior information in association mapping under generalized linear regression framework, enabling incorporation of other covariates. This method also capitalizes on the reduced test burden of admixture mapping relative to association mapping. We genotyped our initial study sample of 722 AA IBD cases and 317 healthy controls for 181,268 autosomal SNPs using the Immunochip. Local and global admixture was estimated using WINPOP (Pasaniuc et al 2009). No SNP reached significance of $p < 2e\text{-}6$ with traditional logistic regression after adjusting for global ancestry and significant principal components. We employed a routine to use the Shriner joint test and identified four significant SNPs on 3 loci: 2 SNPs on 1q23, 1 on 8q22, and 1 on 22q13. These four SNPs showed only modest association by logistic regression ($5e\text{-}5 < p < 1e\text{-}3$). The 1q23 and 8q22 loci are novel IBD loci, and the 22q13 locus was previously established in European IBD (Franke et al. 2010). Global West African ancestry differed little between cases (80.7%) and controls (81.6%). However, West African ancestry about the 22q13 locus was significantly lower for cases vs. controls (76.4% vs. 80.2%, respectively, p=3e-4). Our findings highlight the value of local admixture adjustment, use of Bayesian analytic approaches and evaluation for novel IBD loci in the understudied, admixed African American population.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1382W

Comparison of phylogenetic and haplotype methods for the study of genotype/phenotype association in genome-wide studies. *M.G. Johnson[1], D.L. Swofford[1], M.W. Lutz[2], D.G. Crenshaw[2], A.D. Roses[2].* 1) Duke University, Biology Department, Box 90338 Durham, NC. 27708; 2) Duke University, Neurology Department, Box 3503 Durham, NC 27705.

Numerous approaches have been proposed to test whether polymorphisms in a genomic region are associated with phenotypic variation. Some of these approaches are based on the principle that evolutionarily closely related haplotypes will share mutations that are important in the expression of phenotype. Assessment and comparison of these methods for detection power and specificity requires a computational framework that allows rapid inclusion of new methods and supports a spectrum of realistic genetic models. Datasets were generated by forward-time simulation, beginning with the HapMap genomes, followed by population expansion under slight negative selection. Case-control datasets were generated from these populations under an additive model with five levels of penetrance. Six analysis methods were analyzed: basic chi-squared tests, ancestral recombination graphs, perfect phylogenies, tree disequilibrium tests, haplotype clustering, and phylogenetic clustering. An analytical pipeline was developed to easily replicate simulation datasets under various test conditions, and generate input files for the analysis programs. Each method was assessed for power (ability to detect the causal marker) and specificity (number of false positives both near and far from the causal marker). All six methods were able to recover an association at the true causal marker in 95% of trials when the Genotype Relative Risk (GRR) was 2.0, even with the strictest score thresholds. The methods were less successful with lower GRR, although the tree-disequilibrium test and phylogenetic clustering were most successful at moderate risk levels. None of the methods were able to consistently recover the true marker at GRR = 1.25. No method found false associations far away from the causal locus, but many markers crossed even the strictest significance threshold for each program near the disease locus. Phylogenetic clustering had by far the fewest false positives, while the tree disequilibrium method had the most. Several methods were able to exceed basic chi-squared tests in detection power both with and without a typed causal marker. However, many of these methods suffered from low specificity; many markers near the true locus crossed a significance threshold. Therefore, use of any method requires a comprehensive understanding of the thresholds at which associations should be accepted. One method, phylogenetic clustering, outperforms the other methods in both power and specificity.

1383T

Benefits of using local whole exome reference panels for the imputation of rare variants in two European populations. *P. Joshi[1], R.M. Fraser[1], V. Vitart[2], C. Hayward[2], R. McQuillan[1], O. Polasek[1], S.H. Wild[1], N. Hastie[2], A.F. Wright[2], H. Campbell[1], I. Rudan[1], C. Haley[2], P. Navarro[2], J.F. Wilson[1].* 1) Centre for population health Sciences, University of Edinburgh, Edinburgh, United Kingdom; 2) MRC Human Genetics Unit, University of EdinburghWestern General Hospital, Crewe Road, Edinburgh, UK.

Increasing focus is being put on the potential for rare variants to explain a significant portion of the heritability of complex diseases, which are a substantial health burden in western countries. While whole genome sequencing remains expensive, imputation provides a cost effective method of inferring rare variants on the basis of existing high density genome-wide SNP genotyping. This study examines whether the use of a local reference panel of whole exome sequences can increase the accuracy of rare exonic variant imputation compared with the use of the 1000 Genomes reference panel alone, in the ORCADES and CROATIA-KORCULA genetic isolate studies. 90 Orcadian and 91 Korculan samples were whole exome sequenced and genotyped using the Illumina Hap300 or 370CNV arrays, and ~200k common autosomal array SNPs were used for imputation. A local reference panel was created from the exomes and the array genotypes together and this was phased using SHAPE-IT. The SNP panel genotypes were phased using SHAPE-IT and then each subject was imputed with the array genotypes only, using IMPUTE2. The imputations used (a) the 1000 Genomes reference panel only(500 Europeans and 1221 non-Europeans), and (b) the 1000 Genomes alongside the local panels of ~90 exomes in each population, however the subject being imputed was excluded from the reference panel used in their imputation, to avoid giving IMPUTE2 the actual genotypic values directly. The accuracy of the two imputations was measured against the rigorously Quality-Controlled exomes. We examine the differences in accuracy by allele frequency and population. This study gives an indication of the improvement in accuracy of using local genome sequences to supplement a general population reference panel, when imputing genotypes for rare variants using genome-wide SNP data.

1384F

Combined linkage and association analyses identify a novel locus for obesity near PROX1 in Asians. *H. Kim[1], Y. Yoo[2], Y. Ju[3], S. Lee[1], S. Cho[4], J. Sung[4], J. Kim[1], J. Seo[1].* 1) Department of Biomedical Sciences, Seoul National University Graduate School, Seoul, Republic of Korea; 2) Department of Mathematics Education, Seoul National University, Seoul, Republic of Korea; 3) Genomic Medicine Institute (GMI), Medical Research Center, Seoul National University, Seoul, Republic of Korea; 4) Seoul National University School of Public Health, Seoul, Republic of Korea.

Although genome-wide association studies (GWAS) have substantially contributed to understanding the genetic architecture, unidentified variants for complex traits remains an issue. One of the efficient approaches is to improvement of the power of GWAS scan by weighting P-values with prior linkage signals. To identify new susceptibility loci for obesity, we performed a family-based GWAS combined with a genome-wide linkage study in an isolated Mongolian sample of 1,049 individuals from 74 families. We obtained linkage information for body mass index (BMI) and waist circumference (WC) via a multipoint genome-wide linkage scan. For both BMI (LOD = 3.3) and WC (LOD = 2.6), the highest linkage peak was discovered at chromosome 10q11.22. Next, a family-based GWAS, which integrates within- and between-family components, was performed using the genotype data of 756 individuals of the Mongolian sample, and P-values for association were weighted using linkage information obtained previously. Six single-nucleotide polymorphisms (SNPs) for BMI and five SNPs for WC reached a significant level of association (P-value<1^10-5). Of these, only one of the SNPs associated with WC (rs1704198) was replicated in 327 Korean families comprising 1,301 individuals; this SNP was located in the proximity of the prospero-related homeobox 1 (PROX1) gene, the function of which was validated previously in a mouse model. Our powerful strategic analysis enabled the discovery of a novel candidate gene, PROX1, associated with WC in an Asian population.

1385W

A more powerful burden tests to detect Rare Genetic Variants. *J. Kim[1], J. Lee[1], S. Choi[2], M.S. Kwon[2], T. Park[1].* 1) Department of Statistics, Seoul National University, Seoul, South Korea; 2) Interdisciplinary program in Bioinformatics, Seoul National University, Seoul, South Korea.

Analysis of rare variants, typically defined to have a minor allele frequency (MAF) less than 1%, is not straightforward due to data sparsity. To overcome this sparsity problem, a variety of burden tests have been proposed such as Sequence Kernel Association Test (SKAT) (Wu et al., 2011), Gene- or Region-based ANalysis of Variants of Intermediate and Low frequency (GRANVIL) (Morris and Zeggini, 2010), Hoffman's method (Hoffman et al., 2010) and so on. The objective of burden tests is to identify association with complex traits and genetic variants efficiently, not through traditional association analysis of single nucleotide polymorphisms (SNPs), but through condensed analysis of rare variants together combined into genes or regions in a reasonable manner. The main goal of our work is to propose a new burden test based on the mixed effects model which has a more flexible structure of random effects. The proposed test is compared to other burden tests via power and Type 1 error using the exome sequence data set of Genetic Analysis Workshop 17. Our comparison study shows that the burden tests based on the mixed effects models tend to perform better than other tests.

1386T

Efficiently Identifying Significant Associations in Genome-wide Association Studies. *E. KOSTEM, E. ESKIN.* Computer Science Dept., UCLA, LOS ANGELES, CA.

The genome-wide association study (GWAS) is a widely used method for locating genomic regions that are associated with complex disease traits. In a GWAS, single nucleotide polymorphisms (SNPs) are collected across the genome and tested individu- ally to identify significant associations, which may provide insight into the genetic basis of the disease. Association studies attempt to collect information on as many SNPs as possible to cover the whole genome. However, as the number of collected SNPs increases so does the computational burden to identify the significant associations. In expression quantitative trait loci (eQTL) studies, tens of thousands of gene expres- sion levels are measured and the GWAS approach is applied to each gene expression level. This leads to computing billions of tests and requires substantial computational resources. We introduce a novel two-stage testing procedure that identifies all of the significant associations more efficiently than testing all the SNPs. In the first-stage a small number of informative SNPs, or proxies, across the genome are tested. Based on their observed associations, our approach locates the regions which may contain significant SNPs and only tests additional SNPs from those regions. We show through simulations and analysis of real GWAS datasets that the proposed two-stage procedure identifies the significant associations approximately ten times faster than testing all the SNPs. Our method and software packages to perform rapid genome-wide association testing are available at http://genetics.cs.ucla.edu.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1387F

A multi-SNP locus-association method reveals a substantial fraction of the missing heritability. *Z. Kutalik[1,2], G. Ehret[3,4], D. Lamparter[1,2], C. Hoggart[5], J. Whittaker[6], J. Beckmann[1,7], GIANT consortium.* 1) Med Gen, Univ Lausanne, Lausanne, Switzerland; 2) Swiss Institute of Bioinformatics, Switzerland; 3) Division of Cardiology, Geneva University Hospital, Geneva, Switzerland; 4) McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University, Baltimore, Maryland, United States of America; 5) Department of Pediatrics, Imperial College London, London, United Kingdom; 6) Quantitative Sciences, GlaxoSmithKline, Stevenage, UK; 7) Service of Medical Genetics, Centre Hospitalier Universitaire Vaudois, Lausanne, Switzer- land.

There are many known examples of multiple (semi-)independent associations at individual loci, which may arise either because of true allelic heterogeneity or imperfect tagging of an unobserved causal variant. This phenomenon is of great importance in monogenic traits but has not yet been systematically investigated and quantified in complex trait GWAS. We describe a multi-SNP association method that estimates the effect of loci harbouring multiple association signals using GWAS summary statistics. Applying the method to a large anthropometric GWAS meta-analysis (GIANT), we show that for height, BMI, and waist-hip-ratio (WHR) 10%, 9%, and 8% of additional phenotypic variance can be explained respectively on top of the previously reported 10%, 1.5%, 1%. The method also permitted to substantially increase the number of loci that replicate in a discovery-validation design. Specifically, we identified in total 263 loci at which the multi-SNP explains significantly more variance than the best individual SNP at the locus. A detailed analysis of multi-SNPs shows that most of the additional variability explained is derived from SNPs not in LD with the lead SNP suggesting a major contribution of allelic heterogeneity to the missing heritability.

1388W

Leveraging Input and Output Structures For Joint Association Mapping of Epistatic and Marginal eQTLs. *S. Lee, E. Xing.* School of Computer Science, Carnegie Mellon University, Pittsburgh, PA.

Interdependent genetic variations affect gene regulations, and they make the outcomes of many complex diseases and expression phenotypes. Thus, it is essential for association mapping to take into account both additive and epistatic (interaction) effects of multiple candidate genotypes. However, it poses computational challenges as it requires massive computations and a very large sample size. Fortunately, recent development of genetic interaction networks has generated prior knowledge about complex interactions between genotypes, and they can be used to reduce the computational complexity and increase the statistical power of association mapping. We propose a novel computational method called structured input-output lasso that uses prior information on structures on both genetic variations (input structure) and expression phenotypes (output structure). Our method makes use of a novel regularization scheme in a general regression setting, which enables us to find multiple related SNPs (e.g. SNPs in LD) which are jointly associated with multiple related expression traits (e.g. co-expressed genes). In addition, we generalize our model to detect SNPs having epistatic effects with manageable computational complexity by incorporating prior knowledge from genetic interaction networks. We show that our method can effectively use prior information for better search of association SNPs. We also present yeast eQTLs that we identified with the analysis of their biological functions.

1389T

Evaluating the power of single variant association tests for low frequency variants in joint and meta-analysis. *C. Ma, T. Blackwell, M. Boehnke, L.J. Scott, GoT2D investigators.* Biostatistics and Center for Statistical Genetics, University of Michigan, Ann Arbor, MI.

The case-control association testing method for common genetic variants (minor allele frequency [MAF] $\geq$ 5%) is well established: test individual variants using logistic regression and combine results across multiple studies using fixed-effects meta-analysis to increase statistical power. For common variants, meta-analysis has comparable power to joint analysis of the combined individual-level data (Lin & Zeng, 2010). However for less common variants (MAF < 5%), it remains to determine: (1) the single variant test with the greatest power and best calibration (i.e. least conservative type I error rate below the nominal significance threshold), and (2) the relative power of joint and meta-analysis. In this study, we use analytic calculation and computer simulation to compare the type I error rate and power of four logistic regression tests (Wald, score, likelihood ratio, and Firth bias-reduced) in both joint and meta-analysis for less common variants. Using preliminary low-pass whole-genome sequencing data for 499 type 2 diabetes cases and 409 controls from the GoT2D study, we also compare joint analysis of the combined dataset, to meta-analysis of the three cohorts contributing to the study.

Analytic and simulation results demonstrate that for joint analysis of less common variants, the Firth test consistently has the best-calibrated type I error rate in both balanced (equal numbers of cases and controls) and unbalanced studies. For meta-analysis of less common variants in balanced studies, the score test is best-calibrated while the Firth test is conservative. However, for meta-analysis of unbalanced studies, all statistical tests can be anti-conservative. Power correlates roughly with the relative conservativeness of each test. The GoT2D data analysis results are consistent with the analytical and simulated type I error and power findings. In conclusion, if individual-level data are available, for association testing of less common variants in both balanced and unbalanced studies, we recommend joint analysis using the Firth test. If individual-level data are unavailable, for meta-analysis of balanced studies we recommend the score test. For meta-analysis of unbalanced studies, further work is required to identify a well-calibrated statistical test.

1390F

One Thousand Genomes Imputation in the NCI Breast and Prostate Cancer Cohort Consortium (BPC3) Aggressive Prostate Cancer Genome-wide Association Study. *M.J. Machiela[1], C. Chen[1], L. Liang[1], WR. Diver[2], VL. Stevens[2], KK. Tsilidis[3], SJ. Chanock[4], DJ. Hunter[1], P. Kraft[1], National Cancer Institute Breast and Prostate Cancer Cohort Consortium.* 1) Program in Molecular and Genetic Epidemiology, Harvard School of Public Health, Boston, MA; 2) Epidemiology Research Program, American Cancer Society, Atlanta, GA; 3) Cancer Epidemiology Unit, Nuffield Department of Clinical Medicine, University of Oxford, Oxford, UK; 4) Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD.

Genotype imputation substantially increases the number of markers available for analysis in a genome-wide association study (GWAS) by leveraging information on linkage disequilibrium from a reference panel. We sought to (i) investigate the performance of imputation from the August 2010 release of the 1000 Genomes Project (1000GP) in a GWAS of aggressive prostate cancer (PCa), (ii) look for novel associations with PCa risk, (iii) fine-map known PCa susceptibility regions using two complementary approaches (an approximate Bayesian framework and stepwise regression), and (iv) compare power and efficiency of imputation and de novo sequencing designs for future genetic association studies of low-frequency and rare variants. We used a collection of 2,782 aggressive PCa cases and 4,458 controls from the NCI Breast and Prostate Cancer Cohort Consortium (BPC3) aggressive prostate cancer GWAS to infer 5.8 million well-imputed autosomal single nucleotide polymorphisms from the 1000GP reference panel. Imputation quality, as measured by correlation between imputed allele counts and true counts, varied widely from variant to variant, with common variants having higher correlations than rare variants. While our study systematically investigated associations of a subset of 1.2 million well-imputed common and low-frequency 1000GP variants with PCa risk, we found no novel associations. 1000GP imputation provided dense coverage of previously-identified PCa susceptibility regions, highlighting its potential as an inexpensive first-pass approach to fine-mapping. For example, we found a subset of highly probable potential causal variants for further analysis at 5p15 and were able to successfully replicate known multi-locus associations at 8q24 found using denser genotyping approaches. At a whole genome sequencing cost of $2,500, imputation from SNP arrays is a more powerful strategy for SNPs with minor allele frequencies (MAF) above 1%. As sequencing costs fall, whole genome sequencing will become more cost effective at higher MAF's, however; SNP arrays are likely to remain the most cost-effective strategy for common variants. Our study shows 1000GP imputation can accurately identify low frequency and rare variants in large cohorts and stresses the importance of large sample size when attempting to use imputation for association studies with low frequency and rare variants.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1391W

SHAVE - Shrinkage Estimator Measured for Multiple Visits Increases Power in GWAS of Quantitative Traits. *O. Meirelles[1], J. Ding[1], T. Tanaka[1], S. Sanna[2], H. Yang[1], D.B. Dudekula[1], F. Cucca[2], L. Ferrucci[1], G. Abecasis[3], D. Schlessinger[1].* 1) Laboratory of Genetics, National Institute on Aging, Baltimore, MD; 2) Istituto di Ricerca Genetica e Biomedica, Consiglio Nazionale delle Ricerche, Monserrato, Cagliari, Italy; 3) Center for Statistical Genetics, Department of Biostatistics, University of Michigan, Ann Arbor, Michigan.

Measurement error and biological variability generate distortions in quantitative phenotypic data. In longitudinal studies with repeated measurements, the multiple measurements provide a route to reduce noise and correspondingly increase the strength of signals in genome-wide association studies (GWAS). In order to optimize noise correction, we have developed Shrunken Average (SHAVE), an approach using a Bayesian Shrinkage estimator. This estimator uses regression towards the mean for every individual as a function of 1) their average across visits; 2) their number of visits; and 3) the correlation between visits. Computer simulations support an increase in power, with results very similar to those expected by the assumptions of the model. The method was applied to a real data set for 14 anthropomorphic traits in ~6,000 individuals enrolled in the SardiNIA project, with up to 3 visits (measurements) for each participant. Results show that additional measurements have a large impact on the strength of GWAS signals, especially when participants have different number of visits, with SHAVE showing a clear increase in power relative to single visits. In addition, we have derived a relation to assess the improvement in power as a function of number of visits and correlation between visits. It can also be applied in the optimization of experimental designs or usage of measuring devices. SHAVE is fast and easy to run, written in R and freely available online.

1392T

To adjust or not to adjust, how and when to incorporate covariates into GWA studies. *G.P. Page[1], N. Garge[2], E.O. Johnson[3].* 1) Statistics & Epidemiology, RTI International, Atlanta, GA; 2) Genomics, Statistical Genetics, and Environmental Research Program, Research Triangle Institute International, Research Triangle Park, NC; 3) Behavioral Health Epidemiology Program, Research Triangle Institute International, Research Triangle Park, NC.

Diseases and traits that are studied using genome wide association studies often have environmental risk factor in addition to the genetic ones. Motivated by a GWA study of HIV among IV drug users whose other significant risk factors include age, gender, number of sex partners, sex work, receptive anal sex, STIs, and needle sharing, and recruitment pre- or post- availability of antiretroviral therapy, we sought to understand how to incorporate covariates into our analysis with maximal power and minimal error. We studied the impact on type I error and power of various strategies for the incorporation of covariates into GWA studies, including: matching, reducing covariates using latent class and principle component analysis, fully adjusting for covariates in linear and logistics models, and ignoring covariate effects. In linear and logistic models, we explored the trade-off between using a single set of covariates for all SNPs in the GWA verses maximizing the model for each SNP individually. We found essentially the same performance whether we used AIC, BIC, MSE or p-values for selecting covariates. The incorporation of a false covariate only modestly decreased power to detect genetic effects. When genetic variance is 5% of total variance and environmental variance 0% of total variance, power without incorporating false covariate is 0.616 (versus 0.612 when false covariates were considered). However modeling a real covariate markedly increases power as the effect of real covariate increases. At 5% genetic variance, the power produced in the presence of 10% environment variance was 0.659 in comparison to 0.836 at 40% environment variance. The joint consideration of real and false covariates only modestly reduced power (power = 0.832). The addition of neither real (0% genetic, 40% real environmental effects, type 1 error = 0.050) nor false environmental covariates (0% genetic, 40% real environmental effects, plus false environmental effect, type 1 error = 0.050) had an effect on the type 1 error for the genetic studies. Our studies suggest that covariates can and should be incorporated into GWA studies even if there is worry that they will be false for the increase in power due to real covariates far outweighs the reduction in power due to false covariates, and that the incorporation of covariates in to GWA does not lead to type 1 biases. In addition, our studies indicate that models can be built per SNP rather than using a single model for all SNPs.

1393F

Refining association mapping in a heterogeneous population: An example in a Schizophrenia case-control study. *A. Ray[1], LC. Lazzeroni[1, 2].* 1) Psychiatry and Behavioral Sciences Department, Stanford University, Stanford, CA; 2) Pediatrics Department, Stanford University, Stanford, CA.

We propose a refined association mapping method to address the possibility of variable linkage disequilibrium (LD) structure within a heterogeneous population. Here we will consider the situation in which markers associated with a trait fall into different LD blocks although the trait frequency is the same across ancestral populations. In this setup, individuals are allowed to have mixed ancestry originating in more than one subpopulation and subpopulations are allowed to have different LD patterns. Our two-stage association statistic combines aspects of principal component adjusted analyses and stratified analyses in order to address two issues simultaneously. The ancestry proportions will be determined locally in each LD block and stratified regression will be performed with each subpopulation as strata. We will apply our method to a genome-wide study of schizophrenia (GAIN) containing phenotype and genotype data for 1 million SNPs in 2.8K European and 2.2K African American participants. We will use schizophrenia diagnosis and the total DSM score to analyze the two heterogeneous populations separately focusing on candidate genes previously characterized by the Consortium on the Genetics of Schizophrenia (COGS) for association with schizophrenia and related endophenotypes. We will compare the association signals from our proposed two-stage association method in the heterogeneous GAIN populations to the relatively homogeneous Caucasian population in COGS and to one-stage association methods that allow less complex patterns of admixture and LD.

1394W

Improved heritability estimation from genome-wide SNP data with application to epilepsy. *D. Speed[1], G. Hemani[2], M. Johnson[3], D. Balding[1].* 1) UCL Genetics Institute, University College London, London, WC1E 6BT, United Kingdom; 2) Diamantina Institute, University of Queensland, Brisbane, Queensland 4102, Australia; 3) Division of Brain Sciences, Imperial College London, London W6 8RF, United Kingdom.

There has recently been much interest in using genome-wide SNP data to estimate the heritability of complex traits (Yang et. al. Common SNPs explain a large proportion of the heritability for human height, Nat. Gen. 2010). The method has been used to investigate numerous disease and other traits, and yet many of the assumptions underpinning the mixed model on which it is based have not previously been rigorously estimated. Through extensive simulation studies, we found the method to be surprisingly robust to violations of underlying assumptions of three types: the effect size distribution, the relationship between effect size and minor allele fraction, and the number of causal variants. However, we found that the heritability estimates can be highly biased by uneven patterns of LD across the genome. In particular, the contribution of causal variants in areas of high LD (for example in the MHC region) are over-estimated, while the contribution of causal variants in areas of low LD are under-estimated. To address this problem, we propose the use of a novel kinship matrix, in which the contribution of each SNP is weighted according to local LD levels. This LD-adjusted kinship matrix almost entirely removes the biases observed in our simulation study using the standard allelic-correlation kinship matrix. Applying our method to WTCCC data, we find for certain traits the estimates of heritability can be altered by as much as 25% compared with the standard method. SNP-based heritability analysis has been used to investigate the genomic architecture of traits, by comparing the variance explained by particular subsets of SNPs, for example, genic vs inter-genic SNPs or low MAF vs high MAF SNPs. By improving its accuracy, the adjusted analysis gives a better insight into the nature of causal variants, which we illustrate with reference to epilepsy GWAS data. LD-adjusted kinship is likely to be beneficial in other application areas of kinship coefficients, such as mixed model association analysis and genomic selection.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1395T

GLOMS: A mixed model-score test based system for association studies of binary traits with risk covariates in populations of related individuals. *S. Stanhope, M. Abney.* Department of Human Genetics, The University of Chicago, Chicago, IL.

Although a variety of effective and tractable methods have been developed for performing association studies of quantitative traits with risk covariates in populations of related individuals, there are substantially fewer options for studies of binary traits. This is likely due to the comparative technical difficulty of developing methods in the mixed model regression frameworks that are typically used for such analyses, and is compounded with the fact that there is currently an increasing focus on performing not only marker-wise studies of association, but also analogous gene enrichment, gene x environment, or marker x marker studies. In this talk, we describe GLOMS (Genome-wide LOgistic mixed model / Multivariate Score test), a parallelized, computationally efficient and powerful system for performing a wide variety of types of association studies of binary traits with risk covariates in populations of related individuals. We demonstrate the application of GLOMS to several kinds of association studies of hypertension in the Hutterites, a population related though a dense 13-generation pedigree.

1396F

Tiled Regression Improves the False Discovery Rate in Genome-Wide Association Studies. *B. Suktitipat[1,2], Y. Kim[1], A.J.M. Sorant[1], H. Sung[1], A.F. Wilson[1].* 1) Genometrics Section, NHGRI/NIH, Baltimore, MD; 2) Department of Biochemistry, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand.

Most genome-wide association studies (GWAS) aim to identify causal genes based on single variant effects ignoring the effects of all other variants in the genome. Ideally, multiple regression of all the variants in the genome would identify the set of independent causal variants; however, the requirement of complete data makes this approach problematic. Tiled Regression combines simple linear or logistic regression models, stepwise selection of variables and a staged approach. User-defined regions ("tiles") of potentially correlated SNPs (e.g., based on hotspots or genes) are first considered separately and discarded if they show no evidence of association with a trait. Stepwise regression is then used to select independent significant SNPs from the remaining tiles, with higher order stepwise regression used at the chromosome and genome levels. We evaluated the performance of this approach with the Tiled Regression Analysis Program (TRAP v1.0) using simulated data. We simulated 4000 individuals with 306,097 markers spanning 53,060 tiles across the genome, based on phased data from the HapMap CEU population. To evaluate the false positive rate, we simulated 100 replicates of normally distributed random phenotypes. To evaluate the power, we simulated 100 replicates of quantitative traits from seven independent loci with the same effect size but with various locus specific heritability from 0.0005 to 0.01 assuming additive genetic model. A critical value of 0.0001 was used as the selection criteria for the stepwise regression for the tile, chromosome, and genome levels. We compared our results to conventional single marker analysis based on simple linear regression model testing one marker at a time. On average, the type I error for TRAP was 5E-05, about one-half of the nominal type I error rate. TRAP had a false discovery rate (FDR), calculated as a proportion of false positive signals divided by all positive signals, of 0.76, which was better than the FDR of 0.86 from single marker analysis using the same statistical significant level. Power estimated with TRAP was comparable to the power estimated from single marker analysis. In summary, in a GWAS setting, we found that TRAP has comparable power but with a lower FDR when compared to single marker analysis. Thus, the use of TRAP is beneficial in selecting independent markers contributing to quantitative trait and in reducing the false discovery rate in GWAS.

1397W

Gene-based association tests using dimension reduction methods and marker correlation structure. *D.M. Swanson, C. Lange.* Biostatistics, Harvard University, Boston, MA.

The advent of genome-wide association studies has led to many novel disease-SNP associations, opening the door to focused study on their biological underpinnings. Because of the importance and relative simplicity of analyzing these associations, many statistical methods have been devoted to them. However, fewer methods have attempted to associate entire genes or genomic regions with outcomes, which is potentially more useful knowledge from a biological perspective. We develop two methods for quantifying the degree of association between a genomic region and outcome, both of which adequately control for LD structure in the region. One method uses dimension reduction methods when significant LD exists in the region to "filter" redundant information and increase power, while the other controls for LD by scaling marker Z-statistics using knowledge of the correlation matrix of markers. We apply both methods to sequence data of oral cleft and compare our results to previously proposed methods, in particular permutation-based ones. We find a significant association in the sequence data between the 8q24 region and oral cleft using our dimension reduction approach and a borderline significant association using the Z-statistic-based approach.

1398T

Impact of hidden sample structure on meta- and joint-analysis in genome-wide association studies. *P.J. Walter, H.M. Kang.* Biostatistics, University of Michigan, Ann Arbor, MI.

Genome wide association studies (GWAS) have shown the ability to detect novel loci associated with many human diseases but issues of hidden relatedness and population structure continue to engender spurious associations and high Type I error. While joint-analysis across multiple studies may provide high power, meta-analysis is often used in practice for convenient and secure analysis with robust control against heterogeneity among studies. Under confounding effect from population structure, correcting for hidden sample structure via principal component analysis or linear mixed models is typically recommended for joint-analysis. However, it is not clearly understood how these strategies behave in the context of meta-analysis with unknown pairwise genetic relatedness between different study samples. Through simulated phenotypes under a variety of polygenic models across nearly ten thousand individuals, we evaluated the statistical power and Type I error across different strategies of joint- and meta-analyses. We find that meta-analysis provides comparable power to joint-analysis for detecting common causal variants. We also find that, while joint-analysis with mixed models is slightly better at prioritizing causal SNPs than principal component correction, they perform similarly in meta-analysis. Compared to uncorrected analysis, we find that correcting for hidden sample structure within studies is better at prioritizing causal SNPs, and accounting for sample structure across studies, if possible, provides even slightly better results. We also evaluate the impact of different strategies of applying genomic control in meta-analysis. Finally, we evaluate different meta- and joint- analysis strategies focusing on low frequency variants. We expect that our results will serve as a useful guidance in designing genetic analysis of complex traits across multiple studies.

1399F

Natural and Orthogonal association framework to detect parent-of-origin effects. *F. Xiao, J. Ma, C. Amos.* UT M.D Anderson Cancer Center, Houston, TX. 1515 Holcombe Blvd. Houston, TX 77030 1-877-MDA-6789.

Many human diseases, such as cancer, diabetes and obesity may be related to imprinted genes. Traditional association studies always assume equal contribution of the paternal and maternal alleles to the trait. Recently many studies have revealed that the power of GWAS could be improved when the parent-of-origin effect (POE) is considered in the analysis model, while searching for complex disorders related genes. In the natural and orthogonal interaction (NOIA) framework which was developed for the quantitative trait analysis by Alvarez-Castro and Carlborg [Genetics, 2007], the estimates of genetic effects are unrelated, also called orthogonal. Therefore I proposed that the power of NOIA could be improved when POE is considered in the analysis model while keeping its orthogonality. Thus we may detect some novel significant genes that associate with complex disease, which were masked before in usual models. We conducted simulations to evaluate the performance of the statistical models in comparison with the usual functional models. We also compared the power of NOIA models with and without POE considered. The results showed that the power for testing associations while allowing for POE considered using the statistical model is much higher than using usual functional models. And as expected, even with the POE considered, the statistical model is still orthogonal for detecting both overall genetic effect and POE effect. We also applied our approach to family data about oral cleft disorder.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1400W

Efficient calculation for multi-SNP genetic risk scores. *T. Johnson.* GlaxoSmithKline, Stevenage, United Kingdom.

Genetic risk scores based on genotypes at multiple single nucleotide polymorphisms (SNPs) have several applications in association studies for complex human phenotypes. However, for many human diseases and traits of clinical importance, identifying genetic associations has required large sample sizes, so single SNP associations typically use meta-analysis summary results obtained from multiple genotype-phenotype datasets held at different institutions. In such settings, testing association between a multi-SNP genetic risk score and a phenotype is prone to organisational difficulties and potential for analytic error. I have developed a convenient "summary statistic" method, which allows most powerful testing of association between a multi-SNP genetic risk score and a phenotype, when only single SNP summary association statistics are available. The method is thus both statistically and computationally efficient. The method has been applied to genetic risk scores for blood pressure, adiponectin, and lipid levels, in several recent publications (Ehret et al. Nature 478:103–109, Dastani et al. PLoS Genetics 8:e1002607, Voight et al. Lancet online early), but the method itself has hitherto not been described in detail. I show that the "summary statistic" method gives comparable results to direct analyses of subject-specific phenotype and genotype data, shows increased power in genetic risk score analyses compared to alternative methods that combine information across SNPs, and demonstrate a novel diagnostic plot and goodness of fit test for the genetic risk score model. The latter may help to identify SNPs with pleiotropic effects, and may indicate situations where a Mendelian randomisation interpretation is not appropriate. I illustrate the method by application to previously published GWAS meta-analysis results for serum lipid traits and for risk of coronary artery disease (CAD). The results suggest that both serum high density lipoprotein (HDL) cholesterol, and serum triglycerides, are not themselves valid therapeutic targets for primary prevention of CAD. These results complement recent similar results for HDL and myocardial infarction (Voight et al. Lancet online early).

1401T

Gene-gene co-association and pathway-based co-association studies for next generation sequencing. *L. Luo[1], M. Xiong[2].* 1) Division of Epidemiology and Biostatistics, Department of Internal Medicine, University of New Mexico, Albuquerque, NM; 2) Division of Biostatistics, Human Genetics Center, The University of Texas School of Public Health, Houston, TX.

Systematically investigating the genetic architecture of common diseases requires complete knowledge of the human DNA sequence variations including SNPs, CNVs and other structural variations across the entire allele frequency spectrum. It has been increasingly realized that gene-gene interactions and the contribution of rare variants could be two potential factors in finding the unexplained disease heritability. However, to develop powerful statistics for testing gene-gene interaction based on data generated from next-generation sequencing remains great challenges for genetic epidemiologists and biostatisticians: (1)rare variants will be difficult to distinguish from sequencing errors in NGS studies; (2) existing statistical methods are underpowered to detect the interaction; (3) taking single SNPs as the basic units of interaction analysis will encounter severe multiple testing problems and heavy computations. To meet the challenges in testing the interaction between genetic variants, we propose to develop a novel concept of gene-gene co-association in which a gene or a genomic region is taken as a unit of association analysis and employing the stochastic processes to develop a unified framework for testing the co-association with disease of two genes, and multiple genes constituting a pathway, that contain both common and rare variants. We also developed a modified co-association analysis method which takes population stratification into consideration. We performed extensive simulations to validate the null distribution, and results showed that the newly developed statistic has type I error rates not appreciably different from the expected nominal level. Application to two independent GWAS datasets of psoriasis identified a network of gene-gene co-association that can be replicated in both studies. Results also showed that a pair of genes, each with no marginal association, may be jointly associated with disease.

1402F

May consanguinity between parents be a risk factor for insulin resistance in childhood obesity? Role of Mitochondrial Uncoupling protein 2 (UCP2) gene polymorphism. *S. Oguzkan Balci[1], N. Col-Araz[2], M. Nacak[3], M. Araz[4], S. Halime[1], A. Balat[5], S. Pehlivan[1].* 1) University of Gaziantep, Faculty of Medicine, Department of Medical Biology and Genetics, Gaziantep, Turkey; 2) University of Gaziantep, Faculty of Medicine, Department of Pediatrics, Gaziantep, Turkey; 3) University of Gaziantep, Faculty of Medicine, Department of Pharmacology, Gaziantep, Turkey; 4) University of Gaziantep, Faculty of Medicine, Department of Endocrinology and metabolism, Gaziantep, Turkey; 5) University of Gaziantep, Faculty of Medicine, Department of Pediatric Nephrology, Gaziantep, Turkey.

Objectives: Obesity is a global health problem, and frequently associated with metabolic complications such as type 2 diabetes mellitus. Childhood obesity will contribute to an increased prevalence of adult obesity. Uncoupling protein 2 (UCP2), a mitochondrial protein, plays an important role in thermogenesis and energy metabolism that might be related with the early onset of obesity or insulin resistance. The aim of this study was to investigate whether polymorphisms of-866 A/G UCP2 gene may be associated with childhood obesity and insulin resistance. **Methods:** A hundred obese children and adolescent and 100 age-and sex-matched healthy controls were tested for two polymorphisms which were -866 A/G promoter and exon 8 insertion/deletion (I/D) polymorphisms in UCP2 gene. Genotyping were performed by PCR and/or RFLP. Insulin resistance was determined as homeostasis model assessment >3.16 (HOMA). **Results:** Exon 8 I/D polymorphism in UCP2 gene showed an association with obesity. The distribution of DD, ID and II genotypes for the gene was 46.0 %,31.0 % and 23.0 % in obese subjects compared with 65.0 %, 30.0 % and 5.0 % in the controls. II genotype was higher in obese patients ($p=0.0001$), while DD genotype was higher in controls ($p=0.0034$). There were no statistically significant differences in the UCP2 gene -866 A/G promoter polymorphism between groups ($p>0.05$). There was also no statistically significant difference between -866 A/G promoter and exon 8 I/D polymorphisms in UCP2 gene and insulin resistance in childhood obesity ($p>0.05$). But if there was an consanguinity between parents GA genotype and A allele were associated with insulin resistance in chilhood obesity ($p<0.05$). Conclusion: Our results suggest that II genotype, of exon 8 of UCP2 gene was associated with childhood obesity, while -866 G/A poymorphism was not. However, GA genotype and A allele were associated with susceptibility to insulin resistance in case of an consanguineous marriage. However, further studies with larger sample size are needed in order to have a better understanding on the effect of this gene in childhood obesity and insulin resistance.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1403W

A Gene-based Association Statistic for Meta-Analysis. *P. Chanda[1,5], H. Huang[2,3], A. Alonso[4], J.S. Bader[1,5], D.E. Arking[6].* 1) Department of Biomedical Engineering, Johns Hopkins University, Baltimore, Maryland, USA; 2) Analytic and Translational Genetics Unit, Department of Medicine, Massachusetts General Hospital, 185 Cambridge St., Boston, Massachusetts, USA; 3) Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, 7 Cambridge Center, Cambridge, Massachusetts, USA; 4) Division of Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis, Minnesota, USA; 5) High Throughput Biology Center, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA; 6) McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA.

Gene-based association analysis has emerged as a powerful complement to identifying disease associated genes using traditional genome-wide single SNP association tests. In a gene-based analysis, one jointly analyzes all independent variants within a putative gene to obtain a single p-value representing the significance of association of the entire gene. When individual level genotype and phenotype data are not available, often due to restrictions on sharing data across cohorts, only the genotype-phenotype summary statistics are available for further analysis. We present a Bayesian method of greedy model selection combining multiple independent effects within a gene to generate a stronger association signal using summary data from a meta-analysis. Our method uses summary data (regression coefficients and standard errors) to perform efficient model selection (selecting a minimal subset of associated SNPs within a gene), parameter estimation and statistical testing (calculating gene-based p-values and genome-wide significance thresholds). The model selection procedure includes a model complexity penalty based on the number of effective SNPs within a gene and a Bayesian Information Criterion penalty that permits a fast estimate of the model likelihood through the sufficient statistics of the maximum likelihood estimator. In the absence of individual level genotype and phenotype data, the empirical p-values are computed by simulating z-scores under the null using random variates sampled from a multivariate normal distribution with covariance matrix computed from the linkage disequilibrium between the SNPs in an appropriate reference population (e.g. HapMap or 1000 Genomes Project). When applied to meta-analysis from genome-wide scans of Atherosclerosis Risk in Communities (ARIC) electrocardiographic traits, we recovered known associations (gold standard loci) with multi-SNP models with power similar to individual level genotype data. Our preliminary results indicate that the proposed method has the potential to identify additional novel disease-susceptibility genes for complex diseases using summary data from large GWAS datasets.

1404T

Joint testing of rare and common variants in admixed populations. *H. Qin, J. Li, H.-W. Deng.* Center for Bioinformatics and Genomics, Department of Biostatistics and Bioinformatics, Tulane University School of Public Health and Tropical Medicine, New Orleans, LA.

An admixed population and its ancestral populations bear different burdens of a complex disease, partially due to different frequencies of deleterious alleles. Deleterious alleles and allele ancestries are dependently distributed among admixed subjects. Within a susceptible gene, the affected and unaffected groups in an admixed population may differ in the marginal distributions of mutation accounts and allele ancestries as well as in the correlations between genotypes and ancestries. The next-generation sequencing technologies are enabling us to discover nearly all rare and common variants. As evidenced by ample empirical findings, rare and common genetic variants as well as environmental factors may work together to influence complex diseases. We propose a novel gene-based framework to integrate local ancestries with rare and common genetic variants in next-generation sequence data of admixed subjects. In this framework, we combine ancestries, genotypes, ancestry-gene interactions across variant sites in a gene by informative weights (e.g., signal-to-noise ratios) to enrich genetic information. We link the enriched genetic information to disease status through a regression model to adjust for covariates (e.g., population structure and environmental factors) and construct a suitable test statistic. Under the null of no association, the test statistic asymptotically follows the chi-square distribution with one-degree-of-freedom (1-df). Our framework allows for genetic effects of diverse sizes and directions and corrects for arbitrary covariates. It properly controlled type I error rates under extensive high-replicate simulations. Under various modes of disease genetics, our framework proved uniformly more powerful than several popular sequence association tests, e.g., the CAST, the SDWSS and the EREC. We demonstrated the prevailing utility to integrate local ancestries, rare variants, common variants and ancestry-gene interactions. Our novel framework can be potentially applied in next-generation sequence association studies of admixed populations.

1405F

The role of KIR genes polymorphisms in Brazilian patients with rheumatoid arthritis and systemic lupus erythematosus. *A.R. Marrero[1], T.D.J. Farias[1], C.C. Coêlho[1], L.D. Hausmann[1], D.G. Augusto[2], M.L. Petzl-Erler[2], I.A. Pereira[3], A.F. Zimmermann[3], S.C.M.S. Fialho[3], Y.C.N Muniz[1], I.R. Souza[1].* 1) LAPOGE, Federal University of Santa Catarina, Florianopolis, Santa Catarina, Brazil; 2) Federal University of Paraná, Curitiba, Paraná, Brazil; 3) HU, Federal University of Santa Catarina, Florianopolis, Santa Catarina, Brazil.

Rheumatoid arthritis (RA) and systemic lupus erythematosus (SLE) are a common autoimmune diseases that affects worldwide population, with female prevalence and characterized by a complex disease and unknown etiology. RA disease affects mainly the joints but can also compromise heart, lung, and eyes functions, while SLE disease affects kidney, lungs and heart functions and development arthritis and rash. The role of immune system in autoimmune pathogenesis is unclear, but the Natural Killer (NK) cells and its cytokines production behave differently in RA and SLE patients. The Killer-cell Immunoglobulin like Receptors (KIRs) are NK cell surface proteins that play an important role modulating the cytotoxicity on innate and adaptive immune system cells. In this study the genotyping was carried out by PCR-SSP from peripheral blood leucocytes of 145 RA, 166 SLE and 270 controls unrelated donors from Universitary Hospital of Federal University of Santa Catarina (Brazil). To eight KIR's genes, the difference in the presence frequency of individual KIR genes in cases and control was tested for significance using Chi square test and Odd Ratios was calculated, and $p<0.05$ was considered significant. Gene frequencies found in the present study were: KIR3DL3 (1.0), KIR2DL3 (0.942), KIR2DS3 (0.373), KIR2DS5 (0.430), KIR2DL4 (0.957), KIR3DS1 (0.436), KIR2DS4 (0.704) and KIR3DL2 (0.996). These frequencies are very similar to frequencies found in another Brazilian population from a different region. The activating KIR genes KIR2DS5 (OR=4.75, $p<0.0001$), KIR2DS3 (OR=2.117, $p=0.003$), and KIR3DS1 (OR=2.242, $p<0.0001$) were associated only with SLE, and the inhibitory genes KIR2DL3, KIR3DL3, KIR2DL4, KIR3DL2 gene was not associated with RA and SLE ($p>0.05$). These results may be indicating a potential role for NK cells in SLE pathogenesis since KIRs interact through its alleles to enhance the inflammatory response (Financial support: PNPD CAPES, FAPESC, CNPq and UFSC).

1406W

Genome-wide linkage, association and gene expression analysis of antibody levels against 13 common infections. *R. Rubicz[1], R. Yolken[2], E. Drigalenko[1], M.A. Carless[1], T.D. Dyer[1], P.E. Melton[1,3], J.W. Kent, Jr.[1], R. Duggirala[1], J.E. Curran[1], M.P. Johnson[1], S.A. Cole[1], L. Almasy[1], E.K. Moses[1,3], N.V. Dhurandhar[4], E. Kraig[5], J. Blangero[1], C.T. Leach[6], H.H.H. Goring[1].* 1) Dept of Genetics, Texas Biomedical Research Institute, San Antonio, TX; 2) Stanley Division of Developmental Neurovirology, Johns Hopkins University School of Medicine, Baltimore, MD; 3) Centre for Genetic Epidemiology and Biostatistics, The University of Western Australia, Perth, Australia; 4) Infections and Obesity Laboratory, Pennington Biomedical Research Center, Louisiana State University System, Baton Rouge, LA; 5) Department of Cellular and Structural Biology, University of Texas Health Science Center at San Antonio, San Antonio, TX; 6) Department of Pediatrics, University of Texas Health Science Center at San Antonio, San Antonio, TX.

There are individual differences in susceptibility to infectious diseases, which are attributable to a number of potential factors including host genetic variation. We previously demonstrated that differences in antibody titer, which reflect history of infection with a particular pathogen, are significantly heritable. Here we conduct genome-wide analyses using 3 genome-wide gene mapping technologies in order to identify genetic factors influencing variation in these serological phenotypes. Blood samples from >1300 Mexican Americans were quantified for IgG antibody titer to 13 common bacterial and viral infections: *Chlamydia pneumoniae*; *Helicobacter pylori*; *Toxoplasma gondii*; cytomegalovirus; Epstein-Barr virus (EBV); herpes simplex I virus; herpes simplex II virus; human herpesvirus 6; varicella zoster virus; adenovirus 36 (Ad-36); hepatitis A; influenza A; and influenza B. The quantitative antibody titer phenotypes were analyzed for each pathogen and for several measures of seroreactivity to multiple pathogens (i.e., all herpesviruses, all viruses, and all pathogens). Genome-wide variance components linkage yielded significant LOD scores for EBV (on chromosome 6), influenza A (on chromosome 6), and Ad-36 (on chromosome 18). Joint linkage and association analyses (using an additive measured genotype model and ~1 million SNPs) produced genome-wide significant results for EBV on chromosome 6 ($p = 1.1 \times 10^{-12}$). Results for measures of seroreactivity to multiple pathogens, however, were not significant. Microarray-based gene expression data (for >16,000 transcripts, generated from peripheral blood lymphocytes collected at the same time as the blood samples used for serological phenotyping) led to identification of significantly correlated transcripts for approximately half the pathogens. Further investigation of anti-EBV antibody variation indicates this trait is influenced by multiple genetic factors within the human leukocyte antigen (HLA) complex, which are specific to this pathogen and not infection in general. Our results demonstrate that individual loci regulating serological measures of these common infections do exist and can be localized by statistical gene mapping approaches.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1407T

Combining Affected Families, Independent Cases, and Controls to Obtain a Single, More Powerful Test of Association. *W. Stewart[1], J. Cerise[2].* 1) Battelle Center for Mathematical Medicine, The Research Institute, Nationwide Children's Hospital, Columbus, OH; 2) Dept Biostatistics, Columbia University, New York, NY.

Over the last few decades, investigators have begun to accumulate complicated data sets that contain affected families, independent cases, and controls. However, the statistical complexity of these data sets makes them difficult to analyze efficiently, and in effect, this has encouraged investigators to perform suboptimal analyses instead. Therefore, we propose a unified test that is the weighted average of a population-based (e.g. case-control) and a family-based (e.g. TDT) test of association. Since both tests are influenced by a common set of genotypes (namely, the genotypes of cases within affected families), the two tests are correlated. Our unified test accounts for this correlation; and for large samples, its distribution is approximately normal (with mean zero and variance one) under the null hypothesis of no association. To assess the power and the type I error of our unified test, which we have implemented in the program POPFAM, we simulated genotypic and phenotypic data across: different genetic models, minor allele frequencies, and levels of correlation between the test SNP and the trait locus. We compared POPFAM to existing methods, and we applied it to the analysis of a real hypertension study as well. Our simulations show that POPFAM is more powerful than individual case-control or family-based tests of association, and that, for some scenarios POPFAM is more powerful than other methods that combine population-based and family-based evidence of association. Finally, in our re-analysis of a familial arterial hypertension study, we integrated information from publicly available HapMap controls to reduce the p-value at a candidate SNP. Overall, in regions containing prior evidence for linkage, POPFAM provides increased power to detect an association. In addition, the software is user-friendly, well documented, and freely available from the web.

1408F

Mixture modeling of rare variant association within exome sequencing data. *B.A. Logsdon[1], J.Y. Dai[1,2], P.L. Auer[1], S.K. Ganesh[3], N.L. Smith[4], J.G. Wilson[5], T.A. Graubert[6], R.P. Tracy[7], L.A. Lange[8], H. Tang[9], S. Rich[10], G. Lettre[11], C.S. Carlson[1], R. Jackson[12], C. O'Donnell[13], M.M. Wurfel[14], D.A. Nickerson[15], C. Kooperberg[1], A.P. Reiner[1,4], NHLBI Exome Sequencing Project, Exome Sequencing Project Blood Counts Project Team.* 1) Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA; 2) Vaccine and Infectious Diseases Division, Fred Hutchinson Cancer Research Center, Seattle, WA; 3) Division of Cardiology, University of Michigan Health System, Ann Arbor, MI; 4) Department of Epidemiology, University of Washington, Seattle, WA; 5) Department of Physiology and Biophysics, The University of Mississippi Medical Center, Jackson, MS; 6) Department of Internal Medicine, Division of Oncology, Washington University, St Louis, MO; 7) Departments of Pathology and Biochemistry, University of Vermont, Burlington, VT; 8) Department of Genetics, University of North Carolina, Chapel Hill, NC; 9) Department of Genetics, Stanford University, Stanford, CA; 10) Center for Public Health Genomics, University of Virginia, Charlottesville, VA; 11) Montreal Heart Institute, Université de Montréal, Montreal, Canada; 12) Department of Internal Medicine, The Ohio State University, Columbus, OH; 13) Framingham Heart Study of the National, Heart, Lung, and Blood Institute, Framingham, MA; 14) Division of Pulmonary and Critical Care Medicine, Department of Medicine, Harborview Medical Center, University of Washington, Seattle, WA; 15) Department of Genome Sciences, University of Washington, Seattle, WA.

We propose a novel approximate inference methodology to identify the subset of rare variants within a gene that drives association with phenotype. While many burden tests have been proposed to identify genes with an aggregation of rare variants associated with phenotype, they generally lack the capability to identify which rare variants are responsible for the association with phenotype. This is both biologically and statistically relevant for mapping rare variation because previous work indicates that even among non-synonymous mutations, likely only 15–20% are functional. Our approach specifically models the presence of both functional and neutral variants with a discrete mixture distribution. We show through simulations that our approach has greater or comparable power to other popular burden tests such as the Combined Multivariate and Collapsing (CMC) method, the Sequence Kernel Association Test (SKAT), the Weighted Sum Statistic (WSS), the Variable Threshold (VT) method, and the Gene or Region-based Analysis of Variants of Intermediate and Low frequency (GRANVIL) method, while also identifying the subset of rare variants associated with phenotype. Our algorithm leverages a fast variational Bayes approximate inference methodology to scale to exome-wide analyses. To demonstrate the efficacy of our approach we analyze platelet count within the National Heart, Lung, and Blood Institute's Exome Sequencing Project. Platelet count is a potentially important intermediate phenotype for blood and thrombotic diseases and has a reported heritability in the range of 50–70%. Through a recent genome-wide association study with imputed exome data in African Americans, several common and lower-frequency variants associated with platelet count have been identified. Using our novel mixture modeling approach, as well as other popular burden methods including the SKAT, GRANVIL, VT, and WSS methods, we identify novel genes with nominally significant exome-wide associations. For these genes we demonstrate the advantage of our mixture method over competing methods by discriminating between putative functional and neutral variants based on the estimated posterior probability of association for each variant from the mixture model.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1409W

Detecting Association of Rare Variants by Testing an Optimally Weighted Combination of Variants in Family-based Designs. *Q. Sha, S. Fang, S. Zhang.* Mathematical Sciences, Michigan Technological University, Houghton, MI.

Although next-generation sequencing technology will soon allow sequencing the whole genome of large groups of individuals, the development of powerful statistical methods for rare variant association studies is still underway. Even though many statistical methods have been developed for mapping rare variants, most of these methods are for unrelated individuals only, whereas family data have been shown to improve power to detect rare variants. Most of the existing methods for unrelated individuals are essentially testing the effect of a weighted combination of variants with different weighting schemes. The performance of these methods depends on the weights being used. Recently, Sha et al. (Genetic Epi, 2012) proposed a Test to test the effect of an Optimally Weighted combination of variants (TOW) for unrelated individuals. It has been shown that TOW performs better than others, especially when both risk and protective variants are presence or there are a large number of neutral variants. In this article, we extend our previously developed TOW for unrelated individuals to family-based data and propose a novel Test to test the effect of an Optimally Weighted combination of variants for Family-based design (TOW-F). The optimal weights are analytically derived and can be calculated from children's genotypes and phenotypes. TOW-F is based on within-family information and thus is robust to population stratification. Using extensive simulation studies, we evaluate and compare our proposed method with the following three methods (1) Weighted Sum Family-Based Association Test (WS-FBAT) that is based on the weight suggested by Madsen and Browning (PLoS Genet 5:e1000384, 2009); (2) Tmax = max {Ti}, where Ti is the FBAT at the ith variant; (3) Adaptive Weighting FBAT (AW-FBAT) in which, we calculate adaptive weights using between-family information, then use FBAT to test for regional association based on the adaptive weighted combination of variants (Fang et al., Genetic Epi, DOI:10.1002/gepi.21646, 2012). Our results show that the proposed method is robust to population stratification, robust to the direction and magnitude of the effects of causal variants, and is more powerful than the other three methods, especially when both risk and protective variants are presence.

1410T

Haplotype-based methods for detecting uncommon causal variants with common SNPs. *N. Liu[1], W. Lin[2], N. Yi[1], D. Zhi[1], K. Zhang[1], G. Gao[3], H. Tiwari[1].* 1) Dept Biostatistics, Univ Alabama at Birmingham, Birmingham, AL; 2) Institute of Epidemiology and Preventive Medicine, College of Public Health, National Taiwan University, Taipei, Taiwan; 3) Department of Biostatistics, Virginia Commonwealth University, Richmond, Virginia.

Detecting uncommon causal variants (minor allele frequency (MAF) < 5%) is difficult with commercial single-nucleotide polymorphism (SNP) arrays that are designed to capture common variants (MAF > 5%). Haplotypes can provide insights into underlying linkage disequilibrium (LD) structure and can tag uncommon variants that are not well tagged by common variants. In this work, we propose a wei-SIMc-matching test that inversely weights haplotype similarities with the estimated standard deviation of haplotype counts, to boost the power of similarity-based approaches for detecting uncommon causal variants. We then compare the power of the wei-SIMc-matching test with that of several popular haplotype-based tests, including four other similarity-based tests, a global score test for haplotypes (global), a test based on the maximum score statistic over all haplotypes (max), and two newly proposed haplotype-based tests for rare variant detection. With systematic simulations under a wide range of LD patterns, the results show that wei-SIMc-matching and global are the two most powerful tests. Among these two tests, wei-SIMc-matching has reliable asymptotic P values, whereas global needs permutations to obtain reliable P values when the frequencies of some haplotype categories are low or when the trait is skewed. Therefore, we recommend wei-SIMc-matching for detecting uncommon causal variants with surrounding common SNPs, in light of its power and computational feasibility.

1411F

Multivariate Adjusted Sequence Kernel Association Test for Rare Variants Controlling for Cryptic and Family Relatedness. *K. Oualkacha[1,2], R. Li[2], B. Richards[2,3,4,6], A. Ciampi[4], C. Greenwood[2,4,5], UK10K cohorts group.* 1) Mathematics, Université de Québec à Montréal (UQÀM), Montreal, Québec, Canada; 2) Lady Davis Institut, Jewish General Hospital, Montreal, Québec, Canada; 3) Medicine, McGill University; 4) Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Québec, Canada; 5) Oncology, McGill University, Montreal, Québec, Canada; 6) Human Genetics, McGill University, Montreal, Québec, Canada.

We have recently developed a framework for association testing of rare variants in family-based designs. This framework is an adaptation of the sequence kernel association test (SKAT) which allows us to control for family structure. Our adjusted SKAT (ASKAT) combines the SKAT approach and the factored spectrally transformed linear mixed models (FaST-LMM) algorithm to capture family effects based on a linear mixed model incorporating the realized proportion of the genome that is identical by descent between pairs of individuals. ASKAT is fast and can analyze hundreds of thousands of markers. Moreover, regardless the level of the trait heritability and the sample structure, the ASKAT framework has good control of type I and provides good power. Despite the performance of this method, power for rare-variant association tests is always a concern. Most currently used methods analyze a single phenotype at a time. However, often multiple correlated phenotypes are collected and the single-phenotype analyses do not exploit this correlation. Multivariate-phenotype analysis has been shown to provide greater statistical power than single trait analysis in linkage studies and in genome-wide association studies (GWAS) of complex diseases/traits, but has not received attention in rare variant association tests. We have developed a multivariate analytical method for rare variant association study in family-based or mixed designs, which is an extension of the ASKAT framework to high dimensions. Our approach, (MASKAT), controls the sample structure, the polygenic pleiotropy and the environmental effects based on a multivariate unbalanced mixed model. Performance is validated through extensive simulation studies. Under different simulation scenarios, the results show that our MASKAT approach outperforms rare-variant single-phenotype approaches in terms of power. In particular, MASKAT has substantial advantages when a positive phenotypic correlation between traits exists. Overall and learnt from this approach, to gain power when multiple correlated phenotypes are available, we would recommend in practice the use of MASKAT approach. We plan to apply our methodology to the bone density data from the UK10K consortium.

1412W

Robust and powerful tests for rare variants using Fisher's method to combine evidence of association from two or more complementary tests. *A. Derkach[1], J.F. Lawless[2,3], L. Sun[1,2].* 1) Statistics, University of Toronto, Toronto, Ontario, Canada; 2) Division of Biostatistics, Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada; 3) Statistics and Actuarial Science, University of Waterloo, Waterloo, Ontario, Canada.

Many association tests have been proposed for rare variants, but there is uncertainty about the practical choice of a good test due to lack of insight concerning the underlying genetic models. Current methods use either linear statistics, which are powerful when almost all variants are causal and have the same direction of effect, or quadratic statistics which are more powerful in other scenarios. The differences in power, unfortunately, can be significant in many settings, e.g. 20% vs. 60% or vice versa To achieve robustness, we propose using Fisher's method to combine evidence from the complementary linear and quadratic statistics. We can also consider minimum-p as another robust statistic. Analytical results and extensive simulation studies of over 10,000 different models for studies of both binary and quantitative traits show that both Fisher's and minimum-p methods are robust across genetic models with varying proportions of causal, deleterious and protective rare variants, variant frequencies, effect sizes and the relationships between variant frequencies and effect sizes. Further, Fisher's method consistently outperforms the minimal-p approach and the individual linear and quadratic tests in several plausible scenarios, for example, when at least 75% of the causal effects are in the same direction (either deleterious or protective). When a linear or quadratic test has moderate power of 20% or more, Fisher's test has considerable better power than each test for 90% of the models considered, while the maximum absolute power loss is 8% for the remaining 10% of the models. An application to the Genetic Analysis Workshop (GAW) 17 quantitative trait Q2 data based on sequence data from the 1000 Genomes Project shows that, compared with the individual linear and quadratic tests, Fisher's test not only has comparable power for all 13 functional genes, but it also provides the best power for more than half of them. The general concept of using Fisher's method to borrow information from two or more existing but complementary methods applied to the same data, beyond the traditional setting of meta-analysis of multiple data resources, to provide robust and (in some scenarios) more powerful tests can be readily extended and useful to many other genetic and genomic studies.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1413T

Leveraging family history in genome-wide association studies. *A. Ghosh, S. Wacholder, P. Hartge, M. Purdue, S. Chanock, N. Chatterjee.* Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, MD.

Participants in many genome-wide association (GWA) studies provide information on history of cancers in family members, particularly first-degree relatives. Mendelian law implies that genotypes of susceptibility markers will be associated with participants' family history; nonetheless, information on association between genetic markers and family history is ignored in standard analyses aimed at gene discovery.

We propose a novel meta-analysis approach that combines association signals from both the primary case-control outcome and the family-history information. Weights used in standard meta-analysis typically accounts for precision of association parameter estimates. Our modification to standard meta-analysis weights also incorporates expected attenuation of strength of association between genotype and family history due to inference of distribution of relatives' genotypes from that of participants. To evaluate the method empirically we used genotype and family history data from 6,411 men and women from the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial (PLCO) cohort included in previously conducted GWA studies. Specifically, we evaluate the performance of the proposed approach in studying associations with 30 single-nucleotide polymorphisms (SNPs) previously established as prostate cancer susceptibility markers. We report the median and the mean of the ratio of the chi-square statistics from case-control and meta-analysis over 30 SNPs to evaluate relative efficiency.

We show that meta-analysis further strengthens evidence for genetic association, at times by a few orders of magnitude. For rs4242382, the SNP most strongly associated with prostate cancer, the p-values are 1.1×10^{-7} and 4.8×10^{-10} from standard case-control and meta-analysis respectively. The mean and median of the ratio of chi-squares from meta-analysis and case-control analysis are 1.62 and 1.25 respectively, which matches with our expectation from theoretical calculation of asymptotic relative efficiency.

Routinely collected information about family history in studies used in GWA can increase the power of detecting association between a genetic variant and disease at no extra cost.

1414F

Identifying environmental exposures for gene-environment investigations using LASSO penalized regression methods. *E. Mowry[1], X. Shao[2], F.B.S. Briggs[2], B. Acuna[3], L. Shen[3], A. Bernstein[3], C. Schaefer[3], L.F. Barcellos[2].* 1) Department of Neurology, JOhn Hopkins Hospital, Baltimore, MD; 2) Genetic Epidemiology and Genomics Laboratory, Division of Epidemiology, School of Public Health, University of California, Berkeley, Berkeley, CA; 3) Division of Research, Kaiser Permanente, Oakland, CA.

In epidemiological studies with detailed genetic and environmental exposure data, determining key predictors for gene-environment (GxE) analyses is challenging, as the ordinary least squares model may become oversaturated. In addition to multiple testing concerns, the underlying correlation structure between predictors may introduce residual confounding. Thus, we are motivated to implement variable selection to identify trait-associated predictors for investigation. LASSO (least absolute shrinkage and selection operator regression) is a popular penalized statistical method that can address these challenges, and can be summarized as calculating the regression coefficients by minimizing the penalized sum of squares. Therefore, LASSO may be considered as a variable selection method for the OLS model, and there is high probability that the predictors selected contain the true set of effective variables. We applied LASSO to identify a subset of environmental exposures for GxE analyses in a large population-based multiple sclerosis (MS) case-control study. MS is a complex autoimmune disease of the central nervous system, with genetic and environmental components. More than 50 risk loci have been identified, however only a few environmental exposures have been confirmed, including tobacco smoke exposure, history of infectious mononucleosis, and vitamin D intake. LASSO was applied to a large data set of 1503 white female (914 MS cases and 589 controls) members of the Kaiser Permanente Northern California membership. Thirteen environmental exposures were studied, including several cosmetic, occupation, chemical and animal exposures. These predictors, known MS environmental risk factors, socioeconomic status (SES) and age were evaluated by LASSO. Ten-fold cross-validation determined the shrinkage parameter that minimized the variance. Six exposures had non-zero coefficients in LASSO regression. Logistic models, adjusted for known environmental risk factors, age, SES, *HLA-DRB1*15* and the weighted genetic risk score assessed the relationship between the selected predictors and MS risk. Our results did not show a strong effect on MS susceptibility for any of the environmental exposures investigated. Here, we demonstrate the application of LASSO to identify a subset of environmental exposures from a large dataset with reduced penalty for multiple statistical test.

1415W

Application of targeted maximum likelihood estimation using genome-wide association data reveals new genetic predictors of disease severity in multiple sclerosis. *X. Shao[1], M. van der Laan[1], F. Briggs[2], P. De Jager[3,4], L. Barcellos[2], The International Multiple Sclerosis Genetics Consortium.* 1) Division of Biostatistics, School of Public Health, University of California, Berkeley, CA 94720-7356, USA; 2) Division of Epidemiology, School of Public Health, University of California, Berkeley, CA 94720-7356, USA; 3) Program in NeuroPsychiatric Genomics, Center for Neurologic Diseases, Department of Neurology, Brigham & Women's Hospital and Harvard Medical School, Boston, MA 02115, USA; 4) Program in Medical & Population Genetics, Broad Institute of Harvard University and Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

Genome-wide association studies (GWAS) have been widely used for identifying disease associated loci. With high-throughput genotyping technology, GWAS test millions of variants across the human genome. Multiple sclerosis (MS) is a chronic inflammatory disorder of the central nervous system with a strong genetic component which influences both disease onset and progression. Studies have shown that the genetic architecture of MS progression is likely polygenic and comprised of moderate effects. Due to insufficient power and lack of efficient methods to correct for multiple testing, GWAS have not been fruitful in identifying risk loci for disease severity. Therefore, rather than assessing the causal effect of SNPs based on GWAS, it is more reasonable to consider GWAS as a screening tool for disease gene discovery. This study aims to identify candidate genes that may influence disease severity, measured by MS Severity Score (MSSS). A two-stage analytical method was applied to a genome-wide SNP data which consisted of two million SNPs for 1,470 MS cases. The phenotype was less severe disease (MSSS≤5) vs. more severe disease (MSSS>5). In the first stage, univariate logistic regression was applied to the two millions SNPs. Lasso was applied to SNPs with marginal unadjusted p-value<0.05 from univariate tests and was reduced to 1,011 independent SNPs. In the second stage, targeted variable importance measures (tVIM) based on targeted maximum likelihood estimation (TMLE) were obtained for each SNP. tVIM is an updated TMLE estimate based on the initial estimator and the confounding mechanism between the targeted SNP and a set of related SNPs. P-values were computed for the marginal tVIM, and modified FDR control for multi-stage analysis was applied. The candidate loci were selected based on adjusted p-values (P<0.05). The most significant finding was a non-genic SNP: rs11765846 on 7q21.2 (P=1×10-17). The top genic results were *C1orf49* on 1q25.2 (rs12754916: P=2×10-11); *ABCA9* on 17q24.2 (rs17684521, a missense SNP: P=2×10-09); *AMPH* on 7p14-p13 (rs17171399: P=5×10-7); *PTPRD* on 9p23-p24.3 (rs10977017: P=2×10-5; rs1322133: P=0.04); and *PPARGC1A* on 4p15.1 (rs10938963: P=3×10-5). In this GWAS, TMLE based tVIM approach identified 36 variants within 35 genes that associate with MSSS; replication analyses are currently underway. Compared to other machine learning algorithms, this approach provides interpretable statistical inference with valid p-values (FDR adjusted).

1416T

Case-sibling studies that acknowledge unstudied parents and enroll unmatched individuals. *M. Shi, D.M. Umbach, C.R. Weinberg.* Biostatistics Br, NIEHS, Res Triangle Park, NC.

Family-based designs enable assessment of genetic associations without bias from population stratification. However, parents are not always available — especially for diseases with onset later in life and the case-sibling design, where each case is matched with one or more unaffected siblings, can be useful. Analysis typically accounts for within-family dependencies by using conditional logistic regression (CLR). We consider an alternative approach that treats each case-sibling set as a nuclear family with both parents missing by design. Under weak assumptions, one can carry out maximum likelihood analysis by using the Expectation-Maximization (EM) algorithm to account for missing parental genotypes. We show that this approach enables the investigator to incorporate supplemental cases who do not have a sibling available and supplemental controls whose case sibling is not available (e.g. due to death). We compare conditional logistic regression and the proposed missing parents approach under several risk scenarios. Our proposed method offers both improved statistical efficiency and asymptotically unbiased estimation for genotype relative risks and genotype-by-exposure interaction parameters.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1417F

A Variable-Selection-Based Novel Statistical Approach to Identify Susceptible Rare Variants Associated with Complex Diseases with Deep Sequencing Data. *H. Sun, S. Wang.* biostatistics, Columbia University, New York, NY., Select a Country.

Existing association methods on sequencing data have been focused on aggregating variants across a gene or a genetic region in the past two years due to the fact that analyzing individual rare variants is underpowered. However, to identify which rare variants in a gene or a genetic region out of all variants are associated with the outcomes (either quantitative or qualitative) is a natural next step. Here we proposed a variable-selection-based novel approach that is able to identify the locations of the susceptible rare variants that are associated with the outcomes with sequencing data. More specifically, with all *P* rare variants in a gene or a genetic region, we generated the power set of the *P* rare variants except the empty set, that is, subsets $S = 2^P - 1$ of the *P* rare variants. We then treated the *S* subsets of the *P* rare variants as the *S* "new variables" and applied the penalized likelihood estimation using L_1-norm regularization in a regression framework where outcomes were regressed on the *S* subsets. The proposed method only requires the regularization procedure to select the highest impact subset out of the *S* subsets on the outcome without worrying the selection of an optimal regularization parameter. After a subset of *P* rare variants is selected as the most associated subset with the outcome, we applied a permutation procedure specifically designed to assess the statistical significance of the selected subset. The selection performance and power of the proposed method were evaluated through intensive simulation studies where different effect sizes, sample sizes and directions of the effects of the individual rare variants were considered. The results demonstrated that the proposed method is able to select subsets with most of the outcome related rare variants in all simulation scenarios considered. The type I error and power of the subsequent permutation procedure demonstrated the validity and advantage of our selection method. The proposed method was also applied to sequence data on the ANGPTL family of genes from the Dallas Heart Study (DHS). Our proposed method was implemented in an **R** package *RVsel* which will be freely downloaded at *http://www.columbia.edu/~sw2206*.

1418W

Causal-Seek: A strategy for finding causal variants via trans-ethnic fine-mapping. *X. WANG[1], Y. Teo[1,2].* 1) School of Epidemiology and Public Health, National University of Singapore, SINGAPORE; 2) Department of Statistics and Applied Probability, National University of Singapore, SINGAPORE.

Genome-wide association studies (GWAS) have achieved remarkable success in identifying genomic variants that are associated with human traits and complex diseases. However, these variants are unlikely to be the real causal SNPs, but are markers that are in linkage disequilibrium (LD) with the functional variants. SNPs having high LD with the causal SNP often emerge from the analysis with strong association signals with the phenotype of interest. To identify the actual causal SNP, fine-mapping with denser genotyping or sequencing information is performed in the candidate regions identified by GWAS. This procedure is hampered by the extensive LD patterns around the causal SNP, which can yield numerous perfect surrogates that are indistinguishable from the real causal SNP. Trans-ethnic fine-mapping that leverages the diverse LD patterns across multiple ethnic populations can be used to address this difficulty. Here we introduce a novel statistical method to perform trans-ethnic fine-mapping, by calculating the probability that each SNP is the causal variant given the observed GWAS association results as well as the genetic variability in each population. Across multiple populations, Maximum likelihood method is adopted to assess the probability at each SNP position. After leveraging the diverse LD patterns across multiple ethnic populations, the SNP that is consistently in high correlation with the surrogates (SNPs emerged from GWAS with strong association signals) in separate GWAS is most likely to be the real causal SNP. We first implemented a series of simulations under different scenarios to assess the power and false positive rates. Subsequently, we applied this method to genome-wide association data of type 2 diabetes in Chinese, Malay and Asian Indian populations from Singapore as well as the diverse populations from European and Japan, to investigate for the real causal variant for the phenotype of interest.

1419T

Novel insights into the genetics of Parkinson's disease on chromosome 17. *W.WS. Lau.* RESEARCH DEPARTMENT O F GENETICS, EVOLUTION AND EN, UNIVERSITY COLLEGE LONDON, LONDON, United Kingdom.

Parkinson's disease (PD) is the second most common neurodegenerative disease. Several G enome Wide Association (G WA) studies have been preformed to date focusing primarily on sporadic PD. These studies have reported a relatively small number of susceptibility loci even though recent studies have indicated a large genetic contribution. Here w e show significant progress by using high-resolution metric linkage disequilibrium maps to identify new genes on chromosome 17 that harbours the MART gene (strongest replicated G WA signal on 17q). Our analysis of the NIN DS G WA data (Illumina 550, 677 cases and 538 controls) has identified the MAPT gene-region but with an estimated location of the casual variant downstream of MAPT. This localization was replicated with precision using an independent G WA on sporadic PD (Illumina 1 M, 2,000 cases and 1,986 controls). Here w e report several additional novel genes, including ALOX15, RABEP1, PER1, CENTA2, TH E M100, AXIN 2, CD300LE, LG ALS3BP. All these signals have been replicated with great precision and most of our location-estimates point to causal variants within regulatory regions. The use of metric linkage disequilibrium maps can bring a much greater resolution and hence they can accelerate progress in identifying the genetic factors to PD and other complex disorders.

1420F

Joint statistical modeling of multiple phenotypes in samples with related individuals. *Z. Wang.* Division of Biostatistics, Yale University, New Haven, CT.

Genetic association studies have routinely been conducted to search for variants associated with diseases and quantitative phenotypes. Clinical and epidemiological studies typically collect data on a set of correlated phenotypes that may share common environmental and/or genetic factors. Such phenotypes contain more information than univariate phenotypes. Thus joint modeling of multiple phenotypes can potentially have increased power to detect association and increased precision of parameter estimation than univariate analysis. In this study, we develop novel statistical methods for multivariate association mapping in samples that contain arbitrarily related individuals. We address both common and rare variants association. The proposed methods are based on retrospective analysis that is less dependent on model assumptions on phenotypes, thus they are robust to trait model misspecification. The new methods can accommodate the external biological information by integrating graphical models and multivariate analysis, and are computationally affordable. We comprehensively evaluate the proposed methods using simulation studies and compare with existing methods. The proposed methods are applied to analysis of blood lipid levels in a whole genome sequence study on the Sardinian population.

1421W

Association of three polymorphisms of IL-18 genes (137G/C, 607C/A,133C/G) in patients with allergic rhinitis in the Iranian population. *S. Ramazi[1], M. Motovalibashi[1], H. Khazraei[2], M. Hashemzade chaleshtori[2].* 1) Cellular and Molecular Research Center, Shahrekord University of Medical Science, Shahrekord,Iran; 2) Division of Genetics, Department of Biology, Faculty of Science, University of Isfahan, Isfahan.

Background: Allergic rhinitis (AR) is an inflammatory disease of the nasal mucosa induced by an IgE-mediated reaction, following exposure to an allergen. IL-18 is a member of the IL-1 family.It was originally described as IFN-γ -inducing factor (IGIF),and is known to influence the balance of Th1/Th2 immune response. This study aimed to examine the association of three different (SNPs) located in IL-18 gene (-607 C/A, -137 G/C and-133 C/G) on chromosome 11q22 with allergic rhinitis. Methods :to determine the role of IL-18 gene in Allergic rhinitis susceptibility, we conducted a case-control study and genotyped three different (SNPs) in promoter region(-137 G/C (rs 187238), -607 C/A (rs 1946518), -133C/G (rs360721)) in 300 patients with allergic rhinitis and 200 healthy controls frame Iranian population. Polymerase chain reaction-restriction fragment length polymorphism were used for genotyping. Results: The frequency of the GC genotype of the IL-18/-133 gene polymorphism was greater in allergic rhinitis patients than in control. Conclusions: This study suggests that IL-18 gene variants may participate as a risk factor in the pathogenesis of AR or in intermediary phenotypes.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1422T

Accounting for Population Stratification in DNA Methylation Studies. *R.T. Barfield[1,2], L.M. Almli[3], V. Kilaru[3], A.K. Smith[3], K.B. Mercer[3], T. Klengel[4], D. Mehta[4], E.B. Binder[3,4], K.J. Ressler[3], K.N. Conneely[1,5].* 1) Department of Biostatistics and Bioinformatics, Emory University, Atlanta, GA; 2) Department of Biostatistics, Harvard University, Boston, MA; 3) Department of Psychiatry and Behavioral Science, Emory University, Atlanta, GA; 4) Max-Planck Institute of Psychiatry, Munich, Germany; 5) Department of Human Genetics, Emory University, Atlanta, GA.

DNA methylation is an important epigenetic mechanism that has been linked to complex disease, and is thus of great interest to researchers as a potential link between genome, environment, and disease. As the scale of DNA methylation association studies approaches that of genome-wide association studies (GWAS), issues such as population stratification will need to be addressed. Failure to adjust for population stratification in genetic association studies can lead to false positives, but population stratification generally has not been accounted for in DNA methylation studies. Here, we devise and compare several approaches to correct for population stratification in DNA methylation studies. We first analyzed DNA methylation for association with race in 388 individuals (365 African American and 23 Caucasian) from a larger study of post-traumatic stress disorder. Whole blood DNA methylation for these individuals was assessed using the Infinium HumanMethylation450 BeadChip Assay, and genotype data was collected using Illumina Omni-Quad 1M and Omni-Express arrays. We identified widespread associations between methylation and race (12,800 associated CpG sites; FDR<0.05). We then compared three different principal-components-based approaches on their ability to adjust for population stratification, measured as the percent reduction in the number of sites significantly associated with race. We examined principal components computed from 1) genome-wide methylation data, which we first pruned to obtain a set of roughly independent sites, 2) sets of CpG sites within 0 bp, 1 bp, 2bp, 5bp, 10bp, 50bp, and 100bp of a known SNP ("informed pruning"), and 3) conventional principal component analysis based on the genome-wide SNP data (after the usual LD-based pruning). We found that by including principal components based on either GWAS data or informed pruning of methylation data, we reduced the number of sites significantly associated with race by nearly 100%. The informed pruning approach was also less computationally intensive than the other two approaches, which both require correlation-based pruning. We have therefore developed an effective method to account for population stratification in DNA methylation studies that does not require the additional collection of genome-wide SNP data.

1423F

Genetic case-control matching strategies for rare-variants analysis in genome-wide association studies. *A. Lacour[1], T. Becker[1,2].* 1) German Center for Neurodegenerative Deseases (DZNE), Bonn, Germany, Bonn, Germany; 2) Institute for Medical Biometry, Informatics and Epidemiology, University of Bonn, Bonn, Germany.

Population stratification in samples of genome wide association studies give rise to large obliterations in the results of statistical tests. In order to correct for stratification effects we have developed and implemented three structuring strategies for association testing. Taking the identity-by-state matrix as a measure for genetic kinship, we obtain pairwise case-control matchings, groupings with respectively at least one control and one case, and clusters. All strategies have in common that they are based on a `maximum weighted bipartite matching' by making use of the hungarian algorithm which solves the assignment problem in polynomial time. We also present a set of quality control strategies that make sure, that the structures found by the algorithm are not between distinct strata. The matching strategies can not only be performed genome-wide but also window-wise as a `localized matching'. Local window sizes of a few thousand SNPs are enough to guarantee identification of strata. This is of particular interest for the application to the analysis of rare variants since the amount of stratification is expected to vary according to the genomic location. Association P-values are obtained by within-structure case-control permutations. As it turns out from simulation studies, the empirical niveau becomes close to the confidence niveau for within-structure permutation becomes close to the nominal level. Thus, our method leads to an increase of power and simultaneously to a reduced false-positive rate in simulations compared to unstratified analyses. As a byproduct, our implementation strongly outperforms common covariate approaches based on principal component analysis or multidimensional scaling in runtime, and makes genome-wide application possible. Our method for stratified analyses is implemented in the genome-wide interaction analysis software INTERSNP.

1424W

Ancestry Informative Principal Components Analysis in Structured Samples with Known or Cryptic Relatedness. *T. Thornton, M. Conomos.* Biostatistics, University of Washington, Seattle, WA.

A problem that is often encountered in association studies of complex traits is that of identifying and adjusting for population stratification, as it is well known that failure to appropriately account for population structure can lead to spurious association. Principal components analysis (PCA) is a commonly used tool for the detection and correction of confounding due to population structure in association studies. For genetic studies that include related individuals, however, PCA can lead to principal components that do not adequately capture population structure due to the complex correlations among relatives. We consider the problem of PCA in related samples from structured populations. We propose a method, PC-AIR, for obtaining ancestry informative principal components in samples with related individuals. PC-AIR can be used across a variety of study designs, ranging from studies that have a combination of unrelated individuals and pedigrees with known genealogy to studies with partially known or completely unknown pedigrees. In simulations studies with related individuals and population structure, including admixture, we demonstrate that PC-AIR provides a substantial improvement over traditional PCA in a broad range of settings. We apply PC-AIR to the Mexican Americans in Los Angeles, California (MXL) population sample of release 3 of phase III of the International Haplotype Map Project (HapMap) where we demonstrate that the first principal component with PC-AIR captures population structure better than the top 30 principal components with traditional PCA.

1425T

Admixture and Association Mapping Identifies Marker in FAM19A2 Associated with FEV_1/FVC in African-Americans. *M.M. Parker[1], M.G Foreman[2], R.A. Mathias[3], T.H. Beaty[1], C.R. Gignoux[4], E.G. Burchard[4], J.B. Hetmanski[1], E.K. Silverman[5], J.D. Crapo[6], COPDGene Investigators.* 1) Epidemiology, Johns Hopkins University, Baltimore, MD; 2) Morehouse School of Medicine, Atlanta,GA; 3) Allergy and Clinical Immunology, Johns Hopkins Asthma and Allergy Center, Baltimore, MD; 4) Department of Medicine, University of San Francisco, San Francisco, CA; 5) Channing Laboratory, Brigham and Women's Hospital, Harvard Medical School, Boston, MA; 6) National Jewish Medical and Research Center, Denver, CO.

Background: The ratio of forced expiratory volume in 1 second (FEV_1) to forced vital capacity (FVC) is used to diagnose chronic obstructive pulmonary disease (COPD) severity. Racial differences in FEV_1/FVC ratio exist between non-Hispanic Whites and African-Americans and have not been fully explained. Methods: We performed admixture mapping to identify genetic variants associated with FEV_1/FVC ratio in 3260 unrelated African-American smokers in the COPDGene Study. FEV_1/FVC ratio was measured using an EasyOne Spirometer and participants were genotyped on the Illumina OMNI Express platform (728,648 SNPs). Local ancestry was estimated for each individual at each marker using LAMP, which infers local ancestry using a clustering algorithm on sliding windows of contiguous SNPs. Overlapping windows are then combined via majority vote to estimate the most likely ancestral population for each SNP. These estimators of local ancestry were included in a linear regression to identify genomic regions where local ancestry was a significant determinant of FEV_1/FVC ratio. Results: Global percent African ancestry among African-American COPDGene participants ranged from 29.5%–99.9% with a mean of 81.7%. This global measure of African ancestry was significantly associated with FEV_1/FVC ratio (β=-0.3412, p=0.033). The strongest association signal from the regression of FEV_1/FVC ratio on local ancestry at each genetic marker (controlling for age, gender, BMI, current smoking status, pack years smoked, and global ancestry) was on chromosome 12 ($p=6.77\times10^{-6}$). Single-SNP genotype association analysis in this region identified an associated SNP within the intronic region of FAM19A2 (rs11174267, $p=6.10\times10^{-9}$) that remained significant even after adjustment for local ancestry ($p=1.71\times10^{-6}$). FAM19A2 is a member of the TAFA gene family and is involved in immune and nerve cell regulation, but has not been reported to be associated with pulmonary function. Conclusion: Through a combination of admixture mapping and association analysis, we identified a marker (rs11174267) significantly associated with FEV_1/FVC ratio among African-American participants in the COPDGene study. This may partly explain of racial differences in spirometry, although the role of this marker in lung function remains unclear. Funding: U01HL089856 (EKS), U01HL089897 (JDC), HL092601 (MGF).

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1426F

The Gamma Method for gene set analysis of RNA-Seq data: Simulation results and application to a smallpox vaccine study. *B.L. Fridley[1], G.D. Jenkins[1], D.E. Grill[1], J.M. Biernacka[1], R.B. Kennedy[2], G.A. Poland[2], A.L. Oberg[1].* 1) Health Sciences Research, Mayo Clinic, Rochester, MN; 2) General Internal Medicine, Mayo Clinic, Rochester, MN.

Many complex disorders may be controlled by the interplay of multiple genes within the same molecular pathway or gene set (GS). Thus, gene set analysis (GSA) has been widely used for mRNA data from microarrays. Aggregating the association signals for a set of genes within a GS incorporates biological knowledge, reduces the multiple-testing burden, and may increase the association signal, thus increasing the power to detect meaningful associations. With the advent of next-generation sequencing technologies, the use of GSA is even more relevant due to limited statistical power, as RNA-Seq studies are often conducted on a small number of subjects. However, RNA-Seq data analysis comes with some different challenges, as RNA-Seq experiments measure the gene expression levels from the total number of reads detected in a gene. Therefore, the analysis methods for RNA-Seq data are quite different than those for microarray-based data. In our GSA, we use generalized linear models to assess per-gene significance assuming the negative binomial distribution. Using these gene-level results, we propose the use of the Gamma Method (GM) for GSA. The GM is based on summing p-values transformed using an inverse Gamma(ω, 1) transformation (Zaykin et al., 2007). Application of different p-value transformations is achieved by changing the ω parameter which varies the emphasis given to particular p-values, and has been referred to as the soft truncation threshold (STT). When ω is 1, GM is equivalent to Fisher's Method. Due to dependency between genes in a GS, empirical GS p-values are determined using permutation testing. In a simulation study with more than 1,400 scenarios, we found the GM with STT < 0.20 outperforms other commonly used self-contained GSA methods such as Fisher's Method and the global model of Goeman et al. (2004), while maintaining control over the type I error rate. We have applied the GM for GSA to a RNA-seq smallpox study to identify GSs with differences in mRNA expression between high and low responders to smallpox vaccine. The top GSs (with good coverage) included: Biotin metabolism (p=0.0005); non-homologous end-joining (p=0.02); focal adhesion (p=0.04); D-Glutamine and D-glutamate metabolism (p=0.04); and ECM-receptor interaction (p=0.05). The results of this research will lead to a powerful method for GSA of RNA-Seq data and new insights into the genomic basis of individual variations in immune response to the smallpox vaccine.

1427W

Evaluation of SNP and indel imputation quality using reference haplotypes from the 1000 Genomes project. *Q. Duan[1], E.Y. Liu[1,3], K.L. Mohlke[1], D.C. Croteau-Chonka[1], Y. Li[1,2,3].* 1) Department of Genetics, University of North Carolina, Chapel Hill, NC 27599; 2) Department of Biostatistics, University of North Carolina, Chapel Hill, NC 27599; 3) Department of Computer Science, University of North Carolina, Chapel Hill, NC 27599.

Genotype imputation has become a routine in genome-wide association (GWA) studies. As imputation accuracy directly affects downstream analysis, there is an increasing need to better understand the quality of imputation, especially for rare genetic variants. Currently, one can only infer SNP and indel imputability information by conducting imputation, which can be computationally costly for a large number of markers of interest. To facilitate this inquiry, we sought to perform a detailed imputation quality assessment of SNPs and indels discovered in the 1000 Genomes project and to provide the imputability information through a publicly available database. Imputation was performed using MaCH-Admix with 2184 haplotypes from the 1000 Genomes project (version 3 March 2012 release). We mimicked imputation practice in typical GWA studies by treating markers absent from selected commercial genotyping platforms as untyped markers. We imputed the genotypes at these untyped markers one sample at a time using the rest of the samples as reference. Imputation accuracy of each marker was measured by the squared Pearson correlation coefficient between the imputed dosages and the actual (masked) genotypes. The accuracy was evaluated separately for each of four continental groups. Genotyping platforms currently available are Affymetrix 500K, Affymetrix 6.0, Affymetrix Axiom, Illumina Human1M, Illumina Omni5 and Illumina Omni ZhongHua. Our results show that imputation accuracy differs by the choice of genotyping platform. Within each genotyping platform, imputation quality increases with the minor allele frequency (MAF) in terms of number of SNPs passing quality control (QC) and the average of the squared Pearson correlation. Finally, we show that the imputation accuracy at each QC level is affected by both population group and MAF, indicating that to achieve similar imputation performance, different QC thresholds need to be applied by taking ancestry and MAF into consideration. The imputability of SNPs and indels is publicly available through our website (http://www.unc.edu/~yunmli/imputability.html).

1428T

An efficient design to detect transplant donor and recipient genetic interactions. *W. Guan[1], A.K. Israni[2].* 1) University of Minnesota, Minneapolis, MN; 2) Hennepin County Medical Center, Minneapolis, MN.

Long-term renal and liver allograft failure remains a major problem in transplantation. Besides recipient characteristics, donor-related factors also play important roles to predict allograft failure, accounting for up to 45% of the observed patient-to-patient variation in allograft function. In order to determine the role of donor and recipient genetic factors and potential interactions, we will design a genome-wide association study (GWAS) using the recipient-donor pairs. A deceased donor usually donates two kidneys and/or a liver, typically each to a different recipient. This unique design allows us to distinguish the relative contribution of the donor genetic factors from the recipient and environmental related factors. In the proposed analysis, we will conduct the association tests in two stages. We first identify recipient variants associated with allograft failure using recipient pairs by conditional logistic regression; and then test for association between donor variants and allograft failure, with the donor-recipient genetic interaction being modeled with each of the identified recipient variant. Through computer simulations, we will compare the statistical power using this recipient-donor pair design over traditional case-control design which exhaustively searches for interactions between every donor and recipient variants. We hope that our approach will help to identify the donor and recipient genetic factors for all solid organ transplants and lead to future development of novel immunosuppression therapies/treatments.

1429F

Dissecting features of causal variants. *S. Sengupta, X. Wen, G. Abecasis.* Biostatistics Dept, University of Michigan, Ann Arbor, MI.

Genome-wide association studies, which examine millions of genetic variants in thousands of individuals, have identified many complex trait associated loci. Most of these loci include many strongly associated variants and, often, variants which are not tested for association but are known to be in strong linkage disequilibrium with the variants exhibiting strongest evidence of association. The large number of variants actually or potentially showing evidence for association in each locus makes it challenging to prioritize likely functional variants at each locus.

We reasoned that causal variants for each trait might share certain genomic features. For example, causal variants for lipid traits might preferentially overlap transcription factor binding sites active in liver, where important steps in lipid metabolism take place. More generally, causal variants for many traits might be non-synonymous variants that alter protein coding sequences. We develop a hierarchical model that identifies genomic features enriched among associated many associated loci and uses this information to prioritize likely functional variants in each locus.

Our model can be fitted to summary statistics from individual studies, making it convenient to incorporate in ongoing genome-wide association study meta-analysis that can include 100,000s of individuals distributed across dozens of studies. We evaluate our method using simulations and application to genome-wide association study data for a variety of metabolic traits.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1430W

Genetic Programming for Detecting Interactions of SNPs with Fuzzy Genotypes in Association Studies. T. *Stöcker[1,2], H. Schwender[1,2]*. 1) Faculty of Statistics, TU Dortmund University, Dortmund, Germany; 2) Faculty of Mathematics and Natural Science, Heinrich Heine University, Düsseldorf, Germany.

In genetic association studies, combinations of single nucleotide polymorphisms (SNPs) are assumed to have a higher impact on the risk of developing a disease than individual SNPs. Therefore, several methods for detecting SNP interactions associated with disease have been proposed, reaching from multiple testing procedures and exhaustive search approaches to procedures such as logic regression (Ruczinski et al., 2004, J. Comput. Graph. Statist. 12, pp. 475–511) and GPAS (Genetic Programming for Association Studies; Nunkesser et al., 2008, Bioinformatics 23, pp. 3280–3288) that search the space of all possible SNP combinations for the interactions influencing the disease risk most substantially by employing a stochastic search algorithm, and therefore, avoid the need to consider all possible combinations. In these methods, the genotypes of the SNPs are typically considered as if they were deterministic. This, however, ignores that in microarray and sequencing studies the genotypes are estimated by a calling algorithm, and hence, specified with uncertainty. It is well-known that taking this uncertainty into account by, e.g., using fuzzy genotypes, i.e. sums over the three possible genotypes weighted by their probability of being the true genotype, can lead to an improved statistical power in association studies. Considering fuzzy genotype becomes even more important (if not mandatory) when imputed SNPs are included in the analysis, as the uncertainty in these genotype calls is even higher. We have therefore extended GPAS to the analysis of fuzzy genotype calls. This procedure called Fuzzy GPAS thus allows not only accounting for the confidence in the genotype calls when searching for SNP combinations associated with disease, but also enables the consideration of both typed and imputed SNPs. In our presentation, we describe Fuzzy GPAS and show how it can be applied to typed and imputed SNPs. Moreover, we propose a measure quantifying the importance of the found SNP interactions, and hence, allowing the ranking of these interactions by their impact on the disease risk. As applications to simulated and real data show, this measure is able to correctly order the interactions by their influence on the disease risk.

1431T

Rare variant analysis for family-based design. *W. Yip[1], G. De[1], I. Ionita-Lazza[2], N. Laird[1]*. 1) Dept of Biostatistics, Harvard School of Public Health, Boston, MA; 2) Dept of Biostatistics, Columbia University, New York, NY.

Many common diseases are now thought to be caused by rare as well as common variants. Genome-wide association studies (GWAS) are designed only to identify the common variants. Recently, a number of methods have been proposed to detect rare variants in population based designs. Here, we proposed a novel method to test for association of rare variants in family-based design by collapsing the standard family-based association test (FBAT) statistic over a region of interest. A weighting scheme based on allele frequency of the SNPs is used to upweigh rare variants compare to common variants. Using simulations, we show that our method performs at par with the population-based methods under no population stratification. By definition, family-based designs are completely robust to population stratification and hence our proposed method is valid even when population stratification is present. The proposed method has been incorporated as an option (-v) in the popular FBAT software. Manuscript submitted to PLos ONE and is currently under review.

1432F

BMI trajectory from early childhood associates with leukocyte telomere length in Northern Finnish women. *S. Das[1], J.L Buxton[2], A. Rodriguez[1,7,8], A. Couto[1], M. Kaakinen[3,4], S. Sebert[1], P. O'Reilly[1], L. Coin[1], A.I.F. Blakemore[2], M-R. Jarvelin[1,3,4,5,6]*. 1) Epidemiology and Biostatistics, Imperial college London, London, UK; 2) Section of Investigative Medicine, Division of Diabetes, Endocrinology and Metabolism, Imperial College London, UK; 3) Institute of Health Sciences, University of Oulu, Finland; 4) Biocenter Oulu, University of Oulu, Finland; 5) National Institute for Health and Welfare, Helsinki, Finland; 6) MRC Health Protection Agency (HPA) Centre for Environment and Health, Imperial College London, UK; 7) Social Genetic Developmental Psychiatry Centre, King's College London, UK; 8) Mid Sweden University, Östersund, Sweden.

Leukocyte telomere length (LTL) is a biomarker for healthy aging, and short telomeres are associated with several age-related conditions, including obesity, type 2 diabetes, cardiovascular diseases and some cancers. Telomeres generally shorten with each cell division, a process exacerbated by oxidative stress. Inverse associations between telomere length in adults and increased body mass index (BMI), waist-to-hip (WHR) ratio and excess visceral fat have been previously reported. There is also some evidence for association between adverse events in childhood and shorter adult LTL. However, there are no studies on the relationship between BMI trajectories from childhood to adulthood and LTL in adulthood. We examined the longitudinal variation in BMI trajectories from childhood (at 5y) to adulthood (at 31y) to identify associations with adult LTL in the Northern Finland 1966 Birth Cohort. Relative mean LTL measurements were obtained using multiplex qPCR in DNA taken from N=5620 participants aged 31 years. We calculated the variation in BMI trajectories as the difference between standardised BMI measurements taken at ~5 y and at 31 y, and tested for association with LTL using regression models adjusting for qPCR plate, and obesity in childhood and in adulthood in consecutive models. Separate analyses were conducted in males and females.We found that changes in BMI across early childhood and adulthood are associated with shorter mean relative LTL in women, independent of obesity status in either early childhood or concurrent at age 31 (P=0.01). We did not find any evidence for an association between BMI trajectories and LTL in men. We also found that the mean relative LTL in men is significantly shorter than that in women (P=1 × 10-4), consistent with previous studies, and with the known antioxidant effects of oestrogen. The identification of novel factors affecting both telomere length and rate of telomere shortening throughout life is important for the elucidation of mechanisms underlying healthy aging. We show here that the degree to which BMI increases after childhood is associated with shorter telomeres at age 31, irrespective of whether individuals are obese at either timepoint. Further longitudinal studies are required to assess the long term effects of weight gain after childhood on LTL and cardiovascular disease risk, and the gender difference observed.

1433W

A model for combining data on de novo mutations with case-control data to identify risk genes. *X. He[1], S. Sanders[2], L. Liu[3], M. State[2], B. Devlin[4], K. Roeder[1,3]*. 1) Lane Center of Computational Biology, Carnegie Mellon University, Pittsburgh, PA; 2) Program on Neurogenetics, Child Study Center, Department of Psychiatry, Yale University School of Medicine, New Haven, CT; 3) Department of Statistics, Carnegie Mellon University, Pittsburgh, PA; 4) Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, PA.

Purpose: De novo mutations, which are not inherited from parents, affect risk for many diseases. An example is autism. Four recent whole-exome sequencing (WES) studies of 932 autism families (mother, father and affected offspring) identified five novel risk genes from a multiplicity of de novo loss-of-function (LOF) mutations and revealed that de novo LOF mutations occurred at a twofold higher rate than expected by chance. Not all WES studies, however, target families and in any case we would expect transmitted variation to also play a role in risk. Therefore an important research challenge is to develop statistical methods that increase the utility of de novo mutations by leveraging additional datasets of inherited or case-control variation. Methods: Our strategy is to model data on de novo mutations together with other forms of data. Case/control WES data, for example, are an unknown mixture of de novo and inherited variation. For simplicity, we focus on LOF mutations because they tend to have large effects on risk (e.g., autism). We develop a statistical model relating the number of de novo mutations to a set of genetic parameters such as mutation rate and relative risk; in addition, the statistical framework analyzes additional datasets, such as case-control data, in an integrated fashion. Inference is based on an empirical Bayes strategy that allows us to borrow information across all genes to infer parameters that would be difficult to estimate from individual genes. Results and Discussion: We validate our statistical strategy using simulations mimicking rare, large-effect mutations. We found that using either de novo mutations or case/control data alone is relatively inefficient. Moreover a naïve method of combining the two datasets using sum of log p-values (Fisher's method) has limited benefits. In contrast, our model substantially increases statistical power and dominates all other methods in almost all settings we examined. We will apply the model to WES data from autism to evaluate its performance on real data.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1434T

Variable Selection in High-dimensional Meta-Analysis for Genetic Data. *Q. He[1], H. Zhang[2], D.Y. Lin[1], C.L. Avery[3].* 1) Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel hill, NC; 2) Department of Statistics, North Carolina State University, NC; 3) Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel hill, NC.

Meta-analysis plays an important role in summarizing and synthesizing scientific evidence from multiple studies. When the dimensions of the data are high, such as in the genome-wide association studies or the sequencing studies, it is desirable to incorporate variable selection into meta-analysis to improve model interpretation and prediction. Existing variable selection methods require direct access to raw data, but in reality, it is often difficult to obtain the raw data from different resources. We propose a new approach, Sparse Meta-Analysis (SMA), which performs variable selection for meta-analysis based solely on summary statistics. This method is extremely useful when the information of many studies needs to be integrated in order to identify weak genetic variants. In addition, our method allows the effect sizes of each covariate to vary among studies, thus offering high flexibility for statistical modeling. Simulation studies and real data analysis demonstrate that the proposed methods perform well in both variable selection and prediction. Since summary statistics are far more accessible than raw data, our methods provide a powerful tool for selecting important variables from high-dimensional meta-analysis.

1435F

Integrative approaches for genetic association studies via Bayesian model uncertainty. *M.A. Quintana, D.V. Conti.* Dept of Preventive Medicine, Division of Biostatistics, University of Southern California, Los Angeles, CA.

Recent technological advances within the genetics community have lead to the collection of massive datasets involving numerous types of genomic data on a single set of individuals or across various levels of information. Traditional analyses concentrate on a single data type and relied heavily on an expert to build models for inference based on contextual knowledge. However, with the shear amount of factors available for evaluation it is often impracticable to build a model in this way. With this in mind, we are interested in developing integrative approaches for high-dimensional genetic association studies that incorporate multiple-levels of data, including functional information. In particular, we have developed iBMU, a Bayesian model uncertainty method that formally incorporates multiple sources of data via a multi-stage hierarchical probit model on the probability that each predictor is associated with the outcome of interest. Using simulations, we demonstrate that iBMU leads to a more efficient model search algorithm that yields an increase in power to detect true associations over more commonly used techniques. Finally, we demonstrate the power and flexibility of iBMU for several genome wide association studies involving: (1) the analysis of rare variants with incorporation of genetic annotation; (2) gene-set enrichment analysis for inference of gene- and pathway-level associations; (3) integrated genomic analysis incorporating functional information via biomarkers.

1436W

Meta-analysis of co-regulated subnetworks in transciptomics data: towards functional marker profiles of human ageing. *E.B. van den Akker[1,2], W.M. Passtoors[1], E.W. van Zwet[3], J.J. Goeman[3], M. Hulsman[2], V. Emilsson[4], M. Perola[5], B.T. Heijmans[1], A.B. Maier[6], J.N. Kok[1,7], P.E. Slagboom[1,8], M.J.T. Reinders[2], M. Beekman[1,8].* 1) Molecular Epidemiology, Leiden University Medical Center, Leiden, Zuid Holland, Netherlands; 2) The Delft Bioinformatics Lab, Delft University of Technology, Delft, Zuid Holland, Netherlands; 3) Medical Statistics, Leiden University Medical Center, Leiden, Zuid Holland, Netherlands; 4) Icelandic Heart Association, 201 Kopavogur, Iceland; 5) National Institute for Health and Welfare, Helsinki, Finland; 6) Gerontology and Geriatrics, Leiden University Medical Center, Leiden, Zuid Holland, Netherlands; 7) Algorithms, Leiden Institute of Advanced Computer Science, University of Leiden, Netherlands; 8) Netherlands Consortium for Healthy Ageing, Leiden, 2333 ZC, Netherlands.

Age associated modules within co-expression networks may provide a molecular basis for identification of mechanisms underlying age related diseases. However, age-changes in gene expression profiles typically have low signal-to-noise ratios that seriously hamper a solely data driven approach. We cope with this aspect in two ways. First, we employ a meta-analysis approach on multiple large-scale expression studies for constructing the co-expression network and detecting age associated modules therein. Second, we project the co-expression network onto the Protein-Protein Interaction (PPI) network spanned up by predicted or experimentally validated functional relations between genes to improve the confidence and interpretation of observed co-regulated subnetworks. For the analysis we created a gene expression compendium of 2,539 subjects in an age range of 15 to 95 years belonging to three published expression studies measured in peripheral blood or lymphocytes: SAFHS, IFB cohort and DILGOM. Single gene meta-analysis showed functional enrichment for N-linked glycosylation, besides a few other broadly defined terms. Our approach based on co-regulated subnetworks (CoRSuN), on the contrary, detected 70 age associated CoRSuNs, including a highly significant module containing T-cell differentiation markers (p = 1.3E-143). To investigate whether the detected CoRSuNs mark ageing up to the highest ages, we tested them for association with age in 50 middle-aged and 50 ninety-year-old participants from the Leiden Longevity Study. This confirmed age-changes in five CoRSunNs, including the T-cell CoRSuN, independent of lymphocyte and monocyte cell counts. Since the T-cell differentiation markers are established markers of T-cell lineages, their expression may point to proportions of T-cell subpopulations, rather than to a molecular mechanism. We therefore adjusted the age-associated CoRSuNs for the first principle component of the T-cell differentiation CoRSuN now revealing 15 out of the 70 subnetworks, including a ribosomal (p = 9.7E-4) and mitochondrial module (p = 4.1E-5). We conclude that age-changes in the blood transcriptome likely reflect changes in T-cell populations, potentially confounding the genomic analysis into ageing mechanisms. Furthermore, we demonstrate that de novo grouping of genes using co-expression networks in conjunction with a PPI network yields functional gene modules as relevant markers of human ageing up to very high ages.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1437T

Integrative Analysis of Sequencing and GWAS Data Improves Statistical Power in Detecting Rare Variants Associated With Complex Diseases. *Y.J. Hu[1], Y. Li[2,4], L.A. Lange[4], E.M. Lange[4], C. Bizon[6], P. Auer[5], G. Heiss[3], C. Kooperberg[5], N. Franceschini[3], U. Peters[5], A.P. Reiner[5], L. Hsu[5], S. Jiao[5], C.S. Carlson[5], K.E. North[3], D.Y. Lin[2], NHLBI GO Exome Sequencing Project.* 1) Department of Biostatistics and Bioinformatics, Emory University, Atlanta, GA; 2) Department of Biostatistics, University of North Carolina, Chapel Hill, NC; 3) Department of Epidemiology, University of North Carolina, Chapel Hill, NC; 4) Department of Genetics, University of North Carolina, Chapel Hill, NC; 5) Fred Huchinson Cancer Research Center, Seattle, WA; 6) Renaissance Computing Institute, University of North Carolina, Chapel, Hill, NC.

Biological evidence suggests that low-frequency and rare variants account for a large proportion of the genetic contributions to complex human diseases. Although next-generation sequencing has made it possible to conduct association studies on low-frequency and rare variants, it is still not economically feasible to sequence all subjects in a large cohort. A cost-effective sequencing strategy prioritizes subjects in the extremes of the quantitative trait distribution or those with diseases. In the NHLBI Exome Sequencing Project (ESP), subjects with the highest/lowest values of body mass index, low-density lipoprotein or blood pressures and those with early onset myocardial infarction or ischemic stroke were selected for whole-exome sequencing. The ESP samples were drawn from large well-established cohort studies which had collected GWAS data on most cohort members. By imputing the exome-sequencing data from the GWAS data for the cohort members who were not selected for sequencing, we can dramatically increase the number of subjects with information on low-frequency and rare variants. It is highly challenging to impute rare variants because only a very small proportion of subjects carry a rare mutation. In addition, treating the imputed genetic values as known quantities in downstream association analysis yields inflated type I error and biased estimates of genetic effects with under-estimated variances, especially when the sequenced subjects are not random samples from the underlying cohorts. We propose to impute burden scores (i.e., linear combinations of mutation counts over multiple variant sites within a gene) rather than individual variants and to integrate the imputation of burden scores and the inference of association into a single likelihood framework that properly reflects the actual sampling design. This approach produces the most powerful test among all valid tests and provides unbiased and efficient estimates of genetic effects. The advantages of the proposed methods are demonstrated through extensive simulation studies and applications to the NHLBI ESP.

1438F

Regionally-Smoothed Meta-analysis for GWAS studies. *F. Begum[1], E. Feingold[1,2].* 1) DEPARTMENT OF BIOSTATISTICS, UNIVERSITY OF PITTSBURGH, PITTSBURGH, PA; 2) DEPARTMENT OF HUMAN GENETICS, UNIVERSITY OF PITTSBURGH, PITTSBURGH, PA.

Over the last decade, genome-wide association studies (GWAS) have become a standard tool for gene discovery in human disease research. While debate continues about how to get the most out of these studies and on occasion about how much value these studies really provide, it is clear that many of the strongest results have come from large-scale mega-consortia and/or meta-analyses that combine data from up to dozens of studies and tens of thousands of subjects. While such analyses are becoming more and more common, statistical methods have lagged somewhat behind. There are good meta-analysis methods available, but even when they are carefully and optimally applied, there remain some unresolved statistical issues. Imputation availability and inaccuracy, heterogeneity in the study cohorts, and differences in data cleaning and statistical modeling among datasets are a few of the major unresolved issues. To address some of the current challenges, we propose a two-stage regionally smoothed meta-analysis (RSM) method for GWAS studies. In the first stage of RSM, we use a sliding window method to get a regional statistic. In the second stage, summarized window statistics for different studies are combined using an appropriate meta-analysis model and allowing for study heterogeneity. RSM is the first method that makes it possible to combine different studies with different sets of SNPs. We applied our method to combine several real GWAS data sets with known results. RSM was able to detect the known genes with increased power.

1439W

Phenotype Mapping using Information Explorer. *Y. Arens[1], J.L. Ambite[1], C.N. Hsu[1], L. Lange[2], S. Sharma[1], S. Voinea[1].* 1) Intelligent Systems Div, USC-ISI, Marina del Rey, CA; 2) Department of Genetics, University of North Carolina, NC.

Databases such as dbGaP represent extremely valuable resources of data that have been assembled across multiple cohorts. The increasing development of cost-effective high-throughput genotyping and sequencing technologies are resulting in vast amounts of genetic data. While such databases were formed in order to archive and distribute the results of previously performed genetic association analyses, an increasing number of studies have provided de-identified individual-level genotypic and phenotypic data that are made available to outside researchers who have obtained the appropriate authorization. While the amount of data made available has increased dramatically in recent years, relatively little has been done to facilitate phenotype harmonization across studies. Researchers currently have to analyze and compare increasingly larger numbers of variables with varying degrees of documentation associated with them to obtain the desired information. Information Explorer is a suite of tools we are developing that would allow researchers to (1) Quickly obtain the information needed to assess whether a specific study will be useful for the hypothesis of interest; (2) Exclude variables that do not meet research criteria; (3) Ascertain which studies have combinations of phenotype and genetic information of interest; and (4) More easily expand research questions beyond the most basic main-effects to more complex analyses such as gene-by-environment interactions and multivariate tests incorporating multiple phenotypes. The increased utility will also enable larger meta-analyses to be performed, as researchers will be able to more quickly hone in on outcomes, exclusionary variables and covariates of interest, leading to increased statistical power to detect genetic associations.

1440T

Data reduction techniques to construct new cellular and whole system biomarkers for ageing. *A. Brown[1], Z. Ding[1], L. Parts[2], D. Glass[3], D. Knowles[4], P. Deloukas[1], E. Dermitzakis[5], M. McCarthy[6,7,8], T. Spector[3], J. Winn[9], R. Durbin[1], MuTHER Consortium.* 1) Wellcome Trust Sanger Institute, Cambridge, United Kingdom; 2) Canadian Institute for Advanced Research, University of Toronto, Toronto, Canada; 3) Department of Twin Research and Genetic Epidemiology, King's College London, London, United Kingdom; 4) University of Cambridge, Cambridge, United Kingdom; 5) Department of Genetic Medicine and Development, University of Geneva Medical School, Geneva, Switzerland; 6) Oxford Centre for Diabetes, Endocrinology & Metabolism, University of Oxford, Churchill Hospital, Oxford, United Kingdom; 7) Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom; 8) Oxford NIHR Biomedical Research Centre, Churchill Hospital, Oxford, United Kingdom; 9) Microsoft Research, Cambridge, United Kingdom.

Biological ageing is the progressive decline in the ability to withstand stress, damage and disease; resulting in degeneration. A consensus on how to measure the difference between biological age and elapsed age in years has yet to emerge. Here, we present a novel analysis to derive new biomarkers for ageing from gene expression in skin, and from a selection of ten phenotypes which all show strong correlations with chronological age. Applying guided factor analysis methods we constructed new phenotypes from linear combinations of skin gene expression data from the MuTHER project, so as to capture variation across groups of functionally related genes (pathway activations). We also apply a hierarchical model to ten ageing related phenotypes to create a new meta-phenotype representing biological age; this captures the common ageing structure across several biological systems, including the respiratory and musculo-skeletal systems. By associating the pathway activations with the ageing phenotypes and the derived biological age meta-phenotype, we see significant associations between three pathway activations and the biological age meta-phenotype. One of these concerns glutathione metabolism, and one is an immunity-related pathway. There is therefore the possibility of using gene expression in these pathways as biomarkers for ageing. This analysis also suggests that combining information across several systems can be informative of variation at the cellular level, and could shed light on the interdependencies in the processes which underlie ageing.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1441F

Prioritization of next-generation sequencing variants using data visualization. *A. Bigelow[1,2], M. Meyer[1], N.J. Camp[2].* 1) Scientific Computing and Imaging Institute, University of Utah School of Computing, Salt Lake City, UT; 2) Division of Genetic Epidemiology, University of Utah School of Medicine.

In complex traits, the underlying critical risk variants will likely include some that are involved in regulation and/or that influence risk via novel biological mechanisms. Such variants will not be conducive to discovery using more standard annotation techniques or formal algorithm-based probabilistic methods which focus on prioritization based on well-described biological mechanisms. Filtering based on multiple characteristics of potential importance, such as allele frequency differences from a background or sharing within a pedigree will likely be useful in identifying these more obscure risk variants; however, black-box hard-filters may inadvertently remove variants of potential importance. Data visualization can be extremely useful for informed heuristic prioritization. The observed structure of the data across multiple characteristics can be instructive and guide the sorting and prioritization of variants for follow-up functional studies.

We have developed an interactive visualization tool, *compreheNGSive*, to support investigation and prioritization of next generation sequence variants. The tool requires .VCF file(s) for variants and, optionally, additional variant-level and/or feature-level data. Variant level data can be categorical (e.g. annotations, sharing in a pedigree) or quantitative (e.g. correlations with known associated risk variants or a LOD score) and is read from columnar text format, such as .CSV. Feature-level data is read from .BED or .GFF format. The user defines sample groups for comparison (e.g. case group, control group, background population), which can be incorporated from one or multiple .VCF files. A target group is defined as the focus for comparisons (e.g. cases) and is used to identify the allele of interest at each locus. The viewer includes a scatterplot, parallel coordinates (to display all frequencies and other variant-level data) and a genome browser (for location, refseq gene context and other feature-level data). All displays are linked and update dynamically, and axes are interchangeable between views. Variant selection and provenance mechanisms, as well as the ability to cope with missing data attributes, are provided. Any selected subset can be exported in columnar text format. The software is written using Python and the Qt framework, which is compatible with Windows, Linux, Mac OS X, and potentially mobile operating systems, and is fully scalable to whole genome data.

1442W

Integrating genome-wide gene expression and genotype data to predict HDL cholesterol levels in the Cholesterol and Pharmacogenetic Study (CAP). *E. Holzinger[1,2], S. Dudek[1,2], A. Frase[2], M. Medina[3], R. Krauss[3], M. Ritchie[2].* 1) Ctr Human Gen Res, Vanderbilt Univ, Nashville, TN; 2) Pennsylvania State University, University Park, PA; 3) Children's Hospital Oakland Research Institute, Oakland, CA.

Technology is driving the field of human genetics research with advances in the ability to generate high-throughput data that interrogate various levels of biological regulation (genomic, transcriptomic, proteomic, etc). With this massive amount of data comes the important task of using powerful and creative bioinformatics techniques to sift through the noise to find true, meaningful signals that predict various human traits. A popular analytic method thus far has been the genome-wide association study (GWAS), which assesses the association of each DNA variation with the trait of interest. Unfortunately, GWAS has not been able to explain a substantial proportion of the estimated heritability for most common, complex traits. Due to the inherently complex nature of biology, this phenomenon could be a factor of the simplistic GWAS study design. A more powerful study design for this data may be a systems biology approach that integrates different types of data, or a meta-dimensional analysis. Low HDL cholesterol is a significant risk factor for cardiovascular disease. For this analysis, we integrated ~3 million SNPs and ~25,000 gene expression variables from 480 individuals that participated in the CAP simvastatin clinical trial to generate meta-dimensional models that predict HDL cholesterol levels. First, to reduce the noise in the data set, we used a statistical filtering method that allows for main and interaction effects. Next, we used the filtered variables (591 SNPs and 222 EVs)to find significantly enriched pathways that cluster into non-redundant biological functions. The most significant cluster involved cellular membrane pathways (p-value 1.8E-4). Cellular membrane processes play a major role in HDL-C synthesis and metabolism. Finally, to generate more parsimonious meta-dimensional models, we analyzed the filtered set of SNPs and expression variables using the Analysis Tool for Heritable and Environmental Network Associations (ATHENA). Our best model consisted of 3 expression variables and one SNP, which together explain 9.4% of the variation in HDL cholesterol in this cohort. This systems biology approach demonstrates that with the right methods, different types of high-throughput data can be integrated to generate meta-dimensional models that predict complex human traits.

1443T

ATOMIC - Assess genotype calling quality using R or Affymetrix' genotyping console software. *A. Ziegler, A. Schillert.* Inst Med Biometry & Statistics, Univ Luebeck, Luebeck, Germany.

Single nucleotide polymorphism (SNP) chips are used for the analysis of exomes, targeted SNP typing, and genome-wide association analysis. High data quality is required for especially for SNP chips focusing on rare variants because misclassified genotypes for only a few individuals can result in highly inflated type I errors. Several statistical approaches, such as collapsing methods for rare variants, gene-gene interaction studies, or data mining approaches also require extremely clean input data to avoid false positives. The best approach for quality control is the investigation of signal intensities, but this approach is computationally intensive and has therefore been limited to SNPs with positive association signals. We present the R package ATOMIC, an acronym for AuToMatic Inspection of Cluster plots. Its main features are 1) splitting of signal intensities into smaller, manageable datasets, 2) assessment of the clusterings utilizing the theory of cluster analysis, including tests for unimodality [1,2], an improved implementation of the ACPA [3] algorithm, and 3) modern graphical display of the resulting signal intensity plots [4]. The approaches implemented in the R package will also be available soon in Affymetrix' genotyping console software. [1] Larkin 1979, Behav Res Meth Instr 11:467–468 [2] Engelman and Hartigan 1969, J Am Stat Assoc 64:1647–1648 [3] Schillert et al. 2009, BMC Proc 3:S58 [4] Wickham, 2009, Springer New York.

1444F

Single Center Experience with Peripheral Blood Gene Expression Profiling in a Large Cardiovascular Cohort. *E.A. Burns[1], S. Feng[1], C. Haynes[1], K. Abramson[1], M. Chryst-Ladd[1], E.R. Hauser[1], L.K. Newby[2], S.G. Gregory[1], W.E. Kraus[1,2], S.H. Shah[1,2].* 1) Center for Human Genetics, Duke University Medical Center, Durham, NC; 2) Department of Medicine, Duke University Medical Center, Durham, NC.

BACKGROUND. Gene expression profiling can provide insight into disease processes. Whole genome expression microarrays provide high-throughput capabilities and comprehensive gene profiling. The overarching goal of our study was to assess the relation between gene expression profiles in blood and cardiovascular disease in a large cohort of 2022 individuals referred for evaluation of coronary artery disease at Duke University. Here we describe our experiences and obstacles in carrying out a large peripheral blood gene expression project with disease outcomes. METHODS/RESULTS. We selected a cohort of 2022 sequential individuals referred for diagnostic coronary artery catheterization at Duke University Hospital. RNA was extracted from frozen peripheral blood that had been stored in PAXgene tubes. RNA from 2022 samples was extracted at two different facilities, of those 298 (14.7%) failed extraction, 116 (5.7%) had RIN scores of less than 6, leaving 1608 (79.5%) to be run on the Illumina HT-12 Gene Expression chips. Basic quality control (QC) assessment followed the manufacturer protocol and included basic analysis such as, gene intensity and probes detected, using the Illumina GenomeStudio program. Data was then analyzed at the plate, chip, and individual sample level to identify outliers identified by principal components analysis of the Illumina QC measures. Outliers were excluded at the plate level (195 [12.1%] samples removed); chip level (103 [6.4%] samples removed) and then at the individual level (27 [1.7%] samples removed). In total 325 (20.2%) samples were excluded in this QC process, leaving 1283 samples for further analysis. These 1283 samples were analyzed in batches of 897 and 386 samples in which they were hybridized to the Illumina Gene Expression chips. Analysis showed that RIN score prior to total RNA processing (TotalPrep) and analysis batch did not contribute to quality of the RNA expression signals. CONCLUSIONS. We describe the QC procedures and factors we found to be important in developing high quality gene expression datasets from peripheral blood. Our experience highlights the potential obstacles in dealing with large datasets and demonstrates how incorporation of multiple QC measures is necessary to ensure data quality. These expression data are being used concurrently with GWAs, DNA methylation, metabolomics and clinical outcomes data to discover human systems biology related to cardiovascular disease and its risk factors.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1445W

Best practices and joint calling of the Illumina HumanExome BeadChip: the CHARGE consortium. *M.L. Grove[1], B.J. Cochran[1], T. Haritunians[2], J.C. Bis[3], K.D. Taylor[2], M. Hansen[4], C.J. O'Donnell[5,6], J.I. Rotter[2], E. Boerwinkle[1], CHARGE Exome Chip Genotyping Committee.* 1) Human Genetics Center, University of Texas HSC at Houston, Houston, TX; 2) Medical Genetics Institute, UCLA Cedars-Sinai Medical Center, Los Angeles, CA; 3) Cardiovascular Health Research Unit, University of Washington, Seattle, WA; 4) Illumina, Inc., San Diego, CA; 5) Cardiology Division, Massachusetts General Hospital, Boston, MA; 6) Framingham Heart Study of the National, Heart, Lung, and Blood Institute, Framingham, MA.

Objective: Exome and whole genome sequencing is becoming more affordable and allows for the detection of novel rare variation in the human genome. Yet, genotyping arrays are a more cost effective approach when typing previously-identified genetic polymorphisms in large populations. One limitation of SNP arrays with low population minor allele frequencies is the ability of clustering algorithms to accurately detect and assign genotype calls. Thus, participating studies in the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium consented to have their array data analyzed collectively in order to increase the ability to accurately genotype rare variants. **Methods:** Approximately 60,000 ethnically diverse samples from the following eleven CHARGE cohorts were genotyped with the Illumina HumanExome BeadChip across seven genotyping centers: Age, Gene/Environment, Susceptibility—Reykjavik (AGES) Study, Atherosclerosis Risk in Communities (ARIC) Study, Cardiac Arrest Blood Study (CABS), Cardiovascular Health Study (CHS), Coronary Artery Risk Development in Young Adults (CARDIA), Erasmus Rucphen Family (ERF) Study, Multi-Ethnic Study of Atherosclerosis (MESA), Family Heart Study (FamHS), Framingham Heart Study (FHS), Health, Aging, and Body Composition (HABC) Study, Jackson Heart Study (JHS) and Rotterdam Study (RS). Each cohort genotyped a common set of unrelated HapMap samples for quality control and determination of batch effects. The raw data files for the samples were transferred to a central location and assembled into a single project for joint calling. Performance of the experiment was tested by examining the concordance of genotypes in a subset of individuals with both exome array and exome sequence data. **Results:** Of the 247,870 markers on the array, 247,485 had call rates > 95%. After exclusion of low performing SNPs on the array and non-overlap of SNPs derived from sequence data, 191,557 genotypes were compared in 530 individuals (85,867 were monomorphic). A total of 101,525,210 pairs of genotypes were tested and 99.7% matched, 0.17% had missing data, and 0.13% was discordant. Mismatches were attributed equally to both sequencing and genotyping which suggests the small number of miscalls are due to the random result of genetic detection technology. **Conclusion:** These data document the ability to accurately genotype rare variation using array technology when a large sample size is available and best practices are followed.

1446T

A Supervised Approach for Filtering and Genotyping High Quality Short Indels from Next Generation Sequencing Data. *A. Tan, H.M. Kang.* University of Michigan, Ann Arbor, MI.

Identifying short indels from Next Generation Sequencing (NGS) technologies is much more challenging than single nucleotide variant detection due to complex patterns of systematic errors in sequencing and alignment procedures. Current available indel calling algorithms produce highly discordant call sets, often showing less than 50% concordance between different algorithms on the same set of mapped reads. Extracting high quality indels from one or more call sets with minimal loss of sensitivity is very important to reliably and powerfully identify potentially causal variants in sequenced-based association studies. We developed and implemented a novel method for extracting high-quality short indels from arbitrary sets of candidate indels by utilizing available resources from independent technologies such as Affymetrix Exome Genotyping Arrays or Complete Genomics' public genomes. By annotating each candidate indel variant with informative sequence-based features such as allele balance, excessive heterozygosity, complexity of sequence context and cycle bias, our method builds a support vector machine trained by the external data. Applying our method to 1000 Genomes Phase 1 indel calls with Affymetrix Exome Array training data, we reduced the estimated false positive rate from 13% to 5%. We found that features like posterior probability of likelihood of indels, allele balance, and excessive heterozygosity helped improve the quality of Indels. Our method is publicly available through a user-friendly software pipeline that provides genotyping and variant filtering simultaneously from candidate indels and aligned sequence reads. Along with the ongoing efforts to improve existing individual indel calling algorithms, we anticipate that our method will contribute to increasing our understanding and discovery of genetic risk factors beyond single nucleotide variants.

1447F

Merging Genomic Data for Research in the Electronic MEdical Records and GEnomics Network: Lessons learned in eMERGE. *M.D. Ritchie[1], S. Setia[1], G. Armstrong[1], L. Armstrong[2], Y. Bradford[3], D.C. Crawford[3], D.R. Crosslin[4], M. de Andrade[5], K. Doheny[6], M.G. Hayes[2], G. Jarvik[4], I.J. Kullo[5], R. Li[7], T. Manolio[7], M. Matsumoto[5], C.A. McCarty[8], D. Mirel[9], S. Nelson[4], L. Olson[3], E. Pugh[6], S. Purcell[10], G. Tromp[11], J.L. Haines[3].* 1) Biochemistry & Molecular Biology, Pennsylvania State University, University Park, PA; 2) Northwestern University, Chicago, IL; 3) Vanderbilt University, Nashville, TN; 4) University of Washington, Seattle WA; 5) Mayo Clinic, Rochester, MN; 6) CIDR, Johns Hopkins; 7) NHGRI, Bethesda, MD; 8) Essentia Rural Health, Duluth, MN; 9) Broad Genotyping Center, Boston, MA; 10) Mt Sinai, New York, NY; 11) Geisinger Medical Center, Danville, PA.

Biobanks linked to electronic health records (EHR) is an emerging area of research for dissecting the architecture of complex human traits. Electronic phenotyping algorithms are deployed in large EHR systems to “ascertain” samples for analysis. To achieve success, sample size is an important consideration to maximize statistical power. To this end, merging data generated from different genotyping platforms is often desired. The eMERGE network, an NHGRI funded initiative, has developed a pipeline for merging genomic data generated on a single platform as well as a new pipeline for merging data from different genotyping arrays based on imputation. eMERGE consists of seven sites, each with DNA databanks linked to EHRs. Over 42,000 samples have been genotyped using one of the available Affymetrix and Illumina genome-wide genotyping arrays. These data have been imputed using BEAGLE and the October, 2011 release of the 1000 Genomes cosmopolitan reference dataset. Because of the computational complexity of imputation and the large sample size of the merged eMERGE dataset (42,807 individuals), a distributed imputation pipeline was implemented. In this scheme, the genome was divided by SNPs into 30,000 marker “SNPlets” with 700 markers of overlap on each side, resulting in 510 SNPlets; the data were also divided into sample sets of n=2000 or less. This parallelized pipeline resulted in over 556 billion SNPs (more than 13 million per individual) based on hundreds of thousands of CPU hours. The dataset generated consists of genome-wide SNPs on thousands of individuals all linked to EHR systems where numerous phenotypes can be explored. The genotype data will be available for research in dbGAP. The lessons learned by this group of investigators will be valuable for the genomics community also dealing with the combining of large-scale genomic datasets. We will present the details of our imputation pipeline, including our quality control investigations in data of this scope and magnitude. This merged eMERGE dataset is an invaluable resource for the genomics community to discover genetic risk factors for common, complex diseases and pharmacogenomic traits.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1448W

Detecting Sample Contamination. *M. Flickinger[1], G. Jun[1], K.F. Doheny[2], J. Romm[2], K.N. Hetrick[2], G.R. Abecasis[1], M. Boehnke[1], H.M. Kang[1].* 1) Department of Biostatistics, University of Michigan, Ann Arbor, Michigan 48109-2029; 2) Center for Inherited Disease Research (CIDR), Institute of Genetic Medicine, Johns Hopkins University, Baltimore, MD.

While advances in genotyping and next-generation sequencing have resulted in higher throughput, lower costs, and reduced error rates, DNA sample contamination remains a threat to data quality. Here we describe methods to detect whether a DNA sample contains DNA from more than one individual. Our goal is to detect and quantify low-levels of contamination early in the analysis pipeline for sequencing and genotyping studies. Our methods allow us to estimate contamination given (a) aligned sequence data and genome-wide, array-based genotypes ; (b) aligned sequence data alone; and even (c) prior to sequencing, when only genotype array data are available. Our sequenced based methods assume that two samples are mixed and that the probability of observing a read from one sample is proportional to the mixture proportion. To estimate a contamination proportion, we maximize the likelihood of observing a particular set of reads for a given mixing proportion while also accounting for read quality scores. Detection of contamination using only genotype-array data can either be accomplished by modeling the observed probe intensities as a Gaussian mixture, or by performing a linear regression of the Illumina reported B-allele frequency against a reference population minor allele frequency. To validate the sequence-based methods, we contaminated reads in silico from pairs of CEU individuals sequenced as part of the 1000 Genomes Project to have specific, known levels of contamination ranging from 0.1%–50%. We observed a Pearson correlation coefficient between our estimated contamination proportion and the truth of r=.9996 using both genotype-array and sequence data and r = .9840 using sequence data alone. To validate the genotype-array methods, we analyzed mixtures of DNA samples from HapMap individuals created to have specific, known levels of contamination ranging from 0.5%–10%. We genotyped these samples using the Metabochip genotyping array (a custom Illumina iSelect SNP array with ~200,000 markers in regions previously shown to be associated with metabolic traits). The mixture and regression models produced highly concordant results and correctly identified at least 34 of the 36 mixtures. These methods reliably detect contamination as low as 1%. Application of these DNA contamination detection methods provides a sensitive approach to identify contaminated samples and to maximize data quality.

1449T

Regression Models to "Explain" Departure from Hardy-Weinberg Equilibrium. *D.J. Schaid, JP. Sinnwell, GD. Jenkins.* Health Sciences Research, Mayo Clinic, Rochester, MN.

Testing whether genotype frequencies depart from Hardy Weinberg equilibrium (HWE) plays a major role in evaluating the quality of measured genotypes; gross departures often result from genotype misclassification errors. It is now standard practice to test HWE for SNPs used for genome wide association studies, and discard those with extreme p-values (e.g. p-value < 1e{hyphen}6). Yet, it is not unusual to find a relatively large number of SNPs whose chi-square statistics do not fit the null expected distribution based on quantile-quantile plots, a finding that could result from population stratification. In this situation, it would be helpful to determine whether measured covariates (such as principal components, geographic coordinates, or ethnic group indicators) could "explain" departures from HWE; departure from HWE implies correlation of alleles within subjects. Explained departures could occur when allele frequencies depend on measured covariates, yet there is minimal residual correlation of alleles within subjects. To address this need, we have developed a two-level strategy that first robustly evaluates which covariates should be included in regression models (allowing for the possibility of residual correlation between alleles), and then tests whether the residual correlation significantly differs from zero after adjusting for covariates. The statistical properties of our approach are illustrated by simulations, along with applications to real data. Because our methods are based on standard software for generalized linear models, they can be readily applied in standard practice, adding to the array of useful tools to analyze and understand large-scale genomic data.

1450F

SNP genotyping is a valuable tool for assessing the quality of next generation sequencing data. *J. Park[1], B. Vecchio-Pagán[2], H. Cuppens[3], G. Cutting[1].* 1) Institute of Genetic Medicine, Johns Hopkins School of Medicine, Baltimore, MD; 2) Cellular and Molecular Medicine, Johns Hopkins School of Medicine, Baltimore, MD; 3) Center for Human Genetics, KULeuven, Leuven, Belgium.

Linkage analysis of families can localize rare risk alleles. However, identification of the risk alleles is usually challenging due to the magnitude of the genomic region that has to be investigated. Targeted sequencing using next-generation sequencing (NGS) technology affords a cost-effective method to find rare variants that underlie linkage signals. Comparison with array based genotyping provides an opportunity to assess to the accuracy of NGS results. To assess the quality of targeted sequencing data, we analyzed a 1.6Mb linkage region on 20q13.2 (LOD=5.03) that harbors a modifier of lung function in cystic fibrosis. Four families of 4 individuals (2 parents and 2 affected children) underwent sequencing in 2 batches of 2 families each. Data were generated on Illumina HiSeq2000 after target enrichment with Nimblegen custom array. Approximately 90.4% (4779/5291) and 81.4% (3194/3929) of the variants mapped to the targeted regions at a minimum read depth of 8x (common default threshold) in batch 1 and 2, respectively. On average, 1.02 ± 0.45 and 0.63 ± 0.33 variants were found per kb in each batch. The mean depth of off-target variants found on chr20 was significantly lower (3.5x and 3.7x) compared to that of on-target variants (34.9x and 17.9x) in the 20q13.2 region. Also, the mean mapping quality of off-target variants was lower (36.6 and 37.1) than that of on-target variants (46.8 and 47.1). Overall, poor capturing in the second batch of 8 individuals resulted in lower depth and more off-target variants compared to the first batch of 8 individuals. To evaluate the accuracy of the sequence, NGS calls were compared to the genotypes generated by Illumina 610K array. Of 307 calls from the array in the 20q13.2 linkage region, 296~305 calls were present in batch 1 at a minimum read depth of 8x, representing 98.4% of coverage. Furthermore, 98.9% of the genotype calls were consistent between the array and sequencing. The majority of the discordant calls were due to incorrect variant calling by SAMtools. As expected, the lower sequence coverage obtained for batch 2 resulted in fewer matched SNPs (222~281), representing 82.9% of coverage, and lower accuracy (95.4%). A high-density SNP array data is useful in assessing NGS data quality and evaluating the variants found in targeted sequencing. Based on the SNP array comparison, the minimum depth can be lowered below 8x for high quality NGS data or raised for poor quality data in future analysis.

1451W

Imputation quality thresholds for rare and common variants. *G. Pistis[1,2,3], C. Sidore[1,2,3], A. Mulas[1], M. Zoledziewska[1], R. Berutti[3,4], F. Reinier[4], MF. Urru[4], A. Maschio[1,2], M. Marcelli[4], A. Angius[1,4], C. Jones[4], G. Abecasis[2], S. Sanna[1], F. Cucca[1,3].* 1) Istituto di Ricerca Genetica e Biomedica-CNR, Monserrato (CA), 09042, Italy; 2) University of Michigan, School of Public Health, Ann Arbor, MI, 48109; 3) Università degli Studi di Sassari, Dip Scienze Biomediche, Sassari, 07100, Italy; 4) Center for Advanced Studies, Research and Development in Sardinia - CRS4, Italy.

Genotype imputation allows the analysis of variants discovered by sequencing to extend across much larger numbers of individuals. A key issue in these analyses is deciding which variants can be well imputed. Here, we use a large panel of 1,146 sequenced individuals to drive genotype imputation and investigate the behaviour of commonly used imputation quality thresholds. For this analysis, we compared imputed and real genotypes at 47,010 high quality SNPs from the ExomeChip, typed in 6,020 individuals from the SardiNIA study. We observed that for MAF ≥1% (28,267 SNPs), the standard imputation quality cutoff of RSQR >0.3, suggested for the algorithms implemented in MACH and minimac, discards the highest proportion of bad quality SNPs (genotype-dosage correlation <0.5) while keeping the highest proportion of good quality ones (genotype-dosage correlation >0.9). Indeed, with this threshold, we would discard 96.7% of poor quality SNPs and keep 99.7% of those with good quality. On the other hand, if we consider less common variants (0.5%≤ MAF <1%) (4,071 SNPs), the same RSQR threshold will exclude only 89.7% of the low quality imputed markers. Using instead an RSQR >0.55, the percentage of low quality markers excluded increases to 97.2%, while losing only 2.9% of those imputed with high accuracy. Finally for rare variants (0.05%≤ MAF <0.5%) (9,266 SNPs), an RSQR >0.65 can guarantee 96.7% and 4.7%, respectively. Our results thus suggest that more restrictive quality cutoffs might reduce false positive signals in association analyses, particularly for imputed rare variants.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1452T

Error rates for very low MAF SNP's on the Illumina Exome Array. *E.W. Pugh[1], H. Ling[1], J.M. Romm[1], I.A. McMullen[1], J.R. Huyghe[2], M. Boehnke[2], K.F Doheny[1].* 1) Center for Inherited Disease Research, Johns Hopkins University, Baltimore, MD; 2) Department of Biostatistics and Center for Statistical Genetics, University of Michigan, Ann Arbor MI.

Some GWAS analyses filter low MAF SNPs due to concern about higher error rates and increased type I errors. As genotyping and analyzing low MAF SNPs is the focus of Illumina's Exome array, we wanted to estimate the error rates for the SNPs at different allele frequencies looking at both minor allele frequency and the count of genotypes not in the major homozygote cluster (nmgc). For this analysis we took a subset of samples from a large exome project of over 11,000 samples of European ancestry plus Hapmap samples. We used GenomeStudio version 2011.1, Genotyping Module version 1.9.4, GenTrain Version 1.0 to cluster the 87 unrelated HapMap samples one pair of HapMap duplicates and project together with 4111 randomly selected investigator samples. This created a 4200 sample project, which is our median project size for exome content array projects. To assess error in multiple ways, we then applied the clustering from the 4200 samples to 142 additional HapMap samples, run with the project and computed reproducibility and Mendelian inheritance rates. We manually reviewed cluster plots for subsets of the SNPs to determine the proportion that appeared poorly clustered such that a genotype would be misspecified. We also compared the genotypes to 1KGP (Oct 2011 release) for 62 out of the 87 samples that are present in1KGP. In the 4,200 sample data set out of the 247,870 SNP on the exome array, 1,259 (0.5%) SNPs were not called, 131,239 (53%) were monomorphic, 26,271 (11% maf .0001-.0003) were singletons, and 11,051 (4% maf .0002–.0005) were doubletons. There were 13683, 11899, 9568 SNPs with nmac of three to five (maf .0003–.0012), 6–15 (maf .0007–.0035), and 16–50 (maf .0018–.0117) respectively. The median number of singletons heterozygote SNPs per HapMap population for the unique samples were: CEU:38, TSI: 57.5, GIH:42.5, MXL:96, CHB: 147 JPT:159, YRI: 313. The proportion of SNPs with one or more Mendelian inheritance, reproducibility error or genotype discordant with sequencing increased with the count of non major genotypes. This may reflect decreased power for these methods to detect errors rare SNPs. Preliminary results of manual review suggest that the proportion with errors is more consistent across the count of non major genotypes and that intensity based filters (for example low intensity values (R) for singleton heterozygote calls) may be useful to flag additional genotyping errors.

1453F

An Algorithm for Identifying High-Order Gene-Gene Interactions from Pairwise Statistical Epistasis Networks. *T. Hu, A.S. Andrews, M.R. Karagas, J.H. Moore.* Department of Genetics, Institute for Quantitative Biomedical Sciences, The Geisel School of Medicine, Dartmouth College, Lebanon, NH.

Bladder cancer is the fourth most common malignancy in western countries. Besides the major contributor of tobacco smoking, hereditary factors have also been suggested playing an important role causing bladder cancer. Many single-locus biomarkers have been reported but could only be used to explain very limited heritability of this complex disease. The interaction effects among multiple genes, i.e., epistasis, could lead provide insights into the complex genetic architecture of bladder cancer. Detecting and characterizing epistasis is computationally challenging due to the large number of combinations in genome-wide association study (GWAS) data. Previously we identified a statistical epistasis network (SEN) using pair-wise interactions among approximately 1500 single nucleotide polymorphisms (SNPs) spanning nearly 500 cancer susceptibility genes in a large population-based study of bladder cancer in the state of New Hampshire. Unlike many existing techniques, SEN provides a model-free and global pair-wise interaction landscape. Here, we used the structural information from this pairwise SEN and evaluated all the three-way interactions that are directly connected in the network using multifactor dimensionality reduction (MDR). Using the pairwise SEN as a guide, we were able to find significant evidence for three-ways gene-gene interactions. The pairwise SEN significantly reduced the computational complexity of exhaustive evaluation all possible three-way combinations. In addition, the model we found also outperformed the results of a number of other commonly used filters such as ReliefF or Chi-square. Our study suggests that using knowledge of pairwise epistasis may aid in the detection and characterization of high-order interactions in GWAS data.

1454W

Fine-Mapping in a Covariate-based Genomewide Linkage Scan of Lung Cancer Susceptibility. *C.L. Simpson[1], T. Green[1], B. Doan[1,2], C.I. Amos[3], S.M. Pinney[4], E.Y. Kupert[4], M. de Andrade[5], P. Yang[5], A.G. Schwartz[6], P.R. Fain[7], A. Gazdar[8], J. Minna[8], J.S. Wiest[9], H. Rothschild[10], D. Mandal[10], M. You[4], T.A. Coons[11], C. Gaba[12], M.W. Anderson[4], J.E. Bailey-Wilson[1].* 1) Inherited Disease Res Branch, NHGRI, NIH, Baltimore, MD; 2) Johns Hopkins School of Medicine, Baltimore, Maryland; 3) Department of Epidemiology, University of Texas, M.D. Anderson Cancer Center, Houston, Texas; 4) Medical College of Wisconsin, Milwaukee,Wisconsin; 5) Department of Health Sciences Research, Mayo Clinic Rochester, Minnesota; 6) Karmanos Cancer Institute, Wayne State University, Detroit, Michigan; 7) University of Colorado, Denver, Colorado; 8) University of Texas Southwestern Medical Center, Dallas, Texas; 9) National Cancer Institute, NIH, Bethesda, Maryland; 10) Louisiana State University Health Sciences Center, New Orleans, Louisiana; 11) Saccomanno Research Institute and John McConnell Math & Science Center of Western Colorado, Grand Junction, Colorado; 12) University of Toledo, Toledo, Ohio.

Lung cancer (LC) is a leading cause of death in the developed world, with over 160,000 deaths expected in the US in 2012. Environmental risk factors such as smoking and asbestos exposures are well known. However, only 15% of smokers develop LC, suggesting genetic effects or gene-environment (GxE) interactions. We previously mapped a major LC susceptibility locus to 6q23–q25, and discovered a rare risk haplotype in linked families that exhibits a GxE interaction between the 6q susceptibility locus and smoking. Genome-wide association studies have suggested other loci with common alleles of small effect on LC risk. However, these loci do not explain all familial risk of LC, suggesting that additional risk alleles exist. We have also found additional susceptibility loci using linkage analysis including environmental covariates on 6p (LOD=5.75, 74cM) and 6q (LOD=3.25, 173cM), with novel evidence of linkage on 12q24 (LOD=5.46, 150cM) and 22q11 (LOD=5.19,10cM). Linkage to lung and throat cancer was observed on 9p21 (LOD=4.97, 64cM). All analyses were on microsatellite data. Here we present the results of a fine-mapping linkage analysis, with data from the microsatellite study combined with a dense SNP map. The data were checked for Mendelian inconsistencies and low call rate and the marker allele frequencies were estimated from the data. Linkage analyses of LC (adjusting for personal smoking) to the combined microsatellite/SNP dataset using LODPAL will be presented.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1455T

Meta-analysis of 2,526 individuals demonstrates strong evidence for gene-environment interaction between *NAT1* and tobacco smoke exposure in multiple sclerosis. F.*B.S. Briggs[1], B. Acuna[2], L. Shen[2], P. Ramsay[1], H. Quach[1], A. Bernstein[2], I. Kockum[3], L. Alfredsson[4], T. Olsson[3], C. Schaefer[2], L.F. Barcellos[1].* 1) Genetic Epidemiology and Genomics Laboratory, Division of Epidemiology, School of Public Health, University of California, Berkeley, Berkeley, CA; 2) Division of Research, Kaiser Permanente, Oakland, CA; 3) Center for Molecular Medicine, Department of Medicine, Karolinska Institute, Sweden; 4) Institute of Environmental Medicine, Karolinska Institute, Sweden.

Multiple sclerosis (MS) is a neuro-inflammatory autoimmune disease with a multifactorial etiology, involving both genetic and environmental components. Exposure to tobacco smoke has consistently demonstrated association with increased MS risk. However, not all smokers develop MS, and only some individuals with MS were ever smokers. We hypothesized that host genetics contribute to metabolism of tobacco smoke constituents, and thus, the development of MS in smokers. A gene-environment (GxE) investigation was conducted to assess whether variation within *NAT1*, *NAT2*, *GSTM1*, *GSTP1* and *GSTT1*, modifies the MS risk conferred by tobacco smoke. These loci are known for their critical role in tobacco smoke constituents metabolism, and their variation has been well-established in several cancers. The discovery data set used comprehensive smoke exposure (including smoking at age 20 and childhood passive smoke exposure) and genetic data collected from ~1,600 White MS case and control members of the Kaiser Permanente (KP) Northern California Region HealthPlan. The replication data set used a study population of ~1,000 White MS cases and controls from Sweden (SW). A history of tobacco smoke was significantly associated with greater MS risk (KP: odds ratio [OR]=1.3, p=0.005; SW: OR=1.4, p=0.009). In the discovery analysis, logistic models adjusted for age, gender, education, and population ancestry, investigated GxE interactions using dense genetic data for the 5 loci (745 SNPs). Specific metabolism phenotypes were also explored (*GSTM1* and *GSTT1* null, *NAT1* and *NAT2* slow acetylator phenotypes). 101 NAT1 variants had significant interactions ($p<0.05$). 85 *NAT1* SNPs were in the replication data set. A variant (rs7388368) within a dense transcription factor (i.e. NF-YB, USF1, USF2) binding region had significant interactions in both data sets (KP: odds ratio of interaction (OR_I)=1.7, p=0.004; SW: OR_I=1.6, p=0.03). In the combined meta-analysis, rs7388368 had an OR_I=1.6, p=0.0005, conferring MS risk among smokers only. Similar results were observed for SNPs in linkage disequilibrium: rs4921877 (OR_I=1.4, p=0.01) and rs6586711 (OR_I=1.4, p=0.01); their minor alleles have been correlated with decreased *NAT1* expression in lymphoblastoid cells. We have identified and replicated the first non-MHC GxE interaction in MS, identifying *NAT1* as a MS susceptibility locus important in the context of the environment. These results provide novel insight to underlying disease mechanisms in MS.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1456F

Algebraic Statistics and Markov Bases in Gene-Environment Interactions. *M. Rao, S. Venkatesan.* Environmental Hlth, Univ Cincinnati, Cincinnati, OH.

The data of interest structurally has three components: Disease (Present or Absent); Genotype (AA, Aa, aa); Environment (categorized into a finite number of categories). For a random sample of subjects from a population of interest, data has been collected whose structure has been outlined as above. The goal is to test the null hypothesis of absence of gene-environment interaction. Classical tests such as Cochran-Mantel-Haenszel test are not applicable for a variety of reasons. We will introduce an analogue of Fisher's exact test, normally used in the context of contingency tables, in our data environment. We will outline how Algebraic Statistics, Markov bases, and Monte Carlo Markov chain simulations can be integrated in the development of our exact test. The test is applied to the data collected from an island population, where the disease is Metabolic syndrome and environment is age, fasting glucose level, or a composite index of their dietary habits. A choice number of SNPs are investigated.

1457W

AprioriGWAS, a frequent itemset mining approach, detects angiogenesis gene interacting with CFH in age-related macular degeneration. *Q. Zhang, Q. Long.* Gregor Mendel Institute, Vienna, Austria.

Identifying gene-gene interaction is a hot, albeit difficult, topic in genome-wide association studies. There are two fundamental challenges: (1) How to smartly identify combinations of variants that are associated with a trait from the astronomical number of possible combinations; and (2) how to test for the presence of interaction when large numbers of associated combinations have been identified. We developed an algorithm called AprioriGWAS, which brings two innovations: (1) It is based on Apriori, a successful method in the field of frequent itemset mining in which a specific pattern growth strategy is leveraged to reduce the search space. It can enormously reduce the number of possible genotype patterns to be checked. (2) We developed conditional permutation to test the interaction effect between variants to reduce the artifacts of association between genotype patterns and phenotype outcomes generated by one variant with high marginal effect. Using simulations following Marchini et al's procedure, it is demonstrated that AprioriGWAS can approximately achieve the coverage of all associated pattern as exhaustive search, but with much less CPU time. Also, as a benefit of conditional permutation, AprioriGWAS has much higher power than logistic regression based on exhaustive search when the type I error is adjusted to be the same. We applied AprioriGWAS to age-related macular degeneration (AMD), which has been deemed a good example of a small number of variants explaining a large proportion of heritability. We found that ANGPT1, BBS9, MPP7, MED27, CHRM2 are interacting with a well-known AMD gene, Complement Factor H (CFH). An exciting finding is ANGPT1, which is a protein with important roles in vascular development and angiogenesis. Drugs for anti-angiogenic activity have been approved by the FDA in the United States for the treatment of cancer and AMD. As far as we know, its interaction with CFH is reported for the first time here. Other genes, e.g. BBS9, MPP7, MED27 are also reported to be important for retinal disease. We also found LPHN2 and PLXNA4A, that both play roles in nervous system are interacting, despite an absence of a marginal effect. Genes with similar functions are interacting with each other. Overall, the potentially interacting genes are enriched in the glycosaminoglycan biosynthetic process (P=1.39E-06). In conclusion, interactions in AMD pathways are abundant and AprioriGWAS is a good tool to find epistasis from GWAS.

1458T

Genotype-by-environment interaction for multivariate environments using the Mahalanobis distance. *V.P. Diego, H.H.H. Göring, J. Blangero.* Genetics, Texas Biomedical Research Institute, San Antonio, TX.

Complex traits are determined largely by genetic and environmental effects and their genotype-by-environment interactions (GEI). Here we propose a novel approach for modeling the role of GEI in complex traits in relation to multivariate environments in human pedigrees. We employ multivariate dietary measures as our environment in an example. To quantify an individual's total dietary environment, we used food frequency questionnaires to obtain measures on total-carbohydrate, -protein, -saturated fat, -unsaturated fat intake, and other indices. These variables together make up a multivariate environment. We developed a method that uses the Mahalanobis distance (MD) between each pair of individuals based on their vectors of dietary variables. The resulting MD matrix contains all pair-wise dissimilarities amongst individuals in the sample. It is employed in the context of a stochastic Ornstein-Uhlenbeck process to model shared environmental covariation as a measured environmental kernel for structuring a dietary-derived variance component. We can then extend this model to allow for GEI by adding another variance component representing the interaction between additive genetic effects and multivariate dietary effects. Under this approach, we are also able to model an across-environment genetic correlation. The two GEI null hypotheses under this model are an interaction variance equal to 0 (IV0), and a genetic correlation equal to 1 (GC1). The first null implies that there is no interaction effect whereas the second null implies that the same genes are involved across environmental contrasts. Rejection of either or both is evidence of GEI. As an example, we tested for GEI in relation to the MD representation of 19 diet-related variables in large Mexican American pedigrees from the San Antonio Family Heart Study. Of the 12 metabolic syndrome traits examined, we found evidence of GEI for serum leptin (IV0 $p = 0.02$; GC1 $p = 0.01$), systolic blood pressure (IV0 $p = 0.03$; GC1 $p = 0.04$), and high-density lipoprotein cholesterol (IV0 $p = 0.02$). We also show how this approach can be extended to utilize high-dimensional gene expression data. A straightforward application is to treat a given gene expression network (GEN) as an environment, and to use this model to search for potential GEIs with a focal clinically-related phenotype. This general framework for the measurement of multivariate environments is being incorporated into our statistical genetics software, SOLAR.

1459F

Interactions between maternal genotypes and metabolites are associated with incidence of congenital heart defects. *S. Erickson[1, 2], S. Chowdhury[3], S.L. MacLeod[2], M.A. Cleves[2], S. Melnyk[2], S.J. James[2], P. Hu[4], C.A. Hobbs[2].* 1) Department of Biostatistics, University of Arkansas for Medical Sciences, Little Rock, AR; 2) Department of Pediatrics, University of Arkansas for Medical Sciences, Little Rock, AR; 3) Clinical Molecular Genetics Department, Providence Sacred Heart Medical Center, Spokane, WA; 4) Buccaneer, a General Dynamics Company, Little Rock, AR.

The development of non-syndromic congenital heart defects (CHDs) involves a complex interplay of genetics, metabolism, and lifestyle. Previous studies have implicated maternal single nucleotide polymorphisms (SNPs) and altered metabolism in folate-related pathways as CHD risk factors. We sought to discover associations between maternal SNPs and metabolites involved in the homocysteine, folate, and transsulfuration pathways, and determine if these associations differ between CHD cases and controls. Genetic, metabolic, demographic, and lifestyle information was available for 335 mothers with CHD-affected pregnancies and 263 mothers with unaffected pregnancies. Analysis was conducted on 1160 SNPs, 13 plasma metabolites, and 2 metabolite ratios. A two-stage multiple linear regression was fitted to each combination of SNP and metabolite/ratio. We identified 4 SNPs in the methionine adenosyltransferase II alpha (MAT2A) gene that were associated with methionine levels. Three SNPs in tRNA aspartic acid methyltransferase 1 (TRDMT1) gene were associated with total plasma folate levels. Glutamylcysteine (GluCys) levels were associated with multiple SNPs within the glutathione peroxidase 6 (GPX6) and O-6-methylguanine-DNA methyltransferase (MGMT) genes. The regression model revealed interactions between genotype and case-control status in the association of folic acid, total glutathione (GSH), and free GSH, to SNPs within the MGMT, 5,10-methenyltetrahydrofolate synthetase (MTHFS), and catalase (CAT) genes, respectively. Our study provides further evidence that genetic variation within folate-related pathways accounts for inter-individual variability in key metabolites. We identified specific SNP-metabolite relationships that differed in mothers with CHD-affected pregnancies, compared to controls. Our results underscore the importance of multifactorial studies to define maternal CHD risk.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1460W

A unified generalized multifactor reduction method for detecting gene-gene interactions in family and unrelated samples with application to nicotine dependence. *X. Lou[1], G. Chen[1], N. Liu[1], Y. Klimentidis[1], X. Zhu[2], D. Zhi[1], X. Wang[3].* 1) Dept Biostatistics, Univ Alabama Birmingham, Birmingham, AL; 2) Dept Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, OH; 3) Dept Physics and the Comprehensive Diabetes Center, Univ Alabama Birmingham, Birmingham, AL.

Gene-gene and gene-environment interactions govern a substantial portion of the variation in complex traits and diseases. In convention, a set of either unrelated or family samples are used in detection of such interactions; even when both kinds of data are available, the unrelated and the family samples are analyzed separately, potentially leading to loss in statistical power. In this report, based on the theory of principal components analysis, we propose a generalized multifactor dimensionality reduction method that unifies analyses of nuclear families and unrelated subjects within the same statistical framework to detect gene-gene interactions. Through comprehensive simulations, we demonstrate that the proposed method can remarkably increase power by pooling unrelated and offspring's samples together as compared to each of two individual analysis strategies while it retains a controlled type I error rate in the presence of population structure. In application to a real dataset, we detected two significant (after Bonferroni correction) tetragenic interactions among CHRNA4, CHRNB2, BDNF, and NTRK2 associated with Fagerstorm Test for Nicotine Dependence in the Study of Addiction: Genetics and Environment (SAGE) sample, suggesting the biological role of these genes in nicotine dependence development. (This study is being supported by NIH grants DA025095, GM081488, GM077490, HG003054, and DK080100.).

1461T

The 5p15 locus is associated with bronchial hyperresponsiveness in siblings unexposed to tobacco smoke in early life in French families ascertained through asthma. *R. NADIF[1,2], P. MARGARITTE-JEANNIN[3,4,5], E. BOUZIGON[3,4,5], F. KAUFFMANN[1,2], M. LATHROP[6], F. DEMENAIS[3,4,5], M-H. DIZIER[3,4,5].* 1) INSERM U1018, CESP Centre for research in Epidemiology and Population Health, Respiratory and Environmental Epidemiology Team, F-94807, Villejuif, France; 2) Univ Paris-Sud 11, UMRS 1018, F-94807, Villejuif, France; 3) INSERM, U946, F-75010, Paris, France; 4) Univ Paris Diderot, Sorbonne Paris Cité, Institut Universitaire d'Hématologie, F-75007, Paris, France; 5) Fondation Jean Dausset-Centre d'Etude du Polymorphisme Humain (CEPH), Paris, F-75010, France; 6) Commissariat à l'Energie Atomique, Centre National de Génotypage, F-91057 Evry, France.

A genome-wide linkage analysis conducted in the French Epidemiological study on the Genetics and Environment of Asthma (EGEA) detected four chromosomal regions linked to bronchial hyperresponsiveness (BHR) and interacting with exposure to tobacco smoke (ETS) in early life (Dizier et al, 2005, 2007). Among these four regions, 1q43–q44 and 4q34 showed linkage to BHR in presence of ETS while 5p15 and 17p11 showed linkage in absence of ETS. Our goal was to conduct fine-scale mapping of these 4 regions to identify the genetic variants interacting with ETS that have an effect on BHR. Association analyses using the family-based association test (FBAT) were conducted in the whole sample of 388 EGEA families (322 sibs with BHR), and in sub-sets stratified according to ETS exposure (138 unexposed sibs with BHR and 184 exposed sibs with BHR). We used a two-step strategy including: 1) selection of SNPs showing association signals with BHR ($p<5\times10\text{-}3$) in the subset where linkage was detected; 2) test of homogeneity of association of selected SNPs with BHR between exposed and unexposed siblings using FBAT to detect SNPxETS interaction. We then applied logistic regression analysis to those variants showing significant interaction with ETS by FBAT, as internal validation to confirm the interaction. In the 5p15 and 17p11 regions, 18 and 15 SNPs showed association signals with BHR ($p<5\times10\text{-}3$) in unexposed siblings, while in the 1q43–q44 and 4q34 regions, 27 and 8 SNPs showed association in exposed siblings. Among these SNPs, only one in 5p15 showed significant evidence for interaction with ETS in the stratified analysis ($p=8\times10\text{-}4$ before correction and $p=0.05$ after correction for multiple testing). Logistic regression provided some support for this interaction ($p=0.11$, in 490 siblings with 63% of them having BHR) Although replication of our results is needed, it is interesting to note that the 5p15 locus that is found associated with BHR in ETS-unexposed siblings has been previously reported to be associated with lung cancer in non-smoking women.

1462F

Systematic identification of interaction effects between genome- and environment-wide associations in Type 2 Diabetes Mellitus. *C.J. Patel[1], R. Chen[1], K. Kodama[1], J.P.A. Ioannidis[2], A.J. Butte[1].* 1) Division of Systems Medicine, Department of Pediatrics, Stanford University School of Medicine, Stanford, CA; 2) Stanford Prevention Research Center, Department of Medicine, and Department of Health Research and Policy, Stanford University School of Medicine, Stanford, CA.

Type 2 Diabetes (T2D) may result from the interplay of both environmental and genetic factors. Genetic loci discovered through genome-wide association studies (GWAS) explain only a small portion of the disease risk variance and some of the unexplained risk may be due to gene-environment interactions. However, documenting gene-environment interactions has been difficult due to lack of concurrent genetic and environmental measurements, factor selection bias, and lack of multiple hypothesis control. We address these challenges by focusing on risk factors that emerge from GWAS and environment-wide association studies (EWAS) for T2D. Using data from 2 surveys (N=1003–3201) from the National Health and Nutrition Examination Survey (NHANES), we systematically screened for statistical interaction in association to T2D 18 genetic loci that have reached genome-wide significance in GWAS along with 5 environmental factors, including 2 pollutants and 3 dietary factors that had false discovery rate (FDR) < 1% in EWAS.T2D was determined by fasting blood glucose (≥126 mg/dL) and environmental factors were measured in serum. We determined we had adequate power (median 70%) to proceed and we implemented 2 methods for multiple hypothesis control, including a false discovery rate (FDR) method via the parametric bootstrap. We adjusted all logistic regression models for age, sex, body mass index, and race. We conducted sensitivity analyses, verifying the strength of interactions by race group. In the end, we found 8 interactions with nominal $p<0.05$ of which 4 had FDR<20%. One interaction withstood Bonferroni correction: non-synonymous SNP rs13266634 (SLC30A8) and trans-β-carotene (Bonferroni-corrected interaction $p=0.006$, FDR<1.5%). The per-risk-allele effect sizes, after adjusting for age, sex, body mass index, and race for those with low trans-β-carotene levels (OR=1.8, 95% CI: 1.3, 2.6) were 40% greater than the marginal genetic effect size of the SNP. We also found an interaction between rs13266634 and γ-tocopherol, a nutrient factor found in corn oil (FDR<20%) imparting 20% higher risk than rs1326634 alone. We present an approach for systematically screening for interactions between genetic and environmental factors. Further prospective validation of the identified novel interactions and/or examination in model systems is warranted. Our results suggest re-framing GWAS with consideration of common environmental factors to identify larger effect sizes for T2D risk.

1463W

A New Bayesian Framework for Detecting Gene-Environment Interaction. *X. Wen[1], M. Stephens[2], P. Zheng[1].* 1) Dept of Biostatistics, University of Michigan, Ann Arbor, MI; 2) Dept of Statistics and Human Genetics, University of Chicago, Chicago, IL.

Study of Gene-Environment interaction is critical for understanding the biological processes of complex diseases. In this research, we propose robust and easy-to-interpret Bayesian statistical approaches for detecting potential gene-environment interaction. The proposed framework has two unique features that are distinctly advantageous to existing methodologies: 1. We propose a new statistical formulation to model potential GxE interactions. In particular, we subgroup the samples according to different environment exposures and quantify the interaction as the degree of heterogeneity of genetic effects in subgroups. We show the traditional GxE interaction model is a special case of this general formulation however with a stronger, non-robust assumption. 2. We provide flexible and powerful Bayesian priors to model the relationship of genetic effects in various environmental conditions and expected heterogeneity. This allows us not only to “test” for potential GxE interaction, but also to infer the interaction pattern by Bayesian model comparisons. We demonstrate our approach in detecting eQTLs in different cell environments varied by different Hormone treatment. These interaction eQTLs unveil the interplays between the gene regulatory network and environment conditions.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1464T

Overlapping G × G interactions associated with caries in permanent and primary dentition. *K. Cuenco[1], M.K. Lee[1], Z. Zeng[1], J.R. Shaffer[1], X. Wang[1], E. Feingold[1], D.E. Weeks[1], S.M. Levy[2], R.J. Weyant[1], R.J. Crout[3], D.W. McNeil[3], M.L. Marazita[1].* 1) Univ Pittsburgh, Pittsburgh, PA; 2) Univ Iowa, Iowa City, IA; 3) West Virginia Univ, Morgantown, WV.

Oral health affects quality of life and systemic disease at all ages. Caries disease risk may vary by topologic characteristics of the tooth. Genome-wide association studies of primary and permanent dentition separately have identified a few suggestive candidate genes by single variant examination, but no single SNP has reached genome-wide significance. Clinical effects may be the result of gene-gene interactions (GxG) that have not been investigated. We screened for GxG associated with surface-specific caries measures obtained from children and adults. Tooth surface assessment measurements were obtained for child and adult participants from the Center for Oral Health in Rural Appalachia (COHRA, n = 1017 adult, n = 547 child), and Iowa (IOWA, n= 459 child) studies. Subjects included were Caucasian. Clinical assessment of pit/fissure (PF) and smooth (SM) tooth surfaces was recorded for primary and permanent dentition separately. Four dependent variables (PF & SM primary, PF & SM permanent) were dichotomized and used for GxG assessments. Input candidate gene sets were previously identified top 50 ranked genes from single SNP GWAS of caries and additional genes identified from pathway analyses. GWAS used SNP genotypes generated (Illuimina 610-Quad) or imputed (~1.4 million SNPs) by the NIH Gene, Environment Association Studies Consortium. We used multifactor dimensionality reduction (MDR) software to identify GxG of potential importance. MDR regroups SNPs based on similar direction of effects and evaluates predictive ability of the SNP set to generate evidence of GxG. Putative GxG are being generated for each caries measure and similarity comparison. The identified GxG interactions provide more information about the underlying biology of caries. Supported by U01-DE018903, R01-DE 014899, R01-DE09551, and R01-DE12101.

1465F

Associations and Interactions of genetic polymorphisms in innate immunity genes with early viral infections and susceptibility to asthma and asthma-related phenotypes. *D. Daley[1], J.E. Park[1], J. He[1], J. Yan[1], L. Akhabir[1], D. Stefanowicz[1], A.B. Becker[2], M. Chan-Yeung[3], Y. Bossé[4,5], A.L. Kozyrskyj[6,7], A.L. James[8], A.W. Musk[9], C. Laprise[10,11], R.G. Hegele[12], P.D. Paré[1], A.J. Sandford[1].* 1) Dept Med, i-CAPTURE Ctr, Univ British Columbia, Vancouver, BC, Canada; 2) 2Department of Pediatrics and Child Health and Manitoba Institute of Child Health, University of Manitoba, Winnipeg, MB, Canada; 3) Occupational and Environmental Lung Disease Unit, UBC, Vancouver, BC, Canada; 4) Institut Universitaire de Cardiologie et de Pneumologie de Québec, Québec, Canada; 5) Department of Molecular Medicine, Laval University, Québec, Canada; 6) Department of Pediatrics, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, Alberta, Canada; 7) Department of Pediatrics and Child Health, Faculty of Medicine, University of Manitoba, Winnipeg, Manitoba, Canada; 8) Department of Pulmonary Physiology and Sleep Medicine, Sir Charles Gairdner Hospital, Perth, Western Australia; 9) Department of Respiratory Medicine, Sir Charles Gairdner Hospital, Western Australia; 10) Département des Sciences Fondamentales, Université du Québec à Chicoutimi, Saguenay, Québec, Canada; 11) Université de Montréal Community Genomic Medicine Center, Chicoutimi Hospital, Saguenay, Québec, Canada; 12) Department of Laboratory Medicine and Pathobiology, University of Toronto, Ontario Canada.

The innate immune system is essential for host survival because it recognizes invading pathogens and mounts defensive responses. We used four large samples (n=5565) to investigate associations of 321 single nucleotide polymorphisms (SNPs) in 26 innate immunity genes with four phenotypes: atopy, asthma, atopic asthma and airway hyper-responsiveness (AHR) in three Canadian family-based studies and one Australian population-based case control study. Interactions between innate immunity genes and early viral exposure to three common viruses (parainfluenza, respiratory syncytial virus and picornavirus) were examined using data collected from the Canadian Asthma Primary Prevention Study. Analyses were conducted using both an affected-only family-based transmission disequilibrium test and case-control methods. **Results:** Six SNPs (rs1519309 (TLR3), rs740044 (ILIR2), rs4543123 (TLR1), rs5741812 (LBP), rs917998 (IL18RAP), and rs3136641 (NFKBIB)) are significant (P<0.05, confirmed with 30,000 permutations) in both the combined analysis of main genetic effects and SNP*viral interaction analyses in both case-control and family based methods. The TLR1 variant (rs4543123) is associated with both multiple viruses (RSV and PIV) and multiple phenotypes. **Conclusion:** We identified susceptibility genes for asthma and related traits and interactions between these genes and early life viral infections that may prove to be of predictive value for the development of allergic diseases.

1466W

Evidence for Genetic Interactions Beyond Variability Explained by Single SNPs in Ulcerative Colitis. *M.H. Wang[1], X. Zhu[2], R.H. Duerr[3], C. Fiocchi[1], J.P. Achkar[1], The Wellcome Trust Case Control Consortium.* 1) Gastroenterology and Hepatology, Digestive Disease Institute, Cleveland Clinic, Cleveland, OH, United States; 2) Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, OH, United States; 3) Gastroenterology, Hepatology, & Nutrition, University of Pittsburgh, Pittsburgh, PA, United States.

BACKGROUND: A limited fraction of ulcerative colitis (UC) heritability is explained by current genome wide association studies (GWAS). Logic regression (LR) is a method that identifies multi-dimensional logical single nucleotide polymorphism (SNP) combinations and thus could explain phenotypic variability better than single SNP additive models. **MATERIALS & METHODS**: All UC-associated SNPs from the Wellcome Trust Case-Control Consortium (WTCCC Affy6.0; 2361 UC, 5417 controls) with nominal P values <10E-04 (i.e., 314 SNPs) were selected and recoded as binary predictors (dominant and recessive). We applied LR to identify multi-dimensional logical SNP combinations (called `Tree'). The odds ratio (OR) of the identified Tree was estimated by logistic regression and then we examined the same Tree in an independent replication cohort, the Cleveland Clinic/University of Pittsburgh (CC/UP) GWAS (Illumina Omni-1 Quad; 566 UC, 1428 controls). Ungenotyped markers in the replication cohort were imputed in BEAGLE using the 1000 Genome haplotypes data. The R package Logic.Reg was used to conduct LR analysis. Model comparisons were conducted through the ROC functions in SAS 9.2. Gene by gene interactions were tested using logistic regression models. **RESULTS**: In the WTCCC GWAS dataset, a Tree model built on 3 SNPs (rs16870123, *BTNL2*; rs3129868, *HLA-DRA*; rs9267798, *TNKB*) was identified through LR: rs3129868 in recessive mode or (rs16870123 [dominant] and rs9267798 [dominant]). The UC risk of carrying this Tree is OR 2.23 (95%CI:1.93–2.56), P=4.32×10E-28. After controlling the 3 SNPs, the adjusted OR is 2.46 (95%CI:1.89–3.21), P=1.25×10E-11. Compared with the additive mode of 3 SNPs, model predictability was significantly improved when adding the logical interactions term (i.e., Tree) (P=0.002).We examined the same Tree model in an independent CC/UP GWAS and found a similar significant association: OR 2.63 (95%CI:1.95–3.56), P=4.70×10E-10; adjusted OR 1.94 (95%CI:1.14–3.32), P=0.01. These 3 SNPs in the chr6p21 region are not in linkage disequilibrium (pairwise R^2 <0.2) and not in the same haplotype block. There are significant interactions between rs16870123 (*BTNL2*) and rs9267798 (*TNKB*) in the WTCCC and the replication cohort CC/UP ($P_{interaction}$ 2.34×10E-12 and 0.01, respectively). **CONCLUSION**: Interactions among SNPs can be identified through LR. Similar genetic interactions patterns were observed in the WTCCC and the independent replication cohort CC/UP GWAS datasets.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1467T

Influence of SNPxSNP interaction on BMI in European American adolescents: Findings from the National Longitudinal Study of Adolescent Health. *K.L. Young[1,2], M. Graff[1,2], K.E. North[1,3], A.S. Richardson[2,6], L.A. Lange[3,4], E.M. Lange[3,4], K.M. Harris[2,3,5], P. Gordon-Larsen[2,6].* 1) Department of Epidemiology, University of North Carolina, Chapel Hill, NC; 2) Carolina Population Center, University of North Carolina, Chapel Hill, NC; 3) Carolina Center for Genome Sciences, University of North Carolina, Chapel Hill, NC; 4) Department of Genetics, University of North Carolina, Chapel Hill, NC; 5) Department of Sociology, University of North Carolina, Chapel Hill, NC; 6) Department of Nutrition, University of North Carolina, Chapel Hill, NC.

Adolescent obesity is predictive of future weight gain, obesity, and adult-onset severe obesity (BMI ≥ 40kg/m^2). A greater understanding of underlying factors, including possible gene-gene interactions, can help explain the "missing heritability" in these complex traits, as well as identify potential biological mechanisms affecting weight gain during this critical developmental period. In adolescent participants of European descent (EA) from the National Longitudinal Study of Adolescent Health (N = 5072, ages 12–21, 52.5% female), we tested established BMI-related SNPs (Speliotes et al. 2010) for gene-gene interaction effects on BMI (measured height and weight). We used mixed effects regression, assuming an additive genetic model and adjusting for age, sex, and geographic region, and allowing random effects for clusters by family and school.

The prevalence of obesity was 12.3%. The overall discordance rate across SNPs was 0.3%, and the average call-rate 97.9%. SNPs that diverged from HWE (p<0.001) (1 SNP), showed no association with BMI in the main effects analyses (p>0.2) (17 SNPs), or had main effects which were directionally inconsistent with results from EA adults (1 SNP) were excluded, leaving 23 potential SNPs for analysis. Of 253 SNP*SNP interactions, two had p values <0.01: in/near *POMC-NEGR1* (p=0.004)and in/near *LZTR2-MC4R* (p= 0.005), and an additional nine had interaction p values <0.05. Six SNPs also had nominal evidence for interaction with at least two other loci. For example, there was nominal evidence for rs987237 (*TFAP2B*) interacting with rs2241423 (*MAP2K5*) (p = 0.016), rs4712652 (*PRL*) (p=0.013), and rs3810291 (*TMEM160*) (p=0.045) in association with BMI. We prioritized SNPs with nominally significant interactions and performed bioinformatic analysis to identify possible biological connections among positional candidate genes. Of the 11 nominally significant interactions, six were positive, suggesting that in the context of epistasis, the effect of the BMI increasing alleles is exacerbated. In contrast, five epistatic interactions reduced the effect of the single variant. While our study is limited by small sample size, our results are currently being replicated in whites from the Atherosclerosis Risk in Communities (ARIC) study, and are suggestive of epistatic effects on BMI during adolescence and point to potential underlying biological pathways important in the development of obesity.

1468F

Analyzing Genome-wide SNP interactions using the Random Forest Fishing method. *W. Yang[1], J. Chu[1], C.C. Gu[1,2].* 1) Division of Biostatistics, Washington University in St Louis, St Louis, MO; 2) Department Genetics, Washington University in St Louis, St Louis, MO.

Despite the importance of pervasive interactions among genetic and environment disease risk factors, genome-wide interactions analysis is seldom carried out, mainly due to the lack of powerful and efficient tools to handle the extreme noise level and computational requirement. We have therefore proposed a fast search method called Random Forest Fishing (RFF), which applies a genetic algorithm based on Random Forest, by repeatedly updating a set of predictors to find a variable set predictive of the disease outcome. This set captures both marginal and interactive effects in their collective effect when making predictions. RFF was applied to analysis hypertension status in a large scale genome-wide association study (GWAS) sample, which consisted of 2,768 African Americans (AA; 832,077 SNPs after QC), and 4,986 Caucasians (CA; 828,744 SNPs after QC). Analyses were performed within races, starting with initial SNP sets of either random SNPs (RFFr) or top SNPs in single SNP tests (RFFt). Briefly, (1) Even though started with different set of SNPs, after long search iterations, RFFr and RFFt results began to converge and had a large number of SNPs in common. In the AA sample, RFFt output 597 SNPs, and RFFr output 914 SNPs, with 393 in common; In CA, RFFt ouput 1071 SNPs, and RFFr 787 SNPs, with 395 SNPs in common. (2) Among the 149 blood pressure related genes reported by past GWAS studies (GWAS catalog as of 05/03/2012), 15 genes were re-discovered by RFFt in CA, and 8 re-discovered in AA. (3) Most RFF identified SNPs were likely to be missed in single SNP or pairwise interaction tests. Only 5 SNPs had single SNP test p values less than 1e-6 in both CA and AA, and the minimum p values for pairwise interaction test were 2.2e-7 and 1.7e-6 in CA and AA, respectively. (4) There were only 7 SNPs common to both races by RFFt. However, identified SNPs were in 222 common genes and 264 common pathways, indicating the same genes and pathways likely underlying hypertension risk. (5) Some know biology pathways were found enriched in the RFF results, such as GO:0044267 pathway (cellular protein metabolic process; with 46 SNPs in RFFt AA results, 64 in CA), and has04510 pathway (Focal adhesion, with 10 SNPs in RFFt AA result, 12 in CA). In summary, our test shows that RFF is able to identify sets of biologically interesting SNPs considering both marginal and interaction effects. RFF could serve as a first step to explore complex interactions in the GWAS studies.

1469W

Hunting for rare genetic variants in a complex disease? Do not forget about gene-environment interactions! *R. Kazma, J. Witte.* Department of Epidemiology and Biostatistics and Institute for Human Genetics, University of California San Francisco, San Francisco, CA.

The interest in evaluating the potential association between disease and rare genetic variants, measured by next-generation sequencing or exome chip arrays, is growing rapidly. A critical issue here is the limited power to detect associations due to the low frequency of genetic variants. To improve power, several methods, aggregating the information of rare variants over a region like a gene, have been proposed. However, none of these methods have considered potential gene-environment interactions. To determine if accounting for gene-environment interactions improves the power to detect associations with rare variants, we have extended three methods (Min-P, CAST, and SKAT) and compared their statistical properties under various simulated disease models. To test for association of a group of rare variants with the disease, Min-P uses permutations to determine the null distribution of the lowest P-value within the group of variants, CAST (Cohort Allelic Sum Test) uses an indicator variable of the presence of at least one rare allele (or the sum of the rare alleles) within the group of variants, and SKAT (Sequence Kernel Association test) uses a kernel machine approach. We extended these three methods by incorporating a term for the gene-environment interaction and testing for the genetic association and the gene-environment interaction simultaneously. Accounting for gene-environment interaction improves the power to detect association of complex diseases with rare genetic variants interacting with the environment, in particular if the interaction is strong and the genetic marginal effect is weak. In the absence of gene-environment interaction, the decrease in power of the test accounting for gene-environment interaction is small (1 to 5%) compared to the standard test of genetic association. SKAT (using a weighted linear kernel) had the best overall power and is the fastest method of those three. However, the kernel choice introduces implicit assumptions and can impact power. CAST was sensitive to the arbitrary choice of the threshold, below which variants are aggregated. Finally, Min-P had a power close to the power of SKAT but its long computational time makes it inappropriate for large data.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1470T

Searching for an enrichment of genetic interaction signals in SNPs within known protein-protein interactions. *J.E. Mollon[1,2], B. Lehne[2], S. Sacks[1], T. Schlitt[2], M.E. Weale[1,2].* 1) MRC Centre for Transplantation, King's College London, London, United Kingdom; 2) Medical and Molecular Genetics, King's College London.

The existence of interactions between single-nucleotide polymorphisms (SNPs) has been hypothesized as a possible contributor to the risk of complex disorders. This seems plausible since many physical interactions between proteins have been demonstrated through functional studies such as yeast two-hybrid screening (Y2H). However, detecting such physical interactions statistically at the SNP level in genome-wide association studies (GWAS) in complex disorders has proven difficult. The search space is very large, for example there are approximately 1.25 billion possible interacting SNP pairs from a 500k SNP genotyping panel. The interaction signal required to overcome the burden of genome-wide multiple testing corrections in such data would have to be very strong, and such strong signals have not yet been found. We created software that takes GWAS data, a list of paired SNPs and a SNP-to-gene map, and searches for a general enrichment of low, but not genome-wide significant, interaction p-values between the SNP pairs. We permuted the gene labels many times and recalculated the test statistic, thereby generating an empirical null distribution while maintaining the gene structure with in the data. Application of this method to the Wellcome Trust Case Control Consortium GWAS in Crohn's disease suggested some enrichment, but this depended on the threshold used to define a low interaction p-value for inclusion in our test statistic.

1471F

RAPID Detection of Gene-Gene Interactions Underlying Quantitative Traits. *N. Udpa[1], V. Bafna[1,2].* 1) Bioinformatics and Systems Biology, University of California at San Diego, La Jolla, CA; 2) Computer Science and Engineering, University of California at San Diego, La Jolla, CA, 92093.

In complex phenotypes, the trait under question may not be explained adequately by single locus effects - multiple loci might interact to confer the phenotype. For instance, two proteins may interact in a lock-and-key mechanism, and a mutation in one protein can be offset by a complementary mutation in the other. In this situation, a specific genotype value for either individual SNP does not matter as much as the relationship between the two genotypes. The problem with detecting these interactions is primarily computational - a naive way takes $O(nm^2)$, where n is the number of individuals and m is the number of variants. This is infeasible for a typical human GWAS, where $n \sim 10^3$ and $m \sim 10^6$. Previously, we developed a tool, RAPID (Brinza et al., Bioinformatics, 2010), which improves upon this significantly. RAPID takes advantage of the fact that we can represent genotypes as a series of binary vectors (where each dimension represents an individual haplotype). Given this framework, correlated loci are proximal, and the problem reduces to finding such adjacent pairs efficiently. RAPID uses a geometric approach (locality-sensitive hashing) to solve this without explicitly checking every pair of loci. Here we extend this approach to handle quantitative phenotypes, as opposed to the standard “case” and “control” delineations. The correlation between quantitative phenotypes and multilocus genotypes is usually performed using ANOVA statistics that do not lend themselves to geometric/discrete approaches. To get around this, we follow a two stage strategy: first, we describe an algorithm to partition the continuous phenotypes into two sub-populations under the assumption that the phenotypes are a mixture of Gaussians. Second, we use the RAPID framework to identify epistatic loci within each sub-population. We use simulated datasets to determine the best way to partition these points given different levels of mixture and relative abundance between the two phenotypes. In addition, we use simulations to explore the impact of the underlying genotype-phenotype models on the ability of RAPID to identify correlations. An application of this algorithm to the WTCCC hypertension datasets show the speed and power of our approach.

1472W

Evidence of gene-environment interaction for the chromosome 4 and environmental tobacco smoke in controlling the risk of cleft lip with/without cleft palate. *T. Wu[1], I. Ruczinski[2], J.C. Murray[3], M.L. Marazita[4], R.G. Munger[5], J.B. Hetmanski[2], T. Murray[2], R.J. Redett[6], M.D. Fallin[2], K.Y. Liang[2,7], P.J. Patel[2], S.C. Jin[2], T.X. Zhang[2], H. Schwender[2], Y.H. Wu-Chou[8], P.K. Chen[8], S.S. Chong[9], F. Cheah[9], V. Yeow[10], X. Ye[11,13], H. Wang[1], S. Huang[12], E.W. Jabs[13], B. Shi[14], A.J. Wilcox[15], R.T. Lie[16], S.H. Jee[17], K. Christensen[18], A.F. Scott[6], T.H. Beaty[2].* 1) Epidemiology, Peking University School of Public Health, Beijing, Beijing, China; 2) Biostatistics, Johns Hopkins University, Baltimore, MD; 3) University of Iowa, Children's Hospital, Iowa City, Iowa; 4) Center for Craniofacial and Dental Genetics, School of Dental Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania; 5) Utah State University, Logan, Utah; 6) Johns Hopkins University, School of Medicine, Baltimore, MD; 7) National Yang-Ming University, Taipei, Taiwan; 8) Chang Gung Memorial Hospital, Taoyuan, Taiwan; 9) National University of Singapore, Singapore; 10) KK Women's & Children's Hospital, Singapore; 11) Wuhan University, School of Stomatology, Wuhan, China; 12) Peking Union Medical College, Beijing, China; 13) Mount Sinai School of Medicine, New York, New York; 14) State Key Laboratory of Oral Disease, West China College of Stomatology, Sichuan University, Chengdu, China; 15) NIEHS/NIH, Epidemiology Branch, Durham, North Carolina; 16) University of Bergen, Bergen, Norway; 17) Yonsei University, Epidemiology & Health Promotion, Seoul, Korea; 18) University of Southern Denmark, Odense, Denmark.

Nonsyndromic cleft palate (CP) is one of the most common human birth defects and both genetic and environment risk factors contribute to its etiology. We conducted a genome-wide association study (GWAS) using 550 CP case-parent trios who were ascertained through subjects collected in an international consortium. Family-based association tests of single nucleotide polymorphisms (SNP) and four common maternal exposures (maternal smoking, alcohol consumption, multivitamin supplementation and environment tobacco smoke) were used in a combined 2 *df* test for gene (G) and gene-environment (GxE) interaction simultaneously, plus a separate 1 *df* test for GxE interaction alone. While no SNP achieved genome-wide significance when considered alone, markers in several genes attained or approached genome-wide significance among 253 Asian trios when GxE interaction was included. Among these, *SLC2A9* and *WDR1* on chromosome 4p16.1 showed multiple SNPs resulting in an increased risk if the mother exposed to environment tobacco smoke (ETS) during the peri-conceptual period (3 months prior to conception through the first trimester). When maternal ETS was considered, five of 137 SNPs mapping to *SLC2A9* gave strong evidence of GxETS interaction in the 1 *df* test using conditional logistic regression models ($P<10^{-6}$). SNP *rs3733585* and *rs12508991* yielded $P=9.63\times10^{-8}$ in the 1 *df* test for GxETS interaction. A cluster of 61 SNPs identified a region spanning 125 kb where evidence of GxETS interaction was apparent. In this region, nineteen of these SNPs yielded significance ($P<10^{-3}$) in both the 1 *df* test for GxE interaction alone and in the 2 *df* test for combined effects of G and GxE interaction. *WDR1* was located next to *SLC2A9* and it encompassed 65 SNPs. A block of 12 SNPs (spanning 63 kb) yielded seven SNPs with significant *P*-values in the 1 *df* test for GxE interaction when maternal ETS was considered. SNPs *rs717615* and *rs7699512* yielded $P= 8.30\times10^{-7}$ and $P=6.09\times10^{-7}$ in the 1 *df* test for GxE interaction in conditional logistic regression models and two adjacent SNPs *rs6834555* and *rs6820756* approached genome-wide significance (*rs6834555*, $P= 1.73 \times 10^{-6}$; *rs6820756*, $P= 1.37 \times 10^{-6}$). Another three SNPs located 147kb downstream from *rs7699512*, also yielded significant GxETS interaction at $P<10^{-4}$ level. These results emphasize the need to consider GxE interaction when searching for genes influencing risk to complex and heterogeneous disorders, such as nonsyndromic CP.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1473T

Pathway-based analysis of gene-gene interactions for complex diseases. *J.G. Zhang, J. Li, H.W. Deng.* Department of Biostatistics and Bioinformatics, Tulane University School of Public Health & Tropical Medicine, New Orleans, LA.

Background: Gene-gene interactions play an essential role in determining individual susceptibility to complex diseases, and many genes in the same pathway are involved in such interactions. Hence, identifying gene-gene interactions associated with complex diseases within gene pathways can help us better understand the disease genetics. However, current analytical methods for gene-gene interaction identification either limit their tests only on SNPs showing marginal effects, or may suffer from high computational complexity, especially for genome-wide association (GWA) data and next-generation sequencing (NGS) data. **Results:** In this study, we propose a novel method to identify disease-associated gene-gene interactions within biological pathways. The method first applies principal components analysis (PCA) to reduce the dimension of genetic variants within each gene of a pathway. Pairwise products of the first principal components (PCs) from all genes in a pathway were computed as interaction terms. Then the likelihood ratio test was applied to determine the statistical significance of interactions in the pathway. We have applied our method to GWA data set for studying osteoporosis and have detected 16 pathways enriched with gene interactions potentially contributing to the susceptibility of osteoporosis. **Conclusion:** Our new method proved fast and efficient for identifying pathways with gene-gene interaction effects associated with complex diseases. It will aid in the deciphering of the genetic mechanisms of complex diseases.

1474F

Identification of host gene-gene interaction in saliva flow and *Streptococcus mutans* using MDR. *M. Lee[1,2], K.T. Cuenco[1], Z. Zeng[3], X. Wang[1], J.R. Shaffer[3], E. Feingold[2,3], D.E. Weeks[2,3], R.J. Weyant[4], R.J. Crouot[5], D.W. McNeil[6], M.L. Marazita[1,3].* 1) Center for Craniofacial and Dental Genetics, Department of Oral Biology, School of Dental Medicine, University of Pittsburgh, Pittsburgh, PA; 2) Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA; 3) Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA; 4) Department of Dental Public Health and Information Management, School of Dental Medicine, University of Pittsburgh, PA; 5) Department of Periodontics, School of Dentistry, West Virginia University, Morgantown, WV; 6) Dental Practice and Rural Health, West Virginia University, Morgantown, WV.

Appalachian populations have limited healthcare access and experience more frequent negative health outcomes like oral disease than the general population. As part of its mission to identify risk factors for poor oral health, families recruited through the Center for Oral Health Research in Appalachia (COHRA) participated in the NIH Gene, Environment Association Studies Consortium (GENEVA) genome-wide association study (GWAS) of dental caries. Low saliva flow and risk of Streptococcus mutans infection have been associated with increased caries, each of which is a complex trait likely controlled by host genetics. We investigated host gene-gene interactions responsible for saliva flow and presence of S. mutans in the COHRA study population. We applied a gene-gene search method to saliva flow and Streptococcus mutans data from the COHRA-GENEVA population. 1506 Caucasian subjects of age 5 to 75 years were available with saliva flow measures. 1554 subjects had S. mutans data available. Saliva flow was categorized as abnormal (≤0.6 ml/min) or normal (>0.6 ml/min). 20 genes obtained from previous GWAS findings were selected for evaluation of gene-gene interactions with saliva flow and present of S. mutans . Gene-gene interactions were detected with multifactor dimensionality reduction (MDR) software, which implements a model free approach and k-fold cross-validation after reducing SNP categories by identifying high or low risk SNPs for an outcome and combining these SNPs into a single category. MDR successfully identified host gene-gene interactions associated with saliva flow and S. mutans. Identified host gene-gene interactions provide insight into additional gene relationships and help prioritize genes for further inquiry. Supported by: DE018903, DE014899.

1475W

Meta-analysis of gene by gene interaction. *RG. Hoffmann[1], K. Yan[1], TJ. Hoffmann[2], P. Simpson[1].* 1) QHS, Pediatrics, Medical College of Wisconsin, Milwaukee, WI; 2) Epidemiology and Biostatistics, University of California San Francisco, San Francisco, CA.

Purpose: Since gene by gene interaction always results in substantially smaller numbers in the cells corresponding to the interaction terms, correct estimation of the effect sizes and validation of the results will usually require combination of multiple studies. The purpose of this study was to compare methods for combining QTL studies with gene by gene interaction. Methods: The motivating data consists of age of onset of type I diabetes from two separate Scandinavian countries. Both cases and controls were available; controls were considered as censored at their current age. Thus both case-control and case-only studies could be performed. Methods that were compared were proportional hazards regression, binary logistic regression, and ordinal logistic regression for the case-control meta-analysis. For the case-only meta-analysis linear regression, proportional hazards regression, ordinal logistic regression and a generalized linear model (a lognormal model for when the data is skewed to the left). Simulations based on permuting the observed subjects were used to examine the effect of sample size on type I and type II errors without losing the inherent structure of the data. Results: Small studies tended to be very unstable and often missed the underlying gene by gene interactions. Studies in the 2000 subject range were still somewhat unstable; unless there were more studies to combine. Case-control studies were quite sensitive to the level of overlap of the QTL and to the age of “censoring” in the controls.

1476T

Identification of Gene-Gene Interactions For Alzheimer Disease Using Co-Operative Game Theory. *B. Vardarajan[1], G. Jun[1], R. Mayeux[2], J.L. Haines[3], M.A. Pericak-Vance[4], G. Schellenberg[5], K.L. Lunetta[2], L.A. Farrer[1], Alzheimer's Disease Genetics Consoritum.* 1) Boston University, Boston, MA; 2) Columbia University, New York, NY; 3) Vanderbilt University, Nashville, TN; 4) Hussman Institute of Human Genomics, Miami, FL; 5) University of Pennsylvania, Philadelphia, PA.

The multifactorial nature and large proportion of unexplained genetic variance of Alzheimer disease (AD) suggests that gene-gene interactions contribute to risk of the disorder. Contemporary approaches to detect interactions in genome-wide data are mathematically and computationally challenging. We investigated gene-gene interactions for AD using a novel algorithm based on co-operative game theory applied to 15 genome-wide association study (GWAS) datasets comprising of a total of 11,840 AD cases and 10,931 cognitively normal elderly controls from the Alzheimer Disease Genetics Consortium (ADGC). We adapted this approach, which was developed originally for solving multi-dimensional problems in economics and social sciences, to compute the rank-order values (called Shapely values) in order to identify genetic markers that contribute most to coalitions of SNPs in predicting AD risk. Treating each GWAS dataset as independent discovery, SNPs were ranked according to their contribution to coalitions formed with other SNPs. Using a backward elimination strategy, SNPs with low Shapley values were eliminated and the statistic was recalculated iteratively. We tested all two-way interactions between top Shapley SNPs in regression models which included the two SNPs (main effects) and a term for their interaction. Models yielding a p-value<0.05 for the interaction term were evaluated in each of the other datasets and the results from all datasets were combined by meta-analysis. Statistically significant interactions were observed with multiple marker combinations in the APOE region. Analyses also revealed statistically strong interactions between markers in 6 regions; CTNNA3-ATP11A (p=4.1E-07), CSMD1-PRKCQ (p=3.5E-08), DCC-UNC5CL (p=5.9e-8), CNTNAP2-RFC3 (p=1.16e-07), AACS-TSHZ3 (p= 2.64e-07) and CAMK4-MMD (p=3.3e-07). CTNNA3, CNTNAP2, CAMK4 and CSMD1 have been previously associated with neurological disorders in genetic or functional studies. The Game-Theory algorithm outperformed other methods in detecting a previously identified interaction between APOE and GAB2. Our approach also revealed strong interactions between SNPs in novel genes with weak main effects, which would have been overlooked if only SNPs with strong marginal association were tested. This method will be a valuable tool for identifying gene-gene interactions for other complex diseases and traits.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1477F

Overestimation of relatedness in admixed and ancestrally heterogeneous populations using method of moments estimation. *J. Morrison.* Biostatistics, Univ Washington, Seattle, WA.

Inference of identity by descent (IBD) in pairs of individuals is a common computational task undertaken in studies involving genome-wide SNP data. One of the most commonly used tools for inference of genome-wide allele sharing probabilities is the method of moments technique implemented in PLINK and other software which estimates the proportions of the genome at which two individuals share 0, 1, or 2 alleles IBD. These estimates can be used to check pedigrees, estimate heritability and adjust association analyses. In these last two applications, it is especially important that the estimates are accurate. The method of moments technique is computationally efficient and accurate when the sample size is large and ancestrally homogenous. However, in many cases, the assumption that samples are drawn from a single, homogeneous, randomly mating population does not hold. This assumption is violated if the study sample is small and there is no close population match in publically available datbases, if the study sample contains bi-ancestral pedigrees or if the study sample is drawn from an admixed population. I used publically available genome-wide SNP data to simulate pedigrees under these different conditions. I am able to demonstrate that in these cases method of moments estimation will lead to overestimation of relatedness between many pairs of individuals. Finally, I propose a simple method easily implemented without additional software for improving genome-wide IBD estimates when the assumption of a single, homogeneous population is violated.

1478W

Genome-wide patterns of identity-by-descent sharing in the French Canadian founder population. *H. Gauvin[1,2], C. Moreau[2], JF. Lefebvre[2], H. Vézina[3], D. Labuda[2,4], MH. Roy-Gagnon[1,2].* 1) Dept. of Social and Preventive Medicine, Université de Montréal, Montreal, Quebec, Canada; 2) Sainte-Justine Hospital Research Center, Université de Montréal, Montreal, Quebec, Canada; 3) Dept. of Human Sciences, Université du Québec à Chicoutimi, Chicoutimi, Quebec, Canada; 4) Dept. of Pediatrics, Université de Montréal, Montreal, Quebec, Canada.

In genetics the ability to accurately describe the familial relationships among a group of individuals can be advantageous and sometimes essential. Recent statistical tools succeed in assessing the degree of relatedness up to 6–7 generations with good power using dense genome-wide SNP data to estimate the extent of identity-by-descent (IBD) sharing. However, genome-wide patterns of IBD sharing have not been described for more remote and complex relatedness between individuals, such as that observed in a founder population like Quebec. In this study, we take advantage of the extended genealogical records of the French Canadian founder population to 1) compare different tools to identify regions of IBD, and 2) describe genome-wide IBD sharing and its correlation with genealogical characteristics. We first compared the performance of GERMLINE, Plink, and FastIBD among other tools in a sample of 140 individuals from seven regional populations of Quebec with both genotypic and genealogical data. We calculated genealogical measures such as the kinship and inbreeding coefficients, the number of lowest common ancestors and the minimum and average genetic distance to the lowest common ancestors. We used linear regression models to describe how well the genealogical characteristics are captured by genome-wide IBD sharing. Results showed that the extent of IBD sharing identified with FastIBD correlates best with relatedness measured using genealogical data. Total length of IBD sharing explained 85% of the genealogical kinship's variance. In addition, we observed significantly higher sharing in pairs of individuals with at least one inbred ancestor compared to those without any. Furthermore, patterns of IBD sharing and average sharing were different across regional populations, consistent with the settlement history of Quebec. Our results suggest that, as expected, the complex relatedness present in founder populations is reflected in patterns of IBD sharing. Using these patterns, it is thus possible to gain insight on the types of distant relationships in a sample from a founder population like Quebec.

1479T

Identity-by-descent-based heritability analysis in the Northern Finland Birth Cohort. *S.R. Browning[1], B.L. Browning[2].* 1) Dept Biostatistics, Univ Washington, Seattle, WA; 2) Div Medical Genetics, Univ Washington, Seattle, WA.

For most complex traits only a small proportion of heritability is explained by statistically significant associations from genome-wide association studies (GWAS). In order to determine how much heritability can potentially be explained through larger GWAS, several different approaches for estimating total narrow-sense heritability from GWAS data have recently been proposed. These methods include variance components with relatedness estimates from allele sharing [1], variance components with relatedness estimates from identity by descent (IBD) segments [2], and regression of phenotypic correlation on relatedness estimates from IBD segments [3]. These methods have not previously been compared on real or simulated data. We analyzed the narrow-sense heritability of nine metabolic traits in the Northern Finland Birth Cohort [4] using these methods. We found substantial estimated heritability for several traits, including LDL cholesterol (59% heritability), HDL cholestrol (46%) and fasting glucose levels (39%). When we used a variance components approach, IBD-based estimates of heritability were higher than allele-sharing based estimates for many of the traits, which is consistent with a significant role of rare variants in these traits because rare variants are generally not well tagged by SNPs but are captured by long IBD segments. Estimates of heritability from the regression-based approach are much lower than variance components estimates in these data, which may be explained by the presence of strong population structure in these data. Simulation results substantiate the downward bias of the regression-based approach in the presence of population structure.

References:
1. Yang et al. 2010. Nature Genetics 42: 565-9
2. Price et al. 2011. PLoS Genetics 7: e1001317
3. Zuk et al. 2012. PNAS 109: 1193-8
4. Sabatti et al. 2008. Nature Genetics 41: 35–46.

1480F

Efficient and Flexible IBD Mapping in Genome-wide Association Studies Based on Graph Representation of IBD information. *B. Han[1,2,3], E.Y. Kang[4], V. Ramensky[5], N. Freimer[5], J. Scharf[7], C. Mathews[8], D. Yu[7], A. Ruiz-Linares[9], G. Rouleau[10], N. Cox[11], D. Pauls[7], E. Eskin[4,6], TSAICG.* 1) Department of Genetics, Brigham & Women's Hospital, Harvard Medical School, Boston, MA, USA; 2) Partners HealthCare Center for Personalized Genetic Medicine, Boston, MA, USA; 3) Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA, USA; 4) Department of Computer Science, University of California, Los Angeles, CA, USA; 5) Center for Neurobehavioral Genetics, Semel Institute for Neuroscience and Human Behavior University of California, Los Angeles, California 90095, USA; 6) Department of Human Genetics, University of California Los Angeles, Los Angeles, CA, USA; 7) Psychiatric and Neurodevelopmental Genetics Unit, Massachusetts General Hospital, Boston, MA, USA; 8) Department of Psychiatry, University of California, San Francisco, CA, USA; 9) Department of Genetics, Evolution and Environment, University College London, UK; 10) Centre of Excellence in Neuromics, University of Montreal, Centre Hospitalier de l'Université de Montréal, Montreal, Quebec, Canada; 11) Department of Medicine, University of Chicago, Chicago, IL, USA.

Many methods have been proposed to detect rare variant associations, but they assume expensive sequence data. The identity-by-descent (IBD) mapping can be an inexpensive alternative for detecting regions including causal rare variants and can be an intermediate step suggesting target regions before sequencing. IBD mapping approaches detect associated regions by detecting an excessive haplotype sharing between cases. The widely known IBD mapping method is the pairwise method that compares the IBD rate of case/case pairs to control/control pairs or non-case/case pairs. However, the pairwise method has several limitations. First, the method is susceptible to population structure. Second, the method requires permutation and is inefficient taking years of CPU time for genome-wide data. Third, the analysis cannot include covariates. Fourth, associations to quantitative traits cannot be mapped. Here we propose a new IBD mapping method called GraphIBD that is efficient, general, and robust to population structure. The core idea of our approach is that the IBD information can be represented by a graph where the nodes are individuals and the edges between nodes are pairwise IBDs. Our idea is to split the edges and count the number of edges attached to each node, which we call virtual marker. Then we test association between virtual marker and phenotype. Testing based on virtual marker has a close relationship to the pairwise method that they are exactly equivalent in the balanced case/control study. Our simulations show that the new method has similar power to the traditional method but is thousands times faster taking only hours for genome-wide data. When applied to the isolated population data of Tourette syndrome, our method dramatically reduces the p-value inflation compared to the traditional method showing that our method is more robust to the population structure.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1481W

Optimization of sequencing studies in population isolates using shared ancestral haplotypes. *D. Glodzik[1], P. Navarro[1], V. Vitart[1], R. McQuillan[2], S. Wild[2], H. Campbell[2], A. Wright[1], J. Wilson[2], P. McKeigue[2].* 1) The Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, Midlothian, United Kingdom; 2) Center for Public Health Sciences, University of Edinburgh, Edinburgh, Midlothian, United Kingdom.

In an isolated population individuals are likely to share large genetic regions inherited from common ancestors. The phase and the underlying haplotype sharing can be inferred from SNP genotypes, which is useful in a number of applications, including identifying genetic variants influencing complex disease risk, and planning efficient cohort genotyping and sequencing strategies. We present a method for phasing in population isolates that is based on identity by descent. The purpose of our method is to detect haplotype sharing with maximum accuracy, at low computational effort. Implementation of the algorithm has been made available as the R library ANCHAP. We demonstrate the method using SNP genotypes from three isolated island populations, as well as in a cohort of unrelated individuals. In samples of around 1000 individuals from an isolated population, an average individual shares a haplotype at a genetic locus with 9–12 other individuals, compared to only 2 individuals within the non-isolated population. High genomic kinship in the isolates translates into high potential imputation gain. With sample sizes of ~1000 individuals from an isolated population genotyped using a dense SNP array, and with 20% of these individuals sequenced, 77% of sequences of the unsequenced subjects can be partially inferred. We describe an application of ANCHAP in optimization of resequencing studies.

1482T

Novel statistical model and computational methods for predicting HLA alleles from dense genotype data. *Y. Guan[1], Y. Shen[2].* 1) Pediatrics Molecular & Human Genetics, Baylor College of Medicine, Houston, TX; 2) Department of Biomedical Informatics, Columbia University, New York, NY.

Predicting human leukocyte antigen (HLA) alleles is an important problem in statistical genetics that remains an open challenge. The difficulty lies in the complicated linkage disequilibrium pattern in the major histocompatibility complex (MHC) region of the genome that contains HLA alleles. We present a novel statistical method for predicting HLA subtype and demonstrate its effectiveness via cross validation using real data from Wellcome trust case control consortium. The predicting method relies on a two layer clustering model that approximates the coalescent with recombination to quantify patterns of linkage disequilibrium. The model can jointly infer the haplotype sharing (identity by descent) between each pair of individuals, and the HLA allele type prediction is based on the degree of haplotype sharing between individuals. Since HLA alleles are plenty and many are rare, a large number of clusters is required to capture subtle differences between closely related HLA alleles. This makes the computation prohibitively difficult because the model fitting is quadratic in number of clusters for diploid individuals. We developed a novel computational method that is linear in number of clusters for diploid individuals, which is crucial for successfully predicting HLA alleles.

1483F

Haplotype clusters of rare variants in Korean genomes. *S. Hochreiter, G. Klambauer, G. Povysil, D.-A. Clevert.* Institute of Bioinformatics, Johannes Kepler University Linz, Linz, upper Austria, Austria.

A haplotype cluster corresponds to a DNA segment that is descended from a single founder and, therefore, some haplotypes are similar to each other in this segment. The advent of new sequencing technologies facilitates the identification of rare variants and therefore haplotype clusters containing rare variants ("rare haplotype clusters"). However, LD-related methods fail to extract rare haplotype clusters because of the high variance of LD measures for rare variants. IBD detection methods require sufficiently long IBD regions to avoid high false positive rates. We propose identifying rare haplotype clusters by HapFABIA which uses biclustering to combine LD information across individuals and IBD information along the chromosome. To identify rare haplotype clusters in the Korean population, we applied HapFABIA to data from the Korean Personal Genome Project (KPGP). Genotyping data from the KPGP was combined with those from the 1000-Genomes-Project leading to 1,131 individuals and 3.1 million single nucleotide variants (SNVs) on chromosome 1. HapFABIA identified 113,963 different rare haplotype clusters marked by tagSNVs that have a minor allele frequency of 5% or less. The rare haplotype clusters comprise 680,904 SNVs; that is 36.1% of the rare variants and 21.5% of all variants. The vast majority of 107,473 haplotype clusters contains Africans, while only 9,554 and 6,933 contain Europeans and Asians, respectively. We characterized haplotype clusters by matching with archaic genomes. Haplotype clusters that match the Denisova or the Neandertal genome are significantly enriched by Asians and Europeans. Interestingly, haplotype clusters matching the Denisova or the Neandertal genome contain, in some cases exclusively, Africans. Our findings indicate that the majority of rare haplotype clusters from chromosome 1 are based on ancient founder segments from times before humans migrated out of Africa. The enrichment of Koreans in Neandertal haplotype clusters (odds ratio 10.6 of Fisher's exact test) is not as high as for Han Chinese from Beijing, Han Chinese from South, and Japanese (odds ratios 23.9, 19.1, 22.7 of Fisher's exact test). In contrast to these results, the enrichment of Koreans in Denisova haplotype clusters (odds ratio 36.7 of Fisher's exact test) is is higher than for Han Chinese from Beijing, Han Chinese from South, and Japanese (odds ratios 7.6, 6.9, 7.0 of Fisher's exact test).

1484W

Evaluation of different phasing and imputation strategies in an isolated population with a mixture of unrelated individuals and extended pedigrees. *J. O'Connell[1], O. Delaneau[2], N. Pirastu[3], S. Ulivi[3], P. Gasparini[3], J. Marchini[1,2].* 1) The Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom; 2) Department of Statistics, University of Oxford, Oxford, United Kingdom; 3) IRCCS Burlo Garofolo Hospital, University of Trieste, Italy.

Genotype imputation (and associated haplotype phasing) has become a standard tool in genome wide association study pipelines and there are a number of popular implementations available. Linkage-disequilibrium (LD) based models assume individuals are distantly related. When this assumption does not hold, "long range phasing" techniques can exploit long contiguous segments of identical-by-descent (IBD) loci between individuals to achieve more accurate haplotype estimates (and hence better imputed genotypes). Additionally, genetic isolates may contain a mixture of unrelated individuals and small pedigrees. Pedigrees are highly informative of haplotype phase and standard software does not take advantage of this.

The INGI Friuli-Venezia Giulia cohort contains genotype data from 1476 individuals from six different isolated villages in northern Italy. The cohort consists of a mixture of unrelated (or distantly related) individuals and small pedigrees. We use this data to evaluate two interesting scenarios for phasing and imputation;

-a large group of unrelated individuals from an isolate
-a mixture of unrelated individuals and pedigrees

Popular IBD and LD based phasing methods (and downstream imputation) are compared in both these scenarios. In addition, we investigate modifications to the LD-based phasing algorithms to exploit large contiguous regions of IBD between individuals and approximate methods for handling extended pedigrees.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1485T

Local ancestry inference using identity by descent. *Z. CAI[1], B.L. Browning[2], S.R. Browning[1].* 1) Biostatistics, University of Washington, Seattle, WA; 2) Division of Medical Genetics, Department of Medicine, University of Washington, Seattle, WA.

Disease risk differs among groups that have diverse ancestries. Accurate inference of information regarding genetic ancestry will become a useful tool for improving the diagnosis and treatment in clinical studies. The 1000 Genomes Project provides a great opportunity for us to study the genetics of worldwide populations, including populations of admixed ancestry that are from America. We will present a new method that uses identity-by-descent information to infer local ancestry (ancestry of haplotypes at each point along the genome), with the goal of improving genetic mapping in data from admixed populations.

1486F

Detection of identity by descent based on rare variants. *G. Povysil, G. Klambauer, S. Hochreiter.* Institute of Bioinformatics, Johannes Kepler University Linz, Linz, Austria.

Identity by descent (IBD) between two individuals means that their alleles are identical because they were inherited from a shared common ancestor. Detection of IBD tracts is important for population genetics and association studies. IBD detection methods perform well for family studies where pedigrees are available and for common single nucleotide variants (SNVs). However, recent genotyping projects utilizing next generation sequencing comprise unrelated individuals and detect mostly rare variants. Currently, rare variants are of high interest in genetics because they are assumed to cause complex human diseases. However, their association with a disease is hard to detect as standard tests on rare variants yield low power. IBD mapping can be used to increase the power by two approaches. First, SNVs can be grouped based on IBD and subsequently their joint effect tested for disease association. Secondly, local genetic similarities between individuals can be measured by IBD and used for association tests like implemented in the sequence kernel association test (SKAT).

We compare two of the most commonly used IBD detection techniques, BEAGLE's fastIBD and PLINK, on simulated rare SNVs with implanted IBD tracts of different sizes and different proportions of minor alleles. Both methods miss a large proportion of short tracts and tracts that are tagged by few minor alleles. Overall fastIBD has higher power than PLINK, which is traded off against a higher false discovery rate. Besides correctly identifying IBD tracts, an exact estimation of IBD length and location is essential for identifying disease loci by IBD mapping. In this respect fastIBD systematically overestimates the length of IBD tracts while PLINK is very accurate. Our findings may help to choose an appropriate IBD detection method for studies involving rare variants. Summarizing, fastIBD finds almost all IBD tracts on the cost of many false positives whereas PLINK finds a smaller number of IBD tracts of which only a few are false positives.

1487W

Refined IBD: a new method for detecting identity by descent in population samples. *B.L. Browning[1], S.R. Browning[2].* 1) Medicine, Med Genetics, University of Washington, Seattle, WA; 2) Biostatistics, University of Washington, Seattle, WA.

Identity-by-descent (IBD) from a recent common ancestor is ubiquitous in large population samples. IBD manifests as a long haplotype shared by two or more individuals, and IBD can be detected with high power when the recent common ancestor is within the past 25 generations. The power to detect IBD can be increased by modeling haplotype phase uncertainty, but this modeling has been computationally expensive.

We have developed a new method for IBD detection called Refined IBD that solves the computational problem of incorporating haplotype phase uncertainty. First, candidate IBD tracts are identified by searching for shared haplotypes that exceed a length threshold. Candidate tracts are efficiently identified using IBD detection methods such as GERMLINE or fastIBD that do not model haplotype phase uncertainty. We then compare the likelihood of the genotypes in each candidate tract under IBD and non-IBD models using a likelihood ratio test. This test incorporates haplotype phase uncertainty and identifies the subset of candidate tracts with the greatest evidence for IBD.

We have applied Refined IBD to the WTCCC2 control data (n=5200). After tuning methods to achieve equal false positive rates, Refined IBD detected fourfold more IBD than the GERMLINE and fastIBD algorithms.

Refined IBD is implemented in BEAGLE version 4 (http://faculty.washington.edu/browning/beagle/beagle.html).

1488T

IBD Estimation With Whole-genome Sequence Data. *L. Han, M. Abney.* Human genetics Dept, University of Chicago, Chicago, IL.

Accurate knowledge of identity by descent probabilities (IBDp) between a pair of related individuals at a locus enables sequence imputation methods, improves phasing and detects deletion variants. Next-generation sequencing of personal genomes has revealed numerous rare variants, many of which may play an important role in the risk for common diseases. However, incorporating rare variants while ignoring linkage disequilibrium (LD) affects the accuracy of IBDp estimates. Here we present a method that can estimate IBDp at each locus using whole-genome sequence data. The method incorporates rare variants and extends the standard hidden Markov model to include LD by conditioning on multiple SNPs. At the same time, genotyping errors and missing data are considered in this method. We validate it through simulations and apply it to a real whole genome sequence data set.

1489F

Evaluating type 1 error and power for association and linkage in large complex pedigrees. *A.C. Cummings[1], E. Torstenson[1], M.F. Davis[1], L.N. D'Aoust[1], W.K. Scott[2], M.A. Pericak-Vance[2], W.S. Bush[1], J.L. Haines[1].* 1) Center for Human Genetic Research, Vanderbilt University, Nashville, TN; 2) Hussman Institute of Human Genomics, University of Miami, Miami, FL.

Studying population isolates, like the Amish, with large, complex pedigrees has many advantages for discovering genetic susceptibility loci; however, analyses can be quite complicated. Association tests need to be corrected for relatedness, and linkage analysis requires subdividing and simplifying the pedigree structure. Using simulated pedigrees with the same structure as an Amish pedigree (4998 total, 796 genotyped, 106 affected) we evaluated type 1 error and power for association using the Quasi-Likelihood Score (MQLS) test and for 2-point and multipoint linkage analyses using PedCut for pedigree splitting and Merlin for linkage analyses. To assess type 1 error we simulated null SNP data, and to assess power we simulated dominant, recessive, and additive inheritance models across three odds ratios (1.1, 1.5, and 2.0) for a total of 9 disease models. 1000 replicates were performed for each analysis. We see essentially no inflated type 1 error rate (5.06% for p-values ≤ 0.05; 0.13% for p-values ≤ 0.001) when using MQLS. We also see very low type 1 error rates of 0.01% and 0.02% under dominant and recessive models using a 2-pt heterogeneity lod score (HLOD) threshold of 3. Under both dominant and recessive models, we see type 1 error rates of 2.50% when using multipoint analysis and an HLOD ≥ 3. Power to detect $p \leq 0.05$ using MQLS was >50% for the simulated additive and dominant models with OR's 2.0 (96% add; 91% dom) and 1.5 (66.6% add; 50.2% dom). For the recessive model with OR=2.0, we calculated power to be 26% for $p \leq .05$. Power for all models was less than 50% for $p \leq 5 \times 10\text{-}5$. Power to detect linkage (HLOD>=3) using 2-point and multipoint parametric analysis is very low with these small effect sizes. The highest power to detect a 2-pt and MPT HLOD ≥ 1.0 was 12.3% and 30.7% respectively for the simulated additive model with OR=2.0 when linkage was calculated assuming a dominant model (10.4% assuming a recessive model). These results demonstrate that type 1 error is not inflated using this pedigree structure for either linkage or association analyses. For small effect sizes, characteristic of common complex diseases, there is little to no power to detect linkage using classical criteria, suggesting that association methods are more powerful for detecting variants of modest effect even in isolated populations such as the Amish. More disease models need to be evaluated to make a full assessment of power using these analysis protocols.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1490W

Loss-of-co-Homozygosity (LOcH) mapping: a novel non-parametric linkage analysis leveraging exome sequencing data. *Y. Okada[1,2], N. Gupta[2], D. Mirel[2], S. Gabriel[2], T. Arayssi[3], F. Mouassess[4], WAL. Achkar[4], LA. Kazkaz[5,6], RM. Plenge[1,2].* 1) Division of Rheumatology, Immunology, and Allergy, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA; 2) Program in Medical and Population Genetics, Broad Institute, Cambridge, MA, USA; 3) Weill Cornell Medical College-Qatar, Education City, Doha, Qatar; 4) Molecular Biology and Biotechnology Dept, Human Genetics Division, Damascus, Syria; 5) Tishreen Hospital, Damascus, Syria; 6) Syrian Association for Rheumatology, Damascus, Syria.

Mapping the ancestry-shared genetic regions from pedigrees with genome-wide sequence data represents an effective method to identify causal rare mutations in Mendelian disorders. Although classical linkage analysis is widely used to detect regions harboring causal mutations, its usage is restricted to diseases with typical modes of genetic inheritance. Here, we develop a novel non-parametric linkage analysis we term “Loss-of-co-Homozygosity” (LOcH) mapping that extends homozygosity mapping to include any type of inheritance mode. The fundamental principal underlying LOcH is that a disease-causing mutation in a pedigree resides on a single ancestral haplotype, and that affected cases will carry at least one copy of the mutation - but are never homozygous for the non-mutated allele. That is, genetic markers adjacent to the causal mutation lose homozygous genotypes for at least one of the alleles. LOcH mapping uses genome-wide SNP data to search the regional stretches that lose one or both homozygous genotypes (i.e., lose “co-homozygosity”) in affected cases, which serve as candidate regions harboring causal rare mutations. Within the detected LOcH stretches, our method enables genotype imputation of the mutations identified by exome sequencing, by evaluating the presence of the LOcH stretches in other non-exome-sequenced subjects. To demonstrate our approach, we applied LOcH mapping to a 4-generation, 48-person consanguineous Syrian pedigree in which 8 individuals were affected with rheumatoid arthritis (RA). We used GWAS data and LOcH mapping on a subset of 4 affected and 12 unaffected family members to identify ~12% of the genome in which the same ancestral haplotype was shared among all RA cases. We sequenced the exomes of these 4 RA cases, and identified 13 nonsense or missense mutations shared among all cases but not observed in dbSNP, the 1000 Genomes Project, or the Exome Sequencing Project. All 13 mutations reside within the 12% of the genome defined as LOcH using GWAS data ($P < 10^{-11}$), suggesting that at least one might be the causal mutation. We imputed these 13 candidate mutations in all unaffected family members. We found that 1 mutation in the *phospholipase B1* (*PLB1*) gene at 2p23 preferentially segregated in cases compared to controls ($P < 0.05$). While additional investigation of this mutation is required to confirm the association with risk of RA, the approach highlights a novel method of statistical analysis of genome-wide sequence data.

1491T

Patterns of Indel Variation: Comparison of calling methods and LD with SNPs. *K. Huang[1], M. Ehm[2], N. Bing[2], Y. Liu[3], J. Xu[4], A. Slater[2], D. Fraser[2], J. Novembre[5], J. Li[4], M. Neison[2], Y. Li[1,3,6].* 1) Biostatistics, University of North Carolina - Chapel Hill, Chapel Hill, NC, USA; 2) Quantitative Sciences, GlaxoSmithKline, RTP, NC, USA; Upper Merion, PA, USA; and Stevenage, UK; 3) Computer Science, University of North Carolina - Chapel Hill, Chapel Hill, NC, USA; 4) Human Genetics, University of Michigan, Ann Arbor, MI, USA; 5) Ecology and Evolutionary Biology, University of California - Los Angeles, Los Angeles, CA, USA; 6) Genetics, University of North Carolina - Chapel Hill, Chapel Hill, NC, USA.

Insertion and deletion variants (indels) are an important but poorly studied type of genetic variation. The advancement of massively parallel sequencing technologies and improvement of indel calling methods enables us to systemically measure indels, evaluate their variation and linkage disequilibrium (LD) patterns within and across populations, and assess their impact on human traits. We have investigated the quality and patterns of indel variations within the exons of 202 drug target genes sequenced in over 14,000 individuals. Indels were called using a pipeline involving Dindel, pindel, samtools and GATK. In addition to standard filters from the indel calling algorithms used, we have implemented additional quality control (QC) criterion based on duplicate concordance analysis. Our QC, aimed to reach 96% heterozygous concordance, was confirmed to reach the goal by genotypes determined experimentally through actual genotyping and Sanger sequencing. As expected, a considerable proportion (45.6%) of common (minor allele frequency [MAF] > 0.5%) indels can be tagged by SNPs: 45.6% by GWAS (Affymetrix 6.0) SNPs alone and 58.3% by the combination of GWAS and sequencing identified SNPs. Coding indels are predominantly rare (95.5% with MAF <0.5%), and only a small proportion of coding indels can be tagged by SNPs. Missing genotype calls have relatively little impact on common indel LD inference but cause severe underestimation of LD with rare coding indels. Frequency estimation could be distorted by differential missingness. We have found in our study that excluding indels with call rate < 70% leads to accurate MAF estimation (Pearson correlation > 0.992 when compared with MAF estimated using experimentally determined genotypes with negligible missingness). Our study estimates the accuracy of indel calls with current algorithms across a large sample of sequenced exons, presents the previously undescribed LD pattern for rare indels, and provides insight into the impact of missing genotype calls on LD inference and frequency estimation.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1492F

Estimating disease associations with common, low-risk alleles using pedigree data: Applications to breast cancer. *D. R. Barnes[1], D. Barrowdale[1], J. Hopper[2], D. Goldgar[3], G. Chenevix-Trench[4], A. C. Antoniou[1], G. Mitchell[5], kConFab Investigators, AOCS Group.* 1) Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, United Kingdom; 2) Centre for Molecular, Environmental, Genetic and Analytic Epidemiology, Melbourne School of Population Health, University of Melbourne, Australia; 3) Department of Dermatology, University of Utah School of Medicine, Salt Lake City, Utah, USA; 4) Genetics and Population Health Division, Queensland Institute of Medical Research, Brisbane, Australia; 5) Peter MacCallum Cancer Centre, Melbourne, Australia.

Pedigrees in which several family members are genotyped may be a valuable resource for investigating genetic associations with disease risk. We developed an analytical framework to characterise SNP associations with breast cancer (BC) risk using family data and applied it to evaluate the utility of SNP profiles for risk prediction in the context of families with multiple affected members. We used data from 736 families ascertained due to having multiple cases of BC and in which several (on average 8) family members had been genotyped for 25 SNPs previously associated with BC and from 897 unrelated unaffected controls. Disease incidence was assumed to be a function of age (t) $\lambda(t)=\lambda_0(t)\exp(\beta g+P)$, where g represents the SNP genotype, β is the log-relative risk (RR) and $P\sim N(0,\sigma_R^2)$ is a polygenic component. The sum of the residual polygenic variance (PV), σ_R^2, and variance due to the locus of interest was constrained to agree with external estimates of the total PV for BC. Parameters were estimated by maximising the retrospective likelihood (RL) of observed family genotypes conditional on family disease phenotypes. Models were extended to assess parent-of-origin effects (POE), testing risk differences between maternally and paternally inherited alleles. For comparison purposes, data were analysed by case-control analysis. From the pedigree-RL analyses, 13 SNPs were nominally significant. These analyses generally gave SNP association estimates consistent with those from population-based studies. Logistic regression models comparing affected family members against unrelated unaffected controls and ignoring pedigree structure generally gave larger RRs and p-values. For example, for *FGFR2* SNP rs2981582 the estimated RR was 1.26, 95% CI:1.17–1.36, $p=9\times10^{-10}$ under RL compared with OR= 1.44, 95% CI:1.26–1.65, $p=2\times10^{-7}$ under logistic regression. *LSP1* SNP rs3817198, previously shown to exhibit POE with BC risk yielded maternal and paternal RR estimates similar to those previously reported ($RR_{paternal}$= 1.12 (0.99–1.27), p=0.08, p-one sided=0.04; $RR_{maternal}$=0.94 (0.84–1.06), p=0.3). POE were not observed for any other SNP. We estimated AUC= 0.63 (95% CI:0.60–0.66) for discriminating between affected women with BC family history and unaffected, based on combined SNP profiles alone. This was higher than previous estimates based on random case-control studies, suggesting SNPs are more informative for discrimination in the context of multiple-case families.

1493W

Improving association tests by learning mode of inheritance from parental data. *Z. Yu[1], D. Gillen[1], C. Li[2], M. Demetriou[2,3].* 1) Dept Statistics, Univ California, Irvine, CA; 2) Dept Neurology and Institute for Immunology, Univ California, Irvine, CA; 3) Dept Microbiology and Molecular Genetics, Univ of California, CA.

Information regarding the true underlying genetic model, i.e., the mode of inheritance, is crucial in detecting association. For family-based association studies, we show that the underlying genetic model can be learned from parental data. Specifically, for parental mating type data, we propose a novel statistic to test whether the underlying true genetic model is additive, dominant, or recessive; for parental genotype-phenotype data, we propose several strategies to learn the underlying true genetic model. We then illustrate how to incorporate the learned information into family-based association tests. Because family-based association tests are conducted conditional on parental genotypes, their type I error rates are not inflated by the information learned from parental data, even if such information is weak or learned when the assumption of Hardy-Weinberg equilibrium is violated. Our simulations demonstrate that incorporating parental data into family-based association tests improves the power of association tests. The application of our proposed methods to a candidate-gene study of type 1 diabetes successfully detects a recessive effect that would otherwise be missed using the conventional family-based association tests.

1494T

Enhanced maternal origin of the 22q11.2 deletion in velo-cardio-facial/DiGeorge syndrome. *M. Delio[1], T. Guo[1], D. Mcdonald-McGinn[2], E. Zackai[2,3], S. Herman[1], A. Higgins[4], K. Coleman[5], T. Wang[6], A. Auton[1], R. Shprintzen[4], B. Emanuael[2,3], B. Morrow[1].* 1) Albert Einstein College of Medicine, Bronx, NY; 2) Division of Human Genetics, Children's Hospital of Philadelphia and Department of Pediatrics, University of Pennsylvania School of Medicine, Philadelphia, PA, USA 19104; 3) Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA 19104; 4) Syracuse, NY, USA 13210; 5) Children's Healthcare of Atlanta and Emory University School of Medicine, Atlanta, GA, USA 30322; 6) Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY, USA 10461.

Velo-cardio-facial syndrome/DiGeorge/22q11.2 deletion syndrome (VCFS/DGS/22q11DS) is a congenital malformation disorder associated with a de novo 3 Mb hemizygous 22q11.2 deletion, occurring in 1/4,000 live births. The deletion is mediated by meiotic non-allelic homologous recombination (NAHR) events between flanking low copy repeats also referred to as segmental duplications. One important question has been whether there is a parent of origin biases for mediating the deletion. The mechanism and timing of maternal and paternal meiosis differ, suggesting that there may be gender-based genetic factors that alter risk for the deletion. Risk factors might not reveal themselves in studies of individual nuclear families or smaller cohorts but can be identified by studies of a large number of simplex families. Previous studies exist but they were small in size, with conflicting results. A total of 451 DNAs from probands with 22q11DS and their normal parents were genotyped to determine whether there is a parent of origin bias to the de novo 22q11.2 deletion. We found that 219 (56%) patients with 22q11DS had maternal origin and 177 (45%) had paternal origin of the deletion, a ratio of 1.24, resulting in a statistically significant increase of maternal origin (p = 0.0172). A meta-analysis including previous studies resulted in 355 deletions that were of (56%) of maternal origin and 275 (44%) of paternal origin, a ratio of 1.29, resulting in a statistically significant maternal bias in origin of the deletion, based on Fishers's combined probability test (n= 630; p = 0.0016). The ratio is similar to the female to male genetic distance ratio, suggesting that increased maternal origin may be due to increased meiotic recombination rate. We are currently analyzing the potential association of advanced parental age at the time of conception, which could provide insights into the mechanism of the deletion and offer enhanced genetic counseling for 22q11.2DS. We are in the process of analyzing the conception age for over 700 families.

1495F

Connection between heritability and polygenic risk score: Implications for the genetic architecture of complex traits. *V.V. TRUBETSKOY[1], E.R. GAMAZON[1], M.E. DOLAN[2], N.J. COX[1], H.K. IM[3].* 1) Section of Genetic Medicine, Department of Medicine, The University of Chicago; 2) Section of Hematology and Oncology, Department of Medicine, The University of Chicago; 3) Department of Health Studies, The University of Chicago.

Heritability is a fundamental parameter that may yield insights into the genetic architecture of complex traits. Genome-wide association studies (GWAS) have, to date, identified thousands of loci highly significantly and reproducibly associated with a broad spectrum of complex human traits, but these established loci account for only a small proportion of the heritability for any given disease. A recent analytic development, based on linear mixed models, has enabled quantification of the proportion of phenotypic variance attributable to SNPs contributing to susceptibility that are in linkage disequilibrium with SNPs interrogated in GWAS. Indeed, studies based on this approach have shown that a substantial portion of the missing heritability to at least some complex traits may be explained by a large number of common SNPs of small effect size. A second analytic development, namely polygenic modeling, has established a polygenic component in variation to certain traits, such as schizophrenia, height and body mass index. A polygenic score that is calculated from the risk alleles and the effect sizes of a select set of SNPs in a first GWAS is tested in a second GWAS for its association with the trait. In this study, we establish the connection between a polygenic score and heritability. From this connection, we propose a novel method, PolyScore, for calculating the proportion of phenotypic variation due to additive genetic risk factors. In addition to the insights PolyScore provides into recent analytic approaches on the polygenic component to complex traits, the method provides a computationally efficient approach to calculating heritability. As an example, we applied our method to the phenotype of intrinsic growth rate measured in HapMap cell lines. Despite the small sample size (n=108), we were able to get an estimate of heritability of 10%, which we verified is significantly different from 0 by using permutation. The linear mixed model approach yielded an estimate of heritability consistent between 0 and 1, suggesting that our method may be more robust for small sample sizes. Our method can be extended to handle rare variations, which would allow us to explore the contribution of rare and common variants to a cellular phenotype of fundamental significance to pathophysiology and complex traits.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1496W

Genetic variants in DNA repair pathway genes and upper aerodigestive tract cancers: Combined analysis of data from genome-wide association studies. *M.-C. Babron[1,2], R. Kazma[3], V. Gaborieau[4], J.D. McKay[4], P. Brennan[4], A. Sarasin[5,6,7], S. Benhamou[1,2], INHANCE consortium.* 1) Inserm, U946, Paris, France; 2) Univ Paris Diderot, Institut Universitaire d'Hématologie, UMRS946, Paris, France; 3) Department of Epidemiology and Biostatistics and Institute for Human Genetics, University of California San Francisco, San Francisco, CA, USA; 4) International Agency for Research on Cancer, Lyon, France; 5) University Paris-Sud, Villejuif, France; 6) CNRS UMR8200, Villejuif, France; 7) Institut de Cancérologie Gustave Roussy, Villejuif, France.

Upper Aerodigestive tract (UADT) cancers are a related group of cancers located in the oral cavity, pharynx, larynx and esophagus. While tobacco smoking and alcohol consumption are the prime environmental risk factors, genetic factors and several biological pathways are also thought to contribute to UADT cancer risk. Among these pathways, those involved in DNA repair are good candidates because of their critical role in maintaining genome integrity. Our aim is to assess the role of single nucleotide polymorphisms (SNPs) in genes involved in DNA repair pathways on UADT cancer risk, and their interactions with tobacco smoking and alcohol consumption. We pooled individual data from two genome-wide association studies (GWAS) covering Europe comprising 1,981 cancer cases and 3,151 controls. Demographic information (age, sex, and country of origin) as well as lifetime tobacco smoking and alcohol consumption were available for cases and controls. Cancer site of cases was also documented. We investigated 1,760 SNPs present of the Illumina HumanMap300 chip located in 222 genes involved in 13 DNA repair pathways. First, we tested for disease association with SNPs using unconditional logistic regression adjusting for age, sex, tobacco smoking, and alcohol consumption. To account for population stratification, we included in the regression the eigenvalues of the top 4 axes of the Principal Component Analysis performed on the GWAS data. To account for multiple testing, we used the Bonferroni correction for the effective number of tests, obtained from the Linkage Disequilibrium matrix of the 1,760 SNPs (Li et al, Heredity 2005;95:221–227). Second, we tested for disease association with SNPs grouped by DNA repair genes and pathways using the SNP-set Kernel Association Test (Wu et al, Am J Hum Genet 2010;86:929–942).This method is specifically designed to investigate the disease association of a given set of SNPs. Preliminary results suggest that five distinct DNA repair pathways are involved in UADT cancer risk: telomerases (TEL), DNA polymerases (POL), homologous recombination (HR), base excision repair (BER), and chromatin structure (CHS). Interestingly, three of these pathways (POL, CHS, and HR) are involved in the susceptibility to lung cancer, a cancer in which tobacco smoking also plays a central role (Kazma et al, Carcinogenesis 2012;33:1059–1064).

1497T

Venous Thromboembolism (VTE)-susceptibility Pathways by Gene Set Analyses. *J. Heit[1], S.M. Armasu[2], J.P. Sinnwell[2], D.J. Schaid[2], M. de Andrade[2].* 1) Internal Med/Hematology Res, Mayo Clinic, Rochester, MN; 2) Div Biomedical Statistics and Informatics, Mayo Clinic, Rochester, MN.

Background: Gene set analysis is a promising approach for complementing and interpreting genome-wide association studies. **Objective:** To identify groups of functionally related genes associated with VTE. **Methods:** Genome-wide scan (Illumina 660W Quad [561,408 SNPs]) genotypes from 1270 white adults with objectively-diagnosed VTE (cases; no cancer, venous catheter or antiphospholipid antibodies) and 1302 controls (frequency-matched on case age, gender, race, stroke/myocardial infarction status) were mapped to genes, and genes to the Gene Ontology (GO) and KEGG structures. Gene sets analyses were performed using two methods, a score-statistic method where positive effects of genes on the trait do not cancel, and a p-value combination approach. The analyses were adjusted for age, sex, USA state of residence and stroke/MI status. **Results:** After controlling for multiple comparisons, the following three gene sets were significantly associated with VTE: (1) Positive Regulation of Fibrinolysis (ID 0051919; *KLKB1*, *F11*, *F12*, *PLG*) within the GO Biological Process structures by both analysis methods, either adjusting for covariates only, or for covariates and *F5* rs6025 (Factor V Leiden); relatively strong single SNP association in *F11* may drive this result; (2) Glycoprotein-fucosylgalactoside Alpha-N-acetylgalactosaminyltransferase Activity [ID 0004380; *ABO*] within the GO Molecular Function structures; and (3) Granular Component (ID 0001652; *CDKN2AIP*, *SURF6*, *FBL*) within the GO Cellular Component (nucleolus) structures. None of the gene set structures within KEGG achieved statistical significance. **Conclusion:** Fibrinolysis, glycosyltransferase and nucleolus granular component gene sets are associated with VTE.

1498F

Genetic pathways for ADHD show association to hyperactive/impulsive symptoms. *A. Arias Vásquez[1], J. Bralten[3,7], B. Franke[2], I. Waldan[5], S. Faraone[6], J. Buitelaar[3,4].* 1) Departments of Psychiatry, Human Genetics & Cognitive Neuroscience, Donders Institute for Brain, Cognition and Behavior, Radboud University Nijmegen, Nijmegen, Gelderland, Netherlands; 2) Departments of Psychiatry & Human Genetics, Donders Institute for Brain, Cognition and Behavior Radboud University Nijmegen, Nijmegen, Gelderland, Netherlands; 3) Department Cognitive Neuroscience, Donders Institute for Brain, Cognition and Behavior, Radboud University Nijmegen, Nijmegen, Gelderland, Netherlands; 4) Karakter Child and Adolescent Psychiatry University Center, Nijmegen, The Netherlands; 5) Department of Psychology, Emory University, Atlanta, GA, USA; 6) Department of Psychiatry and Neuroscience and Physiology, SUNY Upstate Medical University, Syracuse, NY, USA; 7) Department of Human Genetics, Radboud University Nijmegen Medical Centre, The Netherlands.

Attention-Deficit Hyperactivity Disorder (ADHD) is a highly heritable disorder but finding genes involved in its etiology has been challenging. Compared to single-SNP or haplotype analysis, pathway based analysis might increase the explained phenotypic variance, thereby boosting the power of genetic studies. Alterations in dopamine, noradrenalin and serotonin neurotransmission have been hypothesized to play a role in ADHD, being associated to behavioral aspects of the disorder. Genes involved in neurite outgrowth that appeared when analyzing the top results of the performed GWAS of ADHD are also of interest. Therefore, we decided to investigate the association of these pathways with ADHD symptom count. Data from the International Multicentre ADHD Genetics (IMAGE) study was used. Phenotypic data were the DSM-IV symptom counts for inattention and hyperactivity/impulsivity. Selection of the genetic pathways was based on Ingenuity Pathway Analysis software in combination with literature. From the imputed (HM2) and pruned (r^2=0.8) genome-wide data 6501 SNPs were selected for the association analysis. The analysis consisted of SNP-by-SNP linear regression including age and gender as covariates and the summation of the individual SNP model effect of the complete set of SNPs. Permutations were ran investigating whether the observed pathway-based statistic fell into the extreme 5% of the distribution and empirical p-values were obtained. Post-hoc analyses were conducted for the Connors' Parent Rating (CPR) Scale, for the separate pathways, to investigate single pathway contribution and for neuropsychological data. Analysis of the three pathways showed significant association to ADHD hyperactive/impulsive symptoms (p=0.0081) but not to inattentive symptoms (p=0.73). Investigating (CPR) hyperactive/impulsive score also showed a significant association (p=0.0049). Post-hoc analyses showed contribution of all pathways (dopamine/noradrenalin p=0.0004, serotonin p= 0.0149, neurite outgrowth p=0.0452) to the hyperactive/impulsive score. The analysis of a neuropsychological measure of inhibition (n=187), also showed association to the three pathways (p=0.0045). These findings show that pathway-based association analyses may overcome power problems in association testing by taking into account allelic heterogeneity and our results suggest a specific genetic contribution to the hyperactive/impulsive component of ADHD.

1499W

Graphical Models and Intervention Calculus for Causal Inference of Genetic Epidemiology Studies. *P. Wang[1], L. Jin[1], M. Xiong[1,2].* 1) MOE Key Laboratory of Contemporary Anthropology, School of Life Sciences, Fudan University, Shanghai, Shanghai, China; 2) Human Genetics Center, University of Texas School of Public Health, Houston, TX, USA.

Complex diseases are caused by genetic variants and environments that are organized into complicated networks. To evaluate their causal relationships among multiple phenotypes and genetic and non-genetic factors plays an essential role in risk prediction and treatment of complex diseases. In the past several years, graphical models coupled with intervention calculus are developed for discovering causal relations among variables and estimate the causal effects. However, such methods can only been applied to thousands of variables. They are difficult to be applied to millions of variables which will be produced by next-generation sequencing. Identifying causal relationships among millions of genetic and non-genetic variables raises great challenges. In this report, we first use high dimensional data reduction techniques to compress genetic information in a gene or genomic region into a variable that follows a Gaussian distribution. We use directed acyclic graph (DAG) model to represent causal relations generating observational data. We then use L1 regularization to estimate the equivalence class of a DAG. Finally, intervention calculus is used to discover causal relations among variables and estimate their causal effects. In summary, we propose to combine high dimension data reduction techniques, graphical models and intervention calculus together to develop a causal inference tool for causal inference of genetic epidemiology studies of complex diseases with next-generation sequencing data. The developed causal inference is applied to the NHLBI's Exome Sequencing Project (ESP) dataset. The preliminary results are encouraging. They demonstrate that the developed causal inference tools will have important implication in risk prediction and treatment of complex diseases.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1500T

Multi-SNP analysis of GWAS reveals important gene-gene interactions and gene networks implicated in breast cancer. *L. Briollais[1], A. Dobra[2], H. Massam[3], H. Ozcelik[1].* 1) Samuel Lunenfeld research Institute, Mouint Sinai Hospital, Toronto, Ontario, Canada; 2) Dept. of Mathematics and Statistics, University of Washington, USA; 3) Dept. of Mathematics and Statistics, York University. ON, Canada.

The emergence of new high-throughput technologies for SNP genotyping has generated great promises that the genetic basis of many common human diseases could be elucidated. These expectations have been partially fulfilled and indeed, genome-wide association studies (GWAS) have identified hundreds of genetic variants associated with common human diseases and complex traits. GWAS have also given rise to many criticisms and increasing skepticism because of its relative difficulty to identify important functional variants and genetic variants that have good prediction ability. The actual paradigm to analyze GWAS is to perform an exhausting testing of all single SNP associations with the response variable. However, there is no theoretical justification of the implicit claim that the resulting set of genetic variants can be converted into a good predictive model and that those variants have any biological relevance. Indeed, many GWAS findings have no clear biological interpretation and fall into inter-genic regions. As a shift to the usual paradigm, we propose a novel statistical method for performing multi-SNP analysis in the context of GWAS. Our working hypothesis is that SNPs interacting together might be more functionally relevant and have a better predictive ability. Our statistical method is based on Bayesian graphical model, which offers a comprehensive framework for studying conditional dependence patterns between discrete random variables. The Bayesian framework allows its application to high-dimensional GWAS data by use of efficient stochastic search algorithms. The multi-SNP models with the highest posterior probability are then selected. Our application to the CGEM breast cancer GWAS data illustrates the advantages of our approach. We were able to detect important SNP-SNP interactions associated with breast cancer, involving important genes previously found implicated in breast cancer (FGFR2, APC, BUB3, NTSR1, etc.). Some of these interactions were validated experimentally and correspond to proteins that interact physically. They could involve both rare and common genetic variants. Importantly, our analysis also suggests important gene networks associated with breast cancer through our SNP discoveries. The predictive ability of our multi-SNP models was superior to previous models with an AUC of 63.5%. We will also discuss how our approach can be further developed for a more systemic investigation of gene networks generated from SNP data.

1501F

Connecting genome-wide association studies with a biological gene interaction network to uncover epistasis. *D.S. Himmelstein[1,2,3], J.S. Witte[1,3,4], O.G. Troyanskaya[2,5], C.S. Greene[2].* 1) Institute for Human Genetics, UCSF, San Francisco, CA; 2) Lewis-Sigler Institute, Princeton University, Princeton, NJ; 3) Biological and Medical Informatics, UCSF, San Francisco, CA; 4) Epidemiology and Biostatistics, UCSF, San Francisco, CA; 5) Computer Science, Princeton University, Princeton, NJ.

Genome-wide association studies (GWAS) have implicated over 1100 genetic loci in conferring susceptibility to 165 complex traits in humans. However, these loci explain only a small portion of the estimated heritability for most complex diseases. Missing heritability may result from the implicit assumption of standard statistical approaches that common allelic variants combine independently to influence phenotype. Such approaches fail to consider potential non-additive relationships between loci that contribute to disease (epistasis). Inappropriately ignoring epistasis may hinder the search for risk loci and lead to overestimation of total disease heritability. Here, we present a method to assess the role of epistasis in a disease's etiology. Our method measures the extent of genetic signal in a GWAS by comparing statistical relationships between gene-pairs and disease to a gold standard of known biological gene-gene relationships. We assess whether the concordance between biological and genetic relationships increases when disease models incorporate the possibility of non-linear interactions in addition to additive gene-disease associations. An increase in concordance indicates epistasis in a disease's etiology. Appropriate permutation of case-control status ensures the findings result from real genetic signals rather than spurious dataset biases. By assuming disease susceptibility genes function within genetic pathways captured by the biological relationship network, our approach uses our understanding of biological interactions to evaluate statistical epistasis. With our method we evaluate the role that epistasis plays in common human diseases. We apply our method to studies of psoriasis, breast cancer and prostate cancer. By using ROC curves to quantify the ability of genetic relationships to predict biological relationship status and by comparing the disease association distributions between biologically related and unrelated gene-pairs, we observe modest but discernible signatures of epistasis in all three GWAS. When contemplating whether complex human disease models should account for genetic interactions, researchers are faced with a paucity of replicated examples of statistical epistasis to guide their analyses. Our method provides a disease specific indication of whether epistasis plays an appreciable role in disease etiology.

1502W

Gene-based collapsing methods may not be powerful for pathway-based association analysis for exome sequencing data. *G. Wu, D. Zhi.* Biostatistics, University of Alabama at Birmingham, Birmingham, AL.

Exome sequencing emerges as a powerful approach for the detection of coding variants, enabling genetic association studies for both common and rare variants. Most current rare variant association methods are gene-based and have insufficient power even for moderately sized genome-wide sequencing data. Pathway association analysis tests the enrichment of phenotype-associating variants and genes in biological pathways, even when few of them reach genome-wide significance individually. Therefore, pathway association may offer insight for most exome sequencing data that were insufficiently powered for gene-base tests. Current pathway association analysis methods, designed for genome-wide association studies (GWAS) data, may not be appropriate for sequencing data. In particular, there may be multiple coding variants which have different directions and magnitude of effects in a gene. Existing pathway association methods typically collapse information at gene-level, and conduct a gene set enrichment test. This approach may lose power when multiple variants in a gene carry association signals. In this study, we evaluate the applicability of GWAS-based pathway association methods to exome sequencing data. In particular, we propose to directly include all variants' information onto the pathway level enrichment analysis for exome sequencing data. Based on extensive simulations using exome sequencing genotype information in 1000 Genome project, and some realistic assumptions on underlying genetic model, we show that, contrary to current beliefs, pathway-based analysis without gene-level collapsing may have higher power than pathway-based analysis with gene-level collapsing. We also provide a guideline in which situation pathway analysis with or without collapsing may be optimal.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1503T

Gene function prediction based on 80,000 expression samples identifies downstream effects of many trait-associated SNPs in an eQTL dataset of 5,311 samples. *J. Karjalainen[1], H. Westra[1], R.S.N. Fehrmann[1], T. Esko[2], M.J. Peters[3], E.M. Festen[1], G.J. te Meerman[1], S.B. Felix[4], H. Yaghootkar[5], J. Kettunen[6], M.W. Christiansen[7], A. Hofman[8], F. Rivadeneira[3], E. Reinmaa[2], R.C. Jansen[9], J. Brody[7], S.A. Gharib[10], A. Suchy-Dicey[7], D. Enquobahrie[11], A.G. Uitterlinden[12,8], C. Wijmenga[1], B.M. Psaty[7], S. Ripatti[6,13], T. Frayling[5], A. Teumer[14], A. Metsepalu[2], J.B.J. van Meurs[3], L. Franke[1].* 1) University Medical Center Groningen, University of Groningen, Groningen, Netherlands; 2) Estonian Genome Center and Institute of Molecular and Cell Biology of University of Tartu, Tartu, Estonia; 3) Department of Internal Medicine, Erasmus Medical Centre, Rotterdam, Netherlands; 4) Department of Internal Medicine B, University Medicine Greifswald, Germany; 5) Genetics of Complex Traits, Peninsula College of Medicine and Dentistry, University of Exeter, Exeter, United Kingdom; 6) Institute for Molecular Medicine Finland FIMM, FI-00014 University of Helsinki, Helsinki, Finland; 7) Cardiovascular Health Research Unit, Departments of Medicine and Epidemiology, University of Washington, Seattle, WA, United States; 8) Department of Epidemiology, Erasmus Medical centre, Rotterdam, Netherlands; 9) Groningen Bioinformatics Center, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Groningen, Netherlands; 10) Center for Lung Biology Department of Medicine University of Washington, Seattle, WA, United States; 11) Department of Epidemiology, University of Washington, Seattle, WA, United States; 12) Netherlands Genomics Initiative-Sponsored by the Netherlands Consortium for Healthy Aging, Rotterdam, Netherlands; 13) Wellcome Trust Sanger Institute, Hinxton, United Kingdom; 14) Interfaculty Institute for Genetics and Functional Genomics, Ernst-Moritz-Arndt University Greifswald, Greifswald, Germany.

High-throughput DNA microarray technology now provides us with a detailed view of the human mRNA transcriptome under different biological conditions. The increasing amount of publicly available microarray data helps in identifying and predicting genes that contribute to the same biological processes. To create a gene co-regulatory model of the human transcriptome, enabling the prediction of gene function, we analyzed 55,000 human, 17,000 mouse and 6,000 rat Affymetrix microarrays from the Gene Expression Omnibus. We created an integrated three-species gene network with 20,000 unique human gene orthologs and developed a principal component based algorithm to predict function for individual genes. We benchmarked the algorithm against several pathway databases (including Gene Ontology, KEGG and Reactome) and observed that empirically validated gene functions could be predicted with high accuracy. As the current method relies solely on expression data, it is not affected by reporting biases in (PPI) databases, abstracts or research questions. We applied our framework to various complex traits for which multiple genome-wide significant SNP associations have been reported. For most traits, our predictions were biologically relevant and often highlighted several pathways that had not been implicated before. For example, “regulation of ossification” was the most significantly enriched pathway among loci associated with human height. In order to pinpoint the most likely causative genes per locus and predict downstream effects of SNPs, we subsequently integrated eQTL data of 5,311 unrelated human peripheral blood samples. We investigated the aforementioned pathways that we had expanded with genes that are likely to be part of these pathways based on the co-expression of the 80,000 samples. When concentrating on SNPs that are known to be associated with complex traits we identified many SNPs for which downstream pathways showed strong enrichments of eQTL signals. These results further support our functional predictions, and indicate that through integration of many gene expression arrays biological knowledge can be obtained for SNPs that have been implicated in disease.

1504F

Constructing biological network using RNA-seq data. *Y. Choi[1], M. Coram[2], H. Tang[1].* 1) Department of Genetics, Stanford University School of Medicine, Stanford, CA; 2) Department of Health Research and Policy, Stanford University School of Medicine, Stanford, CA.

Next-generation technologies that generate large-scale genomic data allow the construction of complex networks that provide a new insight for understanding the molecular process. A number of statistical methods have been developed for constructing networks based on Gaussian assumption that may not be appropriate for non-Gaussian data such as RNA-seq. In this study, we propose a novel statistical approach to model a complex covariance structure on multivariate counts data. This method uses penalized likelihood method to estimate sparse partial correlations. However, the implementation of the method is problematic due to complexity of likelihood. Laplace integration technique is used to approximate the likelihood and its derivatives and Newton-Raphson method is applied to find Maximum-likelihood estimator (MLE). In the simulation studies, we show the proposed method performs better than the Gaussian models, especially in the case that the RNA expression levels are low and the distribution is skewed. We also apply the proposed method to two published RNA-seq data and construct networks that may generate testable hypotheses of genetic regulatory interactions.

1505W

Network Inference, Integrative Dynamic Omics and Personalized Medicine. *G.I. Mias, H. Im, E. Mitsunaga, R. Chen., J. Li-Pook-Than, L. Jiang, M. Snyder.* Genetics Dept, Stanford University, Stanford, CA.

The longitudinal monitoring of global molecular components and physiological states can offer medical insights in the diagnosis and treatment of disease. Such information allows for network inference that can be leveraged to personalized medicine through dynamically following multiple omics technologies. Pathway changes over time and associated network interactions (inferred nodes & connections) were first observed using multi-omic information from a proof-of-principle investigation, which explored the implementation of personalized medicine in an initially healthy individual over a period of (now) 30 months. The included observations spanned multiple healthy and two virally infected states (as well as the clinically diagnosed onset of Type 2 Diabetes). A novel dynamic network framework will be presented, building on a dynamic integrated Personal Omics Profiling analysis, which allows for observing associated known pathway changes from omics information (transcriptome, proteome, metabolome) in any system. The network framework infers novel pathways components and topological changes and has been extended to other longitudinal studies to display changes as an individual or system transitions through dynamical biological states. This observation of dynamic network changes and related pathways will provide a medically relevant interpretation of associated molecular interaction changes that has potential applications to a more personalized, precise and preventative medicine.

1506T

Stress-induced changes in gene interactions in human cells. *R. Nayak[1], M. Kearns[2,4], V.G. Cheung[3,4,5].* 1) Medicine, University of California - San Francisco, San Francisco, CA; 2) Department of Computer and Information Science; 3) Departments of Pediatrics and Genetics; 4) University of Pennsylvania, Philadelphia, PA; 5) Howard Hughes Medical Institute, Philadelphia, PA.

Cells encounter numerous stressors, and failure to cope can lead to human disease. Cellular response to stress includes altering the expression levels of genes and how genes interact with each other. Here, we sought to understand the cellular response in human B cells following exposure to ionizing radiation (IR) or endoplasmic reticulum (ER) stress. We measured the expression levels of over 5,000 genes in over 95 individuals, constructed coexpression networks, and examined interactions among genes. The analyses revealed that many genes change expression levels, and fewer genes change interactions, but genes that change interactions tend to be specific and critical to the stress response. Expression levels of over 70% and 25% of genes changed following ER and IR stress, respectively, which is more than reported previously in the literature. Gene coexpression networks were constructed by measuring correlation in gene expression levels across >95 individuals. Coexpression networks representing stressed states had similar “hubs” but increased clustering coefficients compared to unstressed states; this was true for both ER and IR stress. Further analysis revealed that while thousands of genes changed expression levels, only hundreds of genes changed interactions with other genes. Genes that changed interactions were enriched for known key players that are critical for responding to stress; these genes were also specific to the stress response. For example, key genes that changed interactions only following ER stress included *BIP*, *ERp72*, and *VCP*. Key genes that changed interactions specifically following IR stress included *DDB2*, *Survivin*, and *BNIP3*. Using this knowledge about genes that change interactions, we were able to identify new roles for genes in ER or IR stress, including genes that do not change expression levels. Together, these results suggest that cells tend to use existing interactions in response to acute stress, but will alter the interactions of key genes critical to the stress response. Thus, examining human gene interactions provides information on key genes and their functions in the cellular stress response.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1507F

An exploration of the use of low coverage whole genome sequencing datasets as additional controls for exome sequenced case series. *D. Gurdasani[1, 2], MS. Sandhu[1, 2], P. Kellam[2, 3], PJ. Openshaw[4], JK. Baillie[5], The MOSAIC and GenISIS Investigators.* 1) Department of Public Health and Primary Care, University of Cambridge, Cambridgeshire, UK; 2) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, UK; 3) Division of Infection and Immunity, University College London, London, UK; 4) Centre for Respiratory Infection, National Heart and Lung Institute, Imperial College, London, UK; 5) Division of Genetics and Genomics, The Roslin Institute, University of Edinburgh, UK.

Exome sequencing approaches are increasingly being applied to the study of complex and rare disease. Due to cost limitations, sequencing is often restricted to case series of interest. Exome sequence data from cases with other diseases are often used for comparison, which can bias results. In this context, approaches that utilise a broad range of control sets, including general population controls, may improve power, reduce bias, and lower type I error. Although large publicly available whole genome sequence data can provide additional controls and improve statistical resolution, comparison of data generated using different sequencing methods can also cause bias. We explored the utility of using low coverage 1000 Genomes Project (1000GP) data as controls for a case series of severe influenza. Methods: Cases of severe influenza were recruited from the UK MOSAIC and GenISIS consortia. We carried out deep (80x) exome sequencing on a Hiseq platform. Reads were realigned around indels and base qualities recalibrated. All polymorphic sites within defined exome target regions in 1000GP were called in cases using GATK and Samtools (standard filters). Only sites called in >80% samples with genotype concordance>80% between algorithms were included. Samples with <95% sites called were filtered. Single locus association testing was carried out using logistic regression with principal components as covariates. Wilcoxon Rank Sum test (WRST) was used to assess if p values were statistically lower for variants within genes implicated in influenza replication. Fisher exact test (FET) was used to assess GO term enrichment among the top 100 loci. Results: 1,092,514 sites from 60 influenza cases and 1092 1000GP controls were included in analyses. The genomic inflation factor was 1.04 for variants with MAF>0.05. Variants implicated in influenza replication in previous siRNA studies had lower p values (WRST p<2.2e-16). Variants predicted as deleterious using Polyphen, CONDEL or SIFT and those with predicted functional consequence had lower p values (WRST p<2.2e-16). The top 100 candidates were significantly enriched for GO terms related to inflammatory (FET p=0.001) and immune response (FET p=0.003). Conclusion: In this proof-of-concept study, we show that with stringent filtering thresholds, low coverage sequencing sets may serve as additional sources for controls for exome sequencing case series and identify plausible variants consistent with known pathogenic mechanisms.

1508W

Genome Profile-based Disease Risk Prediction. *P. Hu[1], E. Boerwinkle[2], L. Jin[1], M. Xiong[2,1].* 1) MOE Key Laboratory of Contemporary Anthropology, School of Life Sciences, Fudan University, Shanghai, China; 2) Human Genetics Center, University of Texas School of Public Health, Houston, TX, USA.

There is bit debate about whether DNA variation has value to predict disease. Recently, a study published in Science Translational Medicine claimed that whole genome sequencing is not informative to predict common disease for all. However, many geneticists questioned their methods, models, assumptions and conclusions (http://blogs.nature.com/news/2012/04/questioning-value-personal-genomics.html). The purpose of this paper is to investigate the predictive values of common and rare genetic variants for complex disease by one set of GWAS (psoriasis: 1,346 cases and1, 377 controls, 429,499 SNPs and a set of exome sequencing data (early of myocardial infarction: 611 cases and 940 controls, 278,191 SNPs). The current paradigm for genetic diagnosis is to predict disease risk of individuals by limited loci known to affect risk of complex diseases. Fast and economic next generation sequencing (NGS) technologies will generate millions or even ten millions of genetic variants. Reducing the dimensionality of data without losing intrinsic information is a key step for genome-wide prediction of common diseases. We use genomic information content in a region and LASSO regulation techniques to develop a novel global feature selection method that can reduce dimensions from millions to hundreds to optimally select important variables as prediction sets where the classification accuracy of these variables are assessed. Bayesian discriminant analysis and logistic regression methods are used to predict individual disease risk. We use 10 fold cross validation to quantify the value of prediction of genetic risk to diseases. We find that the media AUC values of classifying psoriasis and early of myocardial infarction in the test sets are 69% and 70.6% by logistic regression analysis, respectively. These empirical results suggest that genetic variants have valuable value to prediction of individual disease risk.

1509T

Whole Genome Sequencing based Imputation: A Comparison between the 1000 Genomes data and the UK10K Sequencing Data. *J. Huang.* Human Genetics, Wellcome Trust Sanger Institute, Hinxton, United Kingdom.

Whole-genome sequencing provides the most-comprehensive characterization of genetic variation, but it is still prohibitive for researchers to conduct whole-genome sequencing on all samples from large population-based cohorts. Statistical inference of missing genotypes (imputation) has been proved successful in Genome-wide association studies. The 1000 Genomes project provides phased haplotypes for more than 1,000 samples covering 11 populations, including 379 samples of European ancestry. As part of the UK10K cohorts project, 4,000 whole genomes are being sequenced at low coverage (average 6x) from two population based UK studies: TwinsUK and ALSPAC. We report a comprehensive evaluation on the imputation performance by using the publically available 1000 genomes reference panel versus the latest sequencing data from the on-going UK10K project. We adopt a 2-step approach by first pre-phasing target genotype data and then running imputation based on haplotype matching between target and reference panels. The pre-phasing step uses SHAPE-IT for each of the 22 autosomal chromosomes without splitting into chunk. The second step uses IMPUTE2, where 5MB segments were extracted for parallel processing. We evaluate imputation quality for 5,667 samples in the WTCCC2 common controls genotyped by the Illumina1M panel, which includes 1,158,122 SNPs for 5,667 unrelated individuals. The imputation quality metrics include imputation information score, correlation coefficient and concordance between true and imputed genotype. We assess the overall genome-wide imputation quality, as well as quality in different spectrum of allele frequency and functional annotations. Our analysis and results will provide empirical information that aids investigators to choose reference panels for imputing their own genotype data.

1510F

EPACTS : A flexible and efficient sequence-based genetic analysis software package. *H.M. Kang, X. Zhan, X. Sim, C. Ma.* Biostatistics Dept, Univ Michigan, Ann Arbor, Ann Arbor, MI.

The advent of next-generation sequencing technologies presents an unprecedented opportunity to identify genetic variants causal for complex traits. At the same time, existing and novel sequence-based statistical methods have major software challenges because innovative statistical methods often stay only as in-house software and fail to become widely available. This is primarily bottlenecked by the implementation cost of supporting a complex variety of genetic variants in a computationally efficient manner. We developed a flexible and efficient sequence-based genetic analysis software package EPACTS (Efficient and Parallelizable Association Container Toolbox for Sequence Data), which allows methods developers to dynamically plug-in user-defined single-variant tests or burden tests in both binary and quantitative traits with simple R script. Methods developers do not need to implement low-level data processing, but leverage efficient and parallelizable back-end libraries to directly access the Variant Call Format (VCF) files, up to orders of magnitude faster than most widely available software.

EPACTS provides users and developers with the following key features: (1) Integrated support of variant annotation and association summary and visualization, including QQ plot, Manhattan plot, locus zoom plots. (2) Built-in implementations of a wide variety of single variant tests and rare variant burden tests, including standard regression, score, and likelihood ratio tests, non-parametric rank test, collapsing methods, variable threshold test, Sequence Kernel Association Test (SKAT), and mixed model methods accounting for hidden relatedness. (3) Efficient and concurrent access to VCF files enabling a genome-wide association in half an hour across thousands of individuals and tens of millions of variants on a single machine. (4) Support both for discrete hard genotype calls and likelihood-based soft genotypes. (5) Support for SNPs, short indels, and structural variants (6) Support for efficient permutation and multiple imputation of missing genotypes, and (7) Convenient utilities for basic quality control and population genetic analysis with visualization support. EPACTS is publicly available with an open interface for users to submit new statistical tests. We provide tutorials demonstrating that EPACTS provides a seamless connection between scientific analysts and statistical methods developers in the era of next-generation association studies.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1511W

A Monte Carlo procedure for assessing co-occurring and anti-co-occurring gene mutations in cancer genome sequencing studies. *P. Liu, X. Hua, H. Xu, Y. Lu.* Dept Physiology and Cancer Center, Medical College of Winconsin, Milwaukee, WI.

Cancers are caused by the accumulation of genomic alterations. Next-generation sequencing technologies have revolutionized cancer genomics research by providing an unbiased and comprehensive method of detecting somatic cancer genome alterations, including nucleotide substitutions, insertions and deletions (indel), copy number alterations, and chromosomal rearrangements. Recent sequencing experiments have brought success in the identification of several cancer genes that were highly frequently mutated in tumors including IDH1, IDH2, DNMT3A, BAP1, ARID1A, ARID2, MLL2, GRIN2A, GRM3 and PBRM1. These experiments have provided a large catalogue of somatic mutations in individual tumors. Mutations in some cancer genes tend to co-occur (termed co-occurring mutations), suggesting that both of them provide additively proliferative advantages to tumors. However, mutations in other genes tend to occur mutually exclusively, suggesting that they may be involved in the same pathway with similar downstream effectors. To better understand these functional interactions among cancer genes, we introduced a Monte Carlo procedure for assessing co-occurring and anti-co-occurring mutations in cancer genes. Our method accounts for variations in gene length and differences in background mutations among individual tumors that were not addressed by previous methods. To demonstrate the utility of our method, we applied it to the analysis of the significantly mutated genes from lung cancer sequencing projects. The results from our analysis provide unprecedented insights into mutational processes and gene networks associated with lung cancer.

1512T

Using whole exome sequencing to identify rare causal variants for oral clefts in multiplex families. *T.H. Beaty[1], I. Ruczinski[1], M.M. Parker[1], J.B. Hetmanski[1], P. Duggal[1], M.A. Taub[1], S. Szymczak[2], Q. Li[2], C. Cropp[2], H. Ling[3], E.W. Pugh[3], Y.H. Wu-Chou[4], J.E. Bailey-Wilson[2], M.L. Marazita[5], J.C. Murray[6], E. Mangold[7], M.M. Noethen[7], K. Ludwig[7], A.F. Scott[8].* 1) School of Public Health, Johns Hopkins Univ., Baltimore, MD; 2) Inherited Disease Branch, NHGRI, NIH, Baltimore, MD; 3) Center for Inherited Disease Research, Baltimore, MD; 4) Chang Gung Memorial Hospital, Taipei, Taiwan; 5) School of Dental Medicine, Univ. of Pittsburgh, Pittsburgh, PA; 6) School of Medicine, Univ. of Iowa, Iowa City, IA; 7) Institute of Human Genetics, Univ. of Bonn, Bonn, Germany; 8) School of Medicine, Johns Hopkins Univ., Baltimore MD.

Non-syndromic oral clefts (cleft lip, cleft palate and cleft lip & palate) are common birth defects with a complex and heterogeneous etiology. Several genes and regions have been associated with risk in case-control and case-family studies, and genome-wide linkage studies of multiplex families have identified additional regions likely to harbor causal genes. This whole exome sequencing (WES) study used 108 affected 2º and 3º relatives drawn from 52 multiplex families originally recruited for linkage (4 families with 3 affecteds and 48 families with 2 affecteds each). While the WES approach has been successful in identifying novel causal genes for Mendelian diseases, it has not previously been applied to oral clefts.

WES was done by the Center for Inherited Disease Research using the Agilent SureSelect v.4 capture reagents & Illumina HiSeq 2000 sequencers. Initially, we focused on truly novel single nucleotide variants (SNVs), i.e. not previously reported, shared by affected relatives within a family (exact genotype matches only) and predicted damaging by SIFT score (≤0.05). A total of 516 novel SNVs were identified as shared between affected relatives in 52 families. Only one truly novel SNV (A→G at hg19 position 3056632) in ZNF764 was shared by affected relatives across 2 families, and these 2 families came from the same recruitment site. We focused on 334 candidate genes (plausible by molecular function, biological process or pathway) identified in Jugessur et al. (2009, PLoS ONE 4:e5385), and identified 5 genes: CDH1 (16q22), FGF8 (10q24), FGFR4 (5q35), GAD2 (10p11), and TRPS1 (8q24). Most of these genes are recognized candidates for oral clefts or cancers. Mutations in FGF8 underlie Kallman syndrome, which can include oral clefts; mutations in CDH1 have been reported in families with oral clefts and diffuse gastric cancer (MIM 192090); mutations in TRPS1 cause the abnormal craniofacial development of tricho-rhino-phalangeal syndrome. Observed SNVs included 1 Stop/Gain and 4 nonsynonymous SNVs. These affected individuals all have apparently inherited forms of oral clefts, and because they are heterozygous for damaging SNVs, they may result from haploinsufficiency of distinct gene products. Additional studies are needed to confirm causality.

1513F

Analyzing Deep Whole Genome Sequence and Genotype Data of >1,000 Individuals from Large Mexican-American Pedigrees in the T2D-GENES Study. *G. Jun[1], M. Almeida[2], P. Cingolani[3], A. Wood[4], C. Fuchsberger[1], T. Teslovich[1], T. Dyer[2], M. Rivas[8], K. Gaulton[8], J. Maller[8], J. Curran[2], J. Grunstad[5], T. Blackwell[1], D. Lehman[6], R. Grossman[5], S. Lincoln[7], J. Laramie[7], M. Boehnke[1], M. McCarthy[8], T. Frayling[4], R. Sladek[3], R. Duggirala[2], J. Blangero[2], G. Abecasis[1].* 1) Dept Biostatistics, Univ Michigan, Ann Arbor, Ann Arbor, MI, United States; 2) Texas Biomedical Research Institute, San Antonio, TX, United States; 3) McGill University, Montreal, QC, Canada; 4) University of Exeter, United Kingdom; 5) University of Chicago, Chicago, IL, United States; 6) University of Texas Health Science Center at San Antonio, San Antonio, TX, United States; 7) Complete Genomics Inc, Mountain View, CA, United States; 8) Oxford University, Oxford, United Kingdom.

Deep whole genome sequencing from large pedigrees has the potential to provide unique insights into the role of rare and private variants in disease. Through the T2D-GENES Consortium, we selected for deep whole genome sequencing 600 of 1,040 individuals from 20 Mexican American pedigrees. Our design allows sequenced data to be combined with GWAS genotypes to impute whole genome data on all remaining family members. Our current data includes deep sequence data on 551 individuals, sequenced by Complete Genomics at an average depth of 60x. Rich longitudinal phenotypic data is available for most individuals, including 316 with type 2 diabetes. We took a series of quality control steps to identify samples with low data quality. Pedigree information was verified by estimating kinship coefficients. Variants are annotated for various sequence quality metrics and filtered by support vector machine to separate likely variants from likely false-positives. We then annotated the impact of each variant on protein structure. This study allows us to examine the segregation of rare and private variants through multiple generations, together with diabetes disease status and related quantitative traits. In the genome of an average individual, we observed 3.4M SNPs, 450K indels, and 2K larger structural variants. Overall, we discovered 24.9M SNPs, including 6.5M singletons, 6.2M indels, and 2.0M multi-nucleotide variants. >50% of SNPs had frequency <1%. After SNP filtering, the Mendelian inconsistency rate was 0.63% and the overall transition to transversion ratio was 2.20. Typically, sample specific rare variants are extremely rare - making it challenging to evaluate their phenotypic effects. Our design means that many sample specific rare variants are seen in many individuals, because they are shared in many individuals that share a particular haplotype as it segregates through a family. Our initial analysis will focus on assessing the impact of identified 170K missense, including 37 variants in genes previously implicated in mature onset diabetes of the young (MODY), and 4K nonsense variants on the risk of type 2 diabetes. Also planned are analyses that evaluate the association between the overall burden of rare variation in individual genes and disease status and analyses that include non-coding variants. Our large set of deeply sequenced large pedigrees provides a unique resource in which to evaluate the association of individually rare variants with type 2 diabetes.

1514W

A Flexible and Effective Poisson Mixed Model Framework for RNA Sequencing and other Count Data. *X. Zhou[1], J. Tung[2], Y. Gilad[1], S. Mukherjee[3, 4], M. Stephens[1, 4].* 1) Department of Human Genetics, University of Chicago, Chicago, IL; 2) Department of Evolutionary Anthropology, Duke University, Durham, NC; 3) Department of Statistics, Duke University, Durham, NC; 4) Department of Statistics, University of Chicago, Chicago, IL.

High throughput sequencing is extremely widely used in genetics and genomics, providing unprecedented insights into many basic biological questions by accurately quantifying RNA abundance, identifying protein-DNA binding sites and measuring open chromatin regions. However, analyzing read count data presents several important statistical and computational challenges. For example, many expression studies have noted and illustrated the importance of taking account of sample structure, batch effects, and unmeasured confounding factors, in the analysis. However, the analytic tools used to take account of these factors are currently only implemented for normally-distributed data, and not for the kinds of count-based data generated by sequencing. As a result, previous count-based models (e.g. based on Poisson, negative binomial, or beta-binomial distributions) fail to account for sample similarity caused by confounding factors and/or individual relatedness. Here, we rectify this problem, using a Poisson mixed effects model with two random effects terms to account for both independent variation and sample relatedness. We present a novel and efficient posterior sampling algorithm for the model, taking advantage of the recently published genome-wide mixed model association (GEMMA) algorithm. We illustrate our model using two applications: estimating gene expression heritability and detecting differentially expressed genes. Using a real data set, we show that the heritability estimates for gene expression obtained from our model closely resemble those obtained from several previous microarray experiments. Using both simulations and real data sets, we further show that our model is much more powerful than several widely used count based approaches in detecting differentially expressed genes.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1515T

Sequence-based analysis identifies variants associated with serum lipids in African Americans. *A. Bentley[1], D. Shriner[1], A. Doumatey[1], J. Zhou[1], H. Huang[1], J. Mullikin[2], R. Blakesley[2], N. Hansen[2], G. Bouffard[2], P. Cherukuri[3], B. Maskeri[2], A. Young[2], A. Adeyemo[1], C. Rotimi[1].* 1) Center for Research in Genomics and Global Health, Natl Human Genome Research Institute, Bethesda, MD; 2) NIH Intramural Sequencing Center, Natl Human Genome Research Institute, Rockville, MD; 3) Office of the Clinical Director, Natl Human Genome Research Institute, Bethesda, MD.

Interethnic differences in serum lipid distributions are consistently observed, with African Americans (AA) generally having healthier lipid profiles than other US ethnicities despite a higher prevalence of risk-associated lifestyle factors. Additionally, interethnic differences in the relationship between lipids and metabolic disorders have been identified. Given the disproportionate burden of these disorders among AA, understanding the genetic determinants of lipid concentrations are critical. To identify genetic factors that influence lipids in AA, we sequenced four lipids-influencing genes (*LPL*, *ABCA1*, *PON1*, and *SERPINE1*) in a subset of Howard University Family Study participants. Participants selected for the sequencing were at the extremes of lipid distribution, either at high risk (high TG and low HDL, n=24) or low risk (low TG and high HDL, n=24). From the variants identified, 103 SNPs were genotyped in the full sample (n=2181). A variety of techniques for analyzing common, rare, and joint common and rare variants were employed to evaluate and confirm associations between these SNPs and serum lipids. A common *LPL* variant rs328 was associated with dramatically increased HDL and homozygosity for the variant allele was associated with 12.7 mg/dl higher HDL (p=0.008). The minor allele of this variant was additively associated with 5% lower TG (p=0.0004). The association with TG was confirmed in a sample of West Africans (3% lower TG, p=0.049). When a method that simultaneously evaluated common and rare variants was used, another common *LPL* SNP, rs1059611, that was not in LD with rs328, was found to be associated with higher HDL (2.6 mg/dl, p=0.02). Rare variant collapsing methods identified positively-associated *ABCA1* variants (n=9, +4% TG, p=0.01) and negatively-associated *PON1* variants (n=2, -8.4 mg/dl HDL, p=0.008) as having roles in serum lipid distributions. In summary, we have identified rare and common variants associated with serum lipids in AA. Beginning the analysis with sequencing has clear advantages for the identification of rare variants; however, this strategy also supports inferences about common variants. For instance, rs328 has been identified previously, but debate exists regarding its role as a causal or proxy SNP; haplotype analysis from this work supports the direct association of rs328 with lipids. Results from this analysis provide new evidence into the genetic determinants of serum lipids in AA.

1516F

Genetic factors associated with levels of immune cell types. *M. Steri[1], V. Orrù[1], E. Fiorillo[1], G. Sole[1], C. Sidore[2,3], F. Virdis[1], M. Dei[1], S. Lai[1], A. Mulas[1], M.G. Piras[1], M. Lobina[1], M. Marongiu[1], M. Zoledziewska[1], M. Congia[4], F. Busonero[1], A. Maschio[1,2], M.F. Urru[5], M. Marcelli[5], R. Atzeni[6], D. Firinu[7], M. Valentini[6], W. Mentzen[1], S. Naitza[1], M.B. Whalen[1], A. Angius[1,5], C.M. Jones[6], D. Schlessinger[8], G. Abecasis[2], S. Sanna[1], F. Cucca[1].* 1) Istituto di Ricerca Genetica e Biomedica (IRGB), CNR, Monserrato, Italy; 2) Center for Statistical Genetics, University of Michigan, Ann Arbor, MI, USA; 3) Dipartimento di Scienze Biomediche, Università di Sassari, SS, Italy; 4) Dipartimento di Scienze Biomediche e Biotecnologie, Università di Cagliari, CA, Italy; 5) CRS4, GSP group, Parco tecnologico della Sardegna, Pula, Cagliari, Italy; 6) CRS4, AGCT group, Parco tecnologico della Sardegna, Pula, Cagliari, Italy; 7) Dipartimento di Allergologia e Immunologia, Università di Cagliari, CA, Italy; 8) Laboratory of Genetics, NIA, Baltimore, MD, United States.

Genome-wide association scans (GWAS) have identified thousands of regions associated with quantitative traits of biomedical interest, but they mostly focused on broadly classified biomarkers and used incomplete genetic maps. We employed a GWAS approach with enhanced genetic and phenotypic resolution to better understand genetic factors regulating the immune system and its involvement in diseases. We evaluated quantitative variation of 272 immune traits, representing the majority of the leukocyte populations like monocytes, granulocytes, circulating dendritic cell subsets as well as lymphocytes subdivided into NK cells, B cells and T cell subsets (regulatory T cells and T cell maturation states), using flow cytometry, in 1,629 volunteers of the SardiNIA project. Subjects were genotyped with Affymetrix arrays, the MetaboChip and ImmunoChip, and ~8.2 Million variants were imputed from a reference panel of 1,656 haplotypes deriving from 828 Sardinian samples sequenced at 3x coverage, on average. Heritability estimates showed that levels of most cells showed considerable heritability, with the genetic component accounting for >40% of the phenotypic variation for most of the traits (ranging from 3% to 87%). We performed a GWAS for each trait and observed, overall, 93 independent variants at 56 loci ($1\times10^{-202}<P<5\times10^{-8}$), including 4 hitherto unreported variants. These 93 variants accounted for 2% of the phenotypic variation, with 14 explaining >5% and 3 >15%. In addition, 32 of them were associated with more than one trait, suggesting the presence of major genetic components that regulate different parts of the immune system. These large genetic effects are in contrast with findings in other GWAS for quantitative traits, for which effect sizes are typically small, leaving a large proportion of heritability unexplained. Indeed, when considering all the independently associated variants, the heritability that remains to be explained was <20% for 8 traits and <50% for 45 traits. Variants at five loci are also associated with autoimmune diseases, such as multiple sclerosis and type 1 diabetes, crucially several associations were found with diseases for which an immunological component is yet unclear. Our results thus demonstrate that sequencing based GWAS, coupled with highly specific phenotype measurements, can improve our understanding of genetic factors and related biological mechanisms underlying physiological and pathological variations in the population.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1517W

Characterization of rare variants in melanoma-associated genes in melanoma-prone families without CDKN2A/CDK4 mutations using exome sequencing data. *R. Yang, K. Jacobs, M. Cullen, J. Boland, L. Burdett, M. Malasky, M. Rotunno, M. Yeager, S. Chanock, M. Tucker, A. Goldstein.* Dept DCEG, NCI, Bethesda, MD.

Approximately 10% of cutaneous malignant melanoma (CMM) cases occur in a familial setting. CDKN2A and CDK4 are the two major susceptibility genes for CMM identified to date. However, these two genes only account for melanoma susceptibility in a small proportion of melanoma-prone families, suggesting the existence of other high-penetrance genes. Exome sequencing has been successful in identifying rare genetic variants underlying rare Mendelian diseases. The goal of this study was to identify additional high-penetrance genes in melanoma-prone families without known mutations. We conducted exome sequencing in blood-derived DNA in 46 CMM cases from 14 melanoma-prone families with three or more affected members. We sequenced five CMM cases in one family, four cases in two families, and three cases in the remaining families. We first filtered out synonymous and common variants that had minor allele frequency >0.05 among our internal controls or in public databases (dbSNP, 1000 Genomes, or NHLBI's Exome Variant Server [ESP]). After this initial filtering step, we observed 1,275 to 3,301 variants in each family. We then applied a dominant segregation model that required variants to be present in all patients (for 3 cases sequenced in a family) or in all but one patient (for >3 cases sequenced in a family) within a family. The number of variants was thus narrowed to 46 to 254 (median 150) in each family. Since disease causing variants are likely to be very rare in the general population, we further filtered variants that were seen more than once among our internal controls (>1/200) or more than five times in ESP. Even with all of these filtering steps, each family still showed many potentially disease-related rare variants (8–84, median 35) and most variants occurred in genes with potential relevance to cancer-related processes. Variants in the same genes were rarely observed in multiple families, suggesting that yet-to-be identified high-penetrance genes may be much rarer than CDKN2A. Our findings suggest that identifying causative variants using exome sequencing for complex diseases like melanoma may be challenging and will require careful prioritization of variants, sequencing of large numbers of families, additional information from other types of evaluations such as gene expression, and potentially functional studies.

1518T

Comparison of variant calling strategies for large-scale exome sequencing projects. *J. Floyd, A. Hendricks, L. Crooks, K. Walter, S. McCarthy, C. Anderson, The UK10K Consortium.* Wellcome Trust Sanger Institute, Cambridge, United Kingdom.

Large-scale sequencing projects in which many hundreds of both affected and unaffected samples are sequenced, as opposed to only one or a handful of disease-affected individuals, are becoming more numerous. When undertaking large-scale whole-genome sequencing projects, it is the norm to analyse the sequence data of all individuals simultaneously in order to accurately call variants. This is because for large-scale whole-genome sequence-based association studies, sequencing depth is typically too low to call variants on a per sample basis. With exome sequence data however, the depth is typically much higher, and calling variants on an individual sample basis becomes possible. However, whether this is always the optimal strategy is not certain, and will depend on the specific aims of the project. For projects with the aim of identifying common (>5% MAF) or low-frequency (1% - 5% MAF) variants affecting common complex disease, multi-sample calling across all sequenced individuals is likely to be the best strategy. However, if the objective is to identify extremely rare high-penetrance variants that may be causing rare Mendelian disease, or even private de novo mutations, then single-sample calling may instead be most powerful. Using data from the UK10K project - a large-scale Wellcome Trust funded sequencing project with the aim of identifying novels genes and pathways for a variety of human diseases - direct comparisons will be made across more than 3,000 exomes called by both single-sample (using a combination of SAMtools and GATK) and multi-sample calling (using SAMtools). Comparisons will be made with respect to how well each method performs in identifying variants at differing minor allele frequency. Comparisons of the filtering stringencies to which the variant calls were subjected will also be performed, since this can have a large effect upon the final number of variants brought forward for analysis. Finally, the call-sets will be assessed with regard to their ability to call a set of known disease causing variants that were validated experimentally.

1519F

A catalogue of structural variants identified in disease-based whole genome sequencing. *A. Kumar[1,2,3], R.D. Pearson[1], R.E. Handsaker[4], K.J. Gaulton[1], S.A. McCarroll[4], M.I. McCarthy[1,2], GoT2D consortium.* 1) WTCHG, University of Oxford, Oxford, Oxfordshire, United Kingdom; 2) OCDEM, University of Oxford, Oxford, United Kingdom; 3) Swiss Tropical and Public Health Institute, University of Basel, Switzerland; 4) Department of Genetics, Harvard Medical School, Boston, Massachusetts, USA.

Structural variants (SVs) are a class of human variation representing deletions, duplications and inversions of genomic regions varying from several hundred bases to megabases in size. They have been implicated in the pathogenesis of numerous disorders. Whole genome sequence (WGS) data affords the opportunity to interrogate SVs for contribution to disease across a wider range of allele frequencies and genomic size than previously possible. We have generated a map of SV deletions using WGS data from the Genetics of Type 2 Diabetes (GoT2D) study, which is characterising the contribution of low frequency and rare variants to T2D using a combination of low-coverage WGS data, deep exome sequencing and Illumina Omni 2.5M array genotyping. Discovery and genotyping of SVs was performed using GenomeSTRiP, which integrates sequence read pair and read depth features to identify deletions, applying quality control criterion similar to those of the 1000 Genomes (1000G) project. In an interim freeze of 918 samples of Northern European origin, we discover a total of 16,136 SVs genome-wide. Comparison of chromosome 20 SVs from these samples reveal high concordance with SVs from 1092 samples (AFR=246, AMR=181, ASN=286, EUR=379) in 1000G phase 1. Of 169 variants with HWE > 1e-10 (143 in 1000G), 16/169 were singletons (13/143 in 1000G) and 43/169 had MAF>0.05 (28/143 in 1000G). On chromosome 20, 1000G reports 22 SVs in Europeans, while GoT2D finds 43 SVs, 21 of them overlapping with 1000G (reciprocal overlap > 0.97). The remaining 22 common SVs have a much smaller length distribution (17 SVs <1000bp) owing to tight insert size distribution. In a preliminary analysis, we identified several rare SVs that overlap genes with known involvement in monogenic and polygenic forms of T2D including HNF1A, PPARG, HMGA2, IGF2BP2 and SLC30A8. We are extending discovery and genotyping to a final set of 2,772 whole-genome low coverage (~4x) samples. This will give an unprecedented opportunity to investigate the role of SVs in type 2 diabetes and related traits as well as providing an improved catalogue of SVs in European populations.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1520W

On the analysis of rare, exonic variation amongst subjects with autism spectrum disorders and population controls. *L. Liu[1], E. Boerwinkle[2], J. Buxbaum[3], E. Cook Jr.[4], B. Devlin[5], G. Schellenberg[6], J. Sutcliffe[7], M. Daly[8], R. Gibbs[2], K. Roeder[1], ARRA Autism Sequencing Consortium.* 1) Department of Statistics, Carnegie Mellon University, Pittsburgh, Pennsylvania, 15213; 2) Human Genome Sequencing Center, Baylor College of Medicine, Houston, Texas, 77030; 3) Seaver Autism Center for Research and Treatment, Mount Sinai School of Medicine, New York, New York, 10029; 4) Department of Psychiatry, University of Illinois at Chicago, Chicago, Illinois, 60608; 5) Department of Psychiatry, University of Pittsburgh Medical School, Pittsburgh, Pennsylvania, 15213; 6) Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, 19104; 7) Vanderbilt Brain Institute, Department of Molecular Physiology & Biophysics and Psychiatry, Vanderbilt University, Nashville, Tennessee 37232; 8) Analytic and Translational Genetics Unit, Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts, 02114.

Whole exome sequencing (WES), performed at the Baylor and Broad sequencing centers, produced data from 1039 subjects diagnosed with autism spectrum disorders (ASD) and 863 controls selected from the NIMH repository to be of similar ancestry to cases. We report on a case-control analysis with several goals: (1) contrast the distribution of rare variation between cases and controls to determine if ASD risk genes can be identified in a sample of this size; (2) investigate the impact of population structure on association test statistics and efficacy of traditional remedies for confounding arising from this structure; and (3), due to the structure of the data, assess the efficiency of meta- versus mega-analysis for common gene-based analyses of association. Data produced by two different centers and using two different sequencing platforms make (3) a particularly rich problem to be solved. Using a variety of gene-based tests and both meta- and mega-analysis, we find no new risk genes for ASD in this sample of cases and controls. Even for known risk genes the evidence from gene-based tests is not sufficiently compelling to “replicate” any gene. Our results suggest that standard gene-based tests will require much larger samples of cases and controls before being effective for gene discovery, even for a disorder like ASD, for which rare variants are known to play an important role in risk. Like other recent studies, we find evidence that population structure can confound case-control studies by the clustering of rare variants in ancestry space. Yet unlike recent studies we find that principal component based analyses are sufficient to control for ancestry and produce test statistics with appropriate distributions. A surprising result from our study involves the performance of meta- versus mega-analysis. It is widely understood that mega- and meta-analysis are typically equivalent for tests of common variant association. But the tests involved are of linear form, whereas many gene-based tests are quadratic. In the quadratic setting we show by theory and simulations that mega-analysis has better power than meta-analysis. The drawback, of course, is that combining datasets presents other challenges, which will be discussed.

1521T

Resolving dependence between overlapping reads in next-generation sequencing data. *Y. Lo[1], G. Abecasis[1], S. Zöllner[1,2].* 1) Biostatistics, University of Michigan, Ann Arbor, MI; 2) Psychiatry, University of Michigan, Ann Arbor, MI.

Generating genotype calls from paired-end next generation sequencing data assumes that each paired-end mate gives independent genotype. In practice read ends often overlap, especially in exome sequencing data, when PCR fragments are short. This overlap violates the independence assumption because overlapping reads replicate the same PCR error. Assuming independence at overlapping bases leads to overestimation of genotype call accuracy. Hence, one of the overlapping reads is typically removed from the data. However, each base in an overlapping region is sequenced twice; each sequencing reaction provides additional evidence about the true underlying genotype. Hence, ignoring one read may lead to substantial data loss. For example, in a low-pass sequencing study, about 13% of the read pairs overlap, and the average length of overlapping region is 44bp out of the 100bp reads. We propose a method (RESCORE) to remove the dependence in overlapping reads, while retaining genotype information from sequencing. To resolve dependence, we calculate a preliminary base quality statistic per overlap position based on the concordance of the bases and on the original quality scores. We turn these preliminary statistics into calibrated quality scores by binning genotypes by their statistic and comparing the reads in each preliminary quality statistic category with the reference genome, excluding known sites of variation. Assuming that all mismatches are genotyping error, we can thus estimate calibrated quality for each category. We apply RESCORE on 1,200 individuals with aligned sequence reads from the 1000 Genomes Project Exome data. We compare genotype calls from original reads, reads modified by RESCORE, as well as reads modified by retaining one read at the overlapping regions. Using RESCORE, quality score at each overlapping base is higher than each original read base, suggesting more accurate estimates of genotype uncertainty. Comparing resulting SNP calls on chromosome 20, RESCORE gives slightly fewer number of singletons, but at higher overall quality. Improved transition-transversion ratio suggests that the call set from RESCORE contains fewer false positives. At the shared singletons, we see higher SNP and genotype qualities using RESCORE. We recommend incorporating RESCORE into standard variant calling pipeline, especially for low-pass sequencing studies.

1522F

Single point and rare variant analysis by whole genome sequencing in 4,000 samples in the UK10K cohorts project. *J.L. Min, The UK10K Consortium (Cohorts Group).* School of Social and Community Medicine, University of Bristol, Bristol, United Kingdom.

The UK10K project is a collaboration between multiple research centres mainly in the UK aiming to uncover rare genetic variants contributing to disease and health status by sequencing 10,000 people. As part of UK10K, 4,000 individuals from two deeply phenotyped cohorts - TwinsUK and the Avon Longitudinal Study of Parents and Children (ALSPAC) - have been sequenced to average 6.5x coverage using next-generation sequencing technology. In the current release of 2,357 whole genome sequences (1,617 from TwinsUK and 740 from ALSPAC), we tested single variants and collapsed sets of variants for association with 46 quantitative traits using several statistical methods. Genotype accuracy is high and there is no evidence for inflation of association summary statistics. Several variants were found confirming previously known genome-wide association results, including CETP for HDL, UGT1A1 for Bilirubin, SLC2A9 for Uric acid, ABO for alkaline phosphatase, APOE for Total cholesterol and LDL and HBS1L-MYB for several blood count measurements. Using six different collapsing methods to test the combined effect of multiple rare coding variants on 10 known loci with common and/or rare variants for various traits also revealed consistent results indicating positive signals for CETP and HDL and SLC2A9 and uric acid (empirical $p<0.01$, $p\sim1e\text{-}5$). As an example of the possibility of extra information at known or new signals found by these methods, preliminary results show that after conditioning on rs1800775 for HDL and rs11727666 for uric acid, additional variants in the allele frequency range of 1–5% and 2–46% appear to contribute to the signal around CETP and SLC2A9 in the collapsing test ($p\sim0.001$). Initial novel signals are being followed up. Currently analysis of a final set of 3910 samples is under way, and variants discovered through genome-wide sequencing of the TwinsUK and ALSPAC cohorts are being imputed into the full genotyped cohorts to increase the power of the UK10K study. Our results provide insights into how large-scale whole-genome sequencing efforts are likely to reveal for the genetic architecture of complex traits and might significantly increase the initial estimates of explained heritability.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1523W

A Statistical Framework for the Evaluation of *De Novo* Variation in Psychiatric Disease. *K.E. Samocha[1, 2, 3, 4], B.M. Neale[1, 2, 3], M.J. Daly[1, 2, 3], ARRA Autism Sequencing Consortium.* 1) Departments of Genetics and Medicine, Harvard Medical School, Boston, MA; 2) Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA; 3) Program in Medical and Population Genetics, Broad Institute, Cambridge, MA; 4) Program in Genetics and Genomics, Biological and Biomedical Sciences, Harvard Medical School, Boston, MA.

With the increasing number of exome sequencing studies focused on *de novo* mutation discovery, it has become critical to have a rigorous method to evaluate the *de novo* findings. Particularly, it is important to determine the potential significance of genes that harbor multiple *de novo* events, especially as the number of sequenced individuals rises. The need for this statistical framework was recently highlighted by four exome sequencing studies of autism spectrum disorders that found an apparent excess of genes that had two loss-of-function (LoF) mutations (Neale et *al*. 2012; Sanders et *al*. 2012; O'Roak et *al*. 2012; Iossifov et *al*. 2012). While rough estimations were used to determine significance, we propose here a systematic approach to estimating individual gene significances as well as the significance of experiment-wide excesses. To properly evaluate the probability of mutation, it is key to precisely estimate the expected mutation rate. To do so, we plan to apply a mutational model based on intergenic SNPs identified by the 1000 Genomes Project to each gene region in the genome. This mutational model incorporates local sequence information to determine the probability of each possible base change for every genomic base. We define the target bases of mutation as those in coding sequence or in the conserved splice sites. We then combine the individual probabilities of each possible substitution to determine the overall probabilities of a synonymous, missense, nonsense, splice site, frameshift, and inframe mutation per gene per trio sequenced. These probabilities can be used directly to evaluate individual gene observations, as well as to simulate efficiently large sequencing studies similar to the ones recently conducted for autism spectrum disorders. While this method is applied to exome sequencing for autism spectrum disorders, it can also be applied more generally. It not only takes into account the targeted sequence context, but also the coverage found in our studies and the number of trios sequenced. These probabilities, therefore, can be scaled and applied to other exome sequencing studies to assess the significance of potentially promising candidate genes for further evaluation.

1524T

Efficient Two-Stage Analysis Approach for Complex Trait Association with Arbitrary Depth Sequencing Data. *S. Yan, Y. Li.* University of North Carolina-Chapel Hill, Chapel Hill, NC.

Sequencing-based genetic association analysis is typically performed by first generating genotype calls from sequence data and then performing association testing on called genotypes. Accurate genotype calling requires either high sequencing depth if a small number of individuals are sequenced or if single individual calling is performed; or computationally intensive multi-sample linkage disequilibrium (LD) aware genotype calling. Here, we propose an economically and computationally efficient two-stage approach in which candidate genetic regions are screened out in the first stage via a rapid likelihood based method on sequencing data directly (without first calling genotypes) and then candidate regions are scrutinized in the second stage by performing association tests on genotypes from LD-aware multi-sample calling. Our simulations show that depth ranging from 4 to 16 benefits most from the 2-stage approach. Among those depth ranges, the overall powers are around 0.5 or more when the 0.05 bonferroni-corrected type I error is achieved with the screening ratio being 0.5 or less. For example, when sequencing depth equals to 16 and screening ratio is only 0.3, the estimated overall power is around 0.73 and the bonferroni-corrected type I error is controlled at 0.05. Genotyping calling only 30% of the markers at the second stage spares a great deal of time compared with the one step testing. We will also show the effectiveness of our two-stage approach using real sequencing data.

1525F

Improving the Accuracy and Efficiency of Sequencing Experiments, by Using Previously Sequenced Genomes. *X. Zhan, H.M. Kang, G. Abecasis.* Center for Statistical Genetics, Department of Biostatisitcs, Univeristy Michigan, 1415 Washington Heights, Ann Arbor, MI 48109.

DNA sequencing is now an indispensable technology in medical genetic research studies. Genome scale sequencing experiments are constrained by the expense required to sequence each sample to high depth, which is required to ensure accurate genotype calls when samples are examined individually. When samples are analyzed jointly, or together with many previously sequenced samples, accurate genotype can be produced by haplotype aware analyses of the short read data.

Here, we report on improvements to our software for the analysis of low pass sequencing experiments. We then use our software to call genotypes at sites with low coverage data (using the 1000 Genomes reference panel to guide haplotype aware analyses) and explore the trade-offs between sequencing depth and genotype accuracy.

We show that moderately deep 12x data allows for >99.5% accuracy at hard to call heterozygous and non-reference homozygote sites. Lower 0.5x coverage, provides 94% genotype accuracy at these hard to call sites using our methods (versus 24% accuracy when using single sample analysis methods).

We also explore the relationship between analysis method, sequencing depth and power. We expect the results are helpful in designing the sequencing experiments of the future.

1526W

Estimating and Testing Genetic Effects for Complex Traits in Sequence-Based Association Studies and Power Comparisons. *J. Zhou, N. Laird.* Harvard Sch Public Health, Boston, MA.

Genetic studies have been haunted by the mystery of “missing heritability”. Although genome-wide association studies (GWAS) have identified many variants associated with some common diseases and traits, these variants typically explain only a small fraction of the heritability. With emergence of the next generation sequencing (NGS) technology, many novel statistical methodologies have been proposed to assess the contribution of rare variations to complex disease etiology. In this paper we explore the application of linear mixed model (LMM), which has been recently used to estimate genetic variance using variants in current GWAS platform, to the estimation and testing of genetic variance based on SNP panels of NGS data. In general, through the estimation of genetic component, LMM provides an association testing strategy that can detect both rare and common variants, deal with both additive and interaction effects, handle both quantitative and dichotomous traits, and incorporate non-genetic covariates. We explore the theoretical connection between LMM and kernel-based association testing such as SKAT and carry out extensive simulation studies to evaluate the power of LMM under various distribution assumptions of traits, proportions of the causal variants and the quality control (QC) stringent thresholds. Our studies show superior performance of LMM under various conditions compared to other existing methods.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1527T

A unified statistical framework to correct for unknown population stratification, family structure and cryptic relatedness in the sequence-based association studies. *Y. Zhu[1], J. Zhao[1], Y. Shugart[2], M. Xiong[3].* 1) Biostatistics and Epidemiology, The University of Oklahoma Health Science Center, Oklahoma City, OK; 2) Unit of Statistical Genomics, Division of Intramural Division Program, National Institute of Mental Health, National Institute of Health, USA; 3) Division of Biostatistics, School of Public Health, the University of Texas Health Science Center at Houston.

In this study, we will address two issues in sequence-based association studies. The first issue is how to correct for population stratification in testing the association of rare variants. Although structured association, principal component analysis (PCA) and manifold learning can be used to correct for population stratification in the test of the association of common variants, how to correct for population structure in the test of the association of rare variants remains a challenge . The commonly used methods such as PCA to correct for population stratification in the GWAS of common variants use summary measures of ancestry as covariates to correct for population stratification. However, many existing statistical methods for testing the association of rare variants are typically incapable of incorporating covariates in the test. To overcome issues, we propose to develop unified, and computationally feasible statistical methods for sequence-based association analysis in samples with arbitrary combinations of related and unrelated individuals using the quasi-likelihood score approach. The second issue is how to use the within family information to correct for family structure and between family information to correct for cryptic relatedness and population stratification. Our approach estimates the information for both within and between family from sequence data and thus can simultaneously correct for family structure, cryptic relatedness and population stratification while maintaining high power. Consequently the proposed novel statistical methods can be applied to several study designs including (1) an isolated population in which individuals may be related, (2) admixed populations, (3) pedigrees, and (4) combination of pedigrees with unrelated individuals. To evaluate theperformance our methods under different study designs, we useintensive simulations to calculate the type 1 error rates of the proposed tests and conduct power comparisons with other existing population stratification correction methods. The proposed statistics have also been applied to the early-onset myocardial infarction (EOMI) exome sequence data from the NHLBI's Exome Sequencing Project. Our results demonstrate that the proposed methods substantially outperform other population structure correction methods.

1528F

Genome-wide association study of primary tooth eruption identifies pleiotropic loci association with craniofacial distances. *G. Fatemifar[1,2], C. Hoggart[3], L. Paternoster[2], J. P. Kemp[1,2], I. Prokopenko[4,5], M. Horikoshi[4,5], J. H. Tobias[6], S. Richmond[7], F. Geller[8], B. Feenstra[8], M. Melbye[8], T. Sørensen[9], A. Zhurov[7], A. M. Toma[7], S. Ring[1], B. S. Pourcain[2], N. J. Timpson[2], G. D. Smith[2], M. Jarvelin[3], D. M. Evans[2].* 1) Social and Community Medicine, University of Bristol, Bristol, United Kingdom; 2) MRC Centre for Causal Analysis in Translational Epidemiology (CAiTE), University of Bristol, United Kingdom; 3) Department of Epidemiology and Biostatistics, Imperial College London, St Mary's Campus, Paddington, London, United Kingdom; 4) Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford, United Kingdom; 5) Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom; 6) Academic Rheumatology, Musculoskeletal Research Unit, Avon Orthopaedic Centre, Southmead Hospital, Bristol, United Kingdom; 7) Department of Applied Clinical Research and Public Health, Cardiff University, Cardiff, United Kingdom; 8) Department of Epidemiology Research, Statens Serum Institute, Copenhagen, Denmark; 9) Institute of Preventative Medicine, Copenhagen University Hospital, Copenhagen, Denmark.

Twin and family studies indicate that the timing of primary tooth eruption is highly heritable, with estimates typically exceeding 80%. To identify variants involved in primary tooth eruption we performed a population based genome-wide association study of `age at first tooth' and `number of teeth at one year' using 5998 and 6609 individuals respectively from the Avon Longitudinal Study of Parents and Children (ALSPAC) and 5403 individuals from the 1966 Northern Finland Birth Cohort (NFBC1966). We tested 2,363,918 SNPs common to both studies. Analyses were controlled for the effect of gestational age, sex and age of measurement. Results from the two studies were combined using fixed effects inverse variance meta-analysis. We identified nine loci reaching genome-wide significance ($p<5\times10-8$) for `age at first tooth' and a further ten loci for `number of teeth at one year'. This included six previously unidentified loci (rs17563 (BMP4), rs1888693 (CACNB2), rs4937076 (CDON), rs1799922 (CALU/OPN1SW), rs997154 (JUB) and rs9316505 (DLEU7)), some of which are known to play a role in tooth and other developmental pathways. Two of these loci (rs17101923 and rs997154) showed evidence of association with craniofacial distances particularly those indexing facial width. Our results suggest that the genome-wide association approach is a powerful strategy for detecting variants involved in dentition, craniofacial growth and development.

1529W

General framework for meta-analysis for sequencing association studies. *S. Lee, X. Lin.* Biostatistics, Harvard school of public health, Boston, MA.

We propose a general statistical framework of meta-analysis for rare variants association test. To identify disease susceptibility genetic variants, meta-analysis has become an effective approach since it effectively combines data from different cohorts to increases the statistical power. In spite of the fact that it has been widely used in genome-wide association studies to identify the disease susceptible common variants, only limited effort has been made for rare variants analysis. To fill the gap, we develop a novel meta-analysis method that extends popular burden tests, sequence kernel association test (SKAT), and more recent optimal unified test (SKAT-O). The proposed method uses gene-level summary statistics to conduct association test and is flexible enough to apply for homogeneous or heterogeneous genetic effects across different cohorts. Furthermore, the proposed method is as powerful as conducting a mega-analysis with all individual level genotype data. Simulation studies and a real data analysis with Dallas Heart Study data will confirm the superior performance of the proposed method.

1530T

General class of family-based association tests for sequence data, and comparisons with population-based association tests. *I. Ionita-Laza[1], S. Lee[2], V. Makarov[3], J.D. Buxbaum[3,4,5], X. Lin[2].* 1) Department of Biostatistics, Columbia University, New York, NY; 2) Department of Biostatistics, Harvard University, Boston, MA 02115; 3) Seaver Autism Center for Research and Treatment, Mount Sinai School of Medicine, New York, NY 10029; 4) Department of Psychiatry, Mount Sinai School of Medicine, New York, NY 10029; 5) Genetics and Genomic Sciences, Mount Sinai School of Medicine, New York, NY 10029.

Recent advances in high-throughput sequencing technologies make it increasingly more efficient to sequence large cohorts in many complex traits. Making sense of these massive data depends critically on the use of efficient study designs and appropriate statistical methods. We discuss a general class of sequence-based association tests for family-based designs, that corresponds naturally to previously proposed population-based tests, including the classical burden and variance-component tests. This framework allows for a direct comparison between the power of sequence-based association tests for family-based, and population-based designs. We show that for dichotomous traits, family-based designs result in similar power levels as the population-based designs (although at increased sequencing costs for the family-based design), while for continuous traits (in random samples, no ascertainment) the population-based designs can be substantially more powerful. A possible disadvantage of population-based designs is that they can lead to increased false-positive rates in the presence of population stratification, and although adjustment for population substructure using principal component analysis can properly adjust for stratification at a small loss in statistical power, in more subtle scenarios, it can fail to completely adjust for such confounding. This is unlike the family-based designs, which are, by definition, robust to population stratification. We also present results from an application to an exome-sequencing family-based study on autism spectrum disorders.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1531F

A novel statistical approach to prioritize variants from deeply sequenced NGS samples using publicly available sequence controls. *L. J. Strug[1,2], A. Derkach[2], T. Chiang[1], L. Addis[3], S. Dobbins[4], I. Tomlinson[5], R. Houlston[4], D. K. Pal[3].* 1) Program in Child Health Evaluative Sciences, the Hospital Sick Children, Toronto, ON, Canada; 2) University of Toronto, Toronto, ON, Canada; 3) Department of Clinical Neuroscience, Institute of Psychiatry, Kings College London, UK; 4) Institute of Cancer Research, London, UK; 5) Wellcome Trust Centre for Human Genetics, Oxford, UK.

Next generation sequencing (NGS) is a powerful tool to identify variants responsible for rare Mendelian disorders. Putative causal variants can be prioritized for follow-up by their novelty and known/predicted functional consequence. However, variant prioritization of NGS data for complex disorders is more challenging: in this situation contributory variants are expected to have an appreciable frequency in healthy subjects. A sufficiently powered case-control design is prohibitively expensive. Population whole genome data is publically available, but its generally low read-depth introduces a bias that complicates comparison with higher read-depth NGS data from cases. We therefore developed novel statistical methodology that adjusts for bias and provide a general framework to use publically available sequence data to prioritize variants in studies of complex disorders. We apply this method to NGS data from a 600kb linkage region for an epilepsy endophenotype present in ~2% of the general population. With DeCode genetics we used long-range PCR and NGS on 27 epilepsy cases, with average read-depth of 197x. We compare the 27 cases with 178 1000 genomes controls (release May 21st 2011), and validate the results by comparison with 200 NGS controls sequenced by Complete Genomics (CG) with ~35x depth. Genotype quality filters were common across the epilepsy and CG datasets; in the 1000 genomes data we used the low coverage calls as released. Using variants seen in both case and control datasets (novel functional variants are prioritized for follow-up by default), we developed a conditional exact score statistic that calculates the probability of observing 2×3 contingency tables with statistics as extreme or more so under the selection scheme. Using this statistic we compared our cases to 1000 genomes, our cases to CG controls, and the control groups to each other. Results using 1000 genomes highlight the need to adjust the statistic for NGS differences, while comparing CG controls to cases shows an absence of bias requiring adjustment. We show that an average genomic control λ=1.65, calculated from re-sampling 1000 sets of variants with pairwise LD<20%, can remove bias and identify the same associated variants as with the CG high-depth control group. These variants can then be prioritized by association and annotation and added to a prioritization list for follow-up, consequently increasing the utility of public NGS controls.

1532W

Variable Selection Based Weighting Schemes for Implicating Rare Variants in Sequence Data. *A.E. Byrnes[1], M. Li[2], M.C. Wu[1], F.A. Wright[1], Y. Li[1,3].* 1) Biostatistics, University of North Carolina, Chapel Hill, NC; 2) Biostatistics and Epidemiology, University of Pennsylvania School of Medicine, Philadelphia, PA; 3) Genetics, University of North Carolina School of Medicine, Chapel Hill, NC.

The improvements in and dropping costs of next generation sequencing technologies allow us to access rare genetic variants (RVs) in increasingly large samples of individuals. This in turn augments our ability to find associations between RVs and human complex traits. Even so, power to detect associations between individual RVs and complex traits is limited and this has motivated the development of burden tests wherein a trait is regressed on the weighted average of multiple RVs across a region. Several such schemes have been proposed, but there is little consensus on which is the best. Here, we compare several weighting schemes using extensive simulations that mimic large-scale sequencing studies of a quantitative trait under a variety of situations. We simulate data in which there are single and multiple truly causal RVs as well as cases where the causal RVs contribute to the quantitative trait in the same and opposite directions. We investigate several existing phenotype-independent and -dependent weighting schemes and introduce variable selection to some of the phenotype-dependent schemes in order to increase the number of zero weights. We assume that most RVs are not associated with the trait under study and, in the hopes of clarifying the association signal, use variable selection to eliminate the weights of the less informative markers by setting their estimates to exactly zero. Our simulations show that phenotype-dependent weighting schemes have increased power over phenotype-independent schemes in most cases. This stands to reason since the phenotype-dependent schemes take the outcome information into account whereas phenotype-independent schemes do not. Further, we show that variable selection can significantly increase the power to detect the effect of multiple RVs in phenotype-dependent schemes since it allows us to select only the most informative variants and to assign non-zero weights preferentially to these. In the majority of the situations we considered with multiple truly causal rare variants, we observed power near or above 80%. Results on a real data set of ~2000 sequenced individuals are also shown.

1533T

Accurate local ancestry inference in exome sequence samples. *Y. Hu, G. Abecasis, H. Kang.* Department of Biostatistics, University of Michigan, Ann Arbor, MI.

Deciphering the ancestry of specific chromosomal regions in admixed individuals is significant for studies of human evolutionary history and disease gene mapping. Typically, this ancestry deconvolution relies on genotypes generated using high-density GWAS arrays; data that is not always available in exome sequenced samples. We show that even a relatively small number of off-target reads from exome sequencing experiments can be used to accurately estimate the ancestry of admixed individuals.

To reconstruct local ancestry, we use a Hidden Markov Model that models sequence data directly, without requiring intermediate genotype calls. We evaluate the accuracy of our method through simulation and the analysis of admixed individuals sequenced by the 1,000 Genomes Project and ongoing disease studies. In African-Americans, we show local ancestry estimates derived using our method are extremely similar to those derived using Illumina's Omni 2.5M genotyping array (average r2 between the two estimates is 0.99); and much improved in relation to estimates that rely use only exome genotypes and ignore off-target sequencing reads.

In 2000 African American individuals, we use our method to evaluate patterns of coding variation in portions of the genome with recent European and African ancestry. We estimate that ~20% of the genome of each individual (range 2–53%) is of European ancestry and that the heterozygosity rate is lower in portions of the exome with recent European ancestry (0.72 per kilboase of coding sequence, compare to 0.87 for African ancestry). Among the heterozygous sites, the ratio of the number of nonsynonymous to synonymous sites is higher in European portions (0.95 versus 0.85). In addition, the number of nonsynonymous and loss of function heterozygous sites per kilobase are higher in portions with European ancestry (195 non-synonymous variants and 10 loss-of-function variants per megabase of coding sequence, compared to 191 non-synonymous and 9 loss-of-function variants per megabase in portions of the genome with recent African ancestry).

Our method should be useful to anyone undertaking exome or targeted sequencing of admixed populations, both for analysis of human population history and for disease gene mapping studies.

1534F

Internal reference panel selection methods for genotype imputation: a novel sequencing study design strategy for genotyped samples. *P. Zhang[1], N. A Rosenberg[2], S. Zöllner[3].* 1) Dept Computational Medicine and Bioinformatics, Univ Michigan, Ann Arbor, MI; 2) Department of Biology, Stanford University, Stanford, CA; 3) Dept Biostatistics, Dept Psychiatry, Univ Michigan, Ann Arbor, MI.

The recent dramatic cost reduction of next-generation sequencing technology enables investigators to identify risk variants for complex diseases with sequencing studies. However, sequencing all individuals in a study is still prohibitively expensive for large sample sizes. For study samples with existing genotype data, such as data from genome-wide association studies (GWAS), a cost-effective approach is to sequence a subset, and use the sequenced subset as an internal reference panel to impute the rest of the study sample. As of today, there are thousands of existing GWAS with a total of millions of individuals involved. Effective use of these genotype data when planning sequencing studies on these individuals make it possible to carry out large-scale sequencing studies with limited cost. We have focused on how to select an internal reference panel in such studies on the basis of the existing genotype data. Using an internal reference panel precludes the possibility of a substantial mismatch of ancestry background between the study population and the reference population. It also allows the variants that are unique to the study sample to be imputed. The goal of this study is to find the optimal internal reference panel for imputing the sequences of the study sample so that we can (1) maximize the number of polymorphic sites in the study template and (2) achieve the optimal imputation accuracy. We propose two strategies for choosing optimal internal reference panels: first, the “most diverse subset,” defined as the subset with maximal phylogenetic diversity, thereby incorporating individuals that span a diverse range of genotypes within the sample. Second, the “most representative subset,” defined as the subset that has the minimum genetic distances to unselected individuals, such that each unselected individual is represented well by at least one individual in the selected subset. Using data both from simulations and from sequences from the 1000 Genomes Project, we show that carefully calibrating the internal reference panel can significantly improve the imputation accuracy, especially for the imputation of rare variants. We further compare the performances of the two panels in genotype imputation for data generated under different scenarios. We seek to provide guidelines for selecting optimal subsets, thereby recommending a novel strategy to obtain sequence information comparable to that available by sequencing the entire study sample.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1535W

Statistical methods for inference in population genetic studies with uncertain genotype data. *E. Han.* Biostatistics, UCLA, Los Angeles, CA.

Next-generation sequencing (NGS) techniques are opening a new era in population genetic studies. However NGS data suffers from errors in read-level base calling, errors in read-mapping, and variable coverage, and these create non-negligible uncertainty in genotype calls in many datasets. Ignoring genotype uncertainty in population genetic inference can lead to serious biases in estimation and mistaken conclusions in subsequent analysis, such as in genome-wide selection scans or studies of population structure. Here we present a statistical approach to infer population genetic properties of the sample such as the number of segregating sites (S), the nucleotide diversity (π) and the population structure fixation index (F_{ST}). Our method computes the posterior probability of the allele frequency at each site, and summary statistics are then calculated by summing over these posterior probabilities. We present software to compute these metrics, compare the behavior of the method to naïve approaches that ignore genotype uncertainty, and explore the properties of these estimators as a function of coverage and base qualities. We have applied the method to next-generation sequencing data from six canid genomes (two dogs, three wolves, and one golden jackal) to detect signatures of adaptation during dog domestication.

1536T

On association analysis of rare variants under population-substructure: An approach for the detection of subjects that can cause bias in the analysis. *D. Qiao[1], M. Mattheisen[2], C. Lange[1,3,4].* 1) Biostatistics, Harvard School of Public Health, Boston, MA; 2) Department of Medicine, Brigham and Women's Hospital, Boston, MA; 3) Institute for Genomic Mathematics, University of Bonn, Bonn, Germany; 4) German Center for Neurodegenerative Diseases (DZNE), Bonn, Germany.

For the analysis of rare variant data in population-based designs, we propose a method to detect study subjects that may create population substructure in the study sample, potentially introducing bias into the association analysis. Our approach is computationally fast and simple, permitting applications to whole-genome sequencing studies. The approach does not require variant data to be in linkage equilibrium and can be applied to all the genetic loci that are available in the study. For both rare and common variants, we assess the performance of our approach by application to HapMap 3 data, 1000 Genome Project phase I data and simulated datasets. The results are compared to those from Principle Component Analysis (PCA). The statistical power of both approaches to detect outliers based on common variant data and based on rare variant data are comparable in most of the scenarios, but the power of PCA in detecting outliers is relatively small in the presence of LD as shown in the simulations. Also, the observed number of false-positive errors appears to be different for the two approaches as the outlier detection algorithm using PCA cannot adjust for the number of subjects included in the dataset. The results of the HapMap analysis and the simulation studies suggest that our approach maintains the type I error. Taking additionally into account the minimal computational requirements of our approach and the ability to incorporate all marker information, the proposed method will have practical utilization for sequencing studies and genome-wide association studies.

1537F

Identity-by-descent analysis of sequence data. *S.M. Smith[1], S.R. Browning[1], B.L. Browning[2].* 1) Biostatistics, University of Washington, Seattle, WA; 2) Medicine, Div of Medical Genetics, University of Washington, Seattle, WA.

Individuals are identical by descent if they share a haplotype by inheritance from a recent common ancestor. A shared haplotype with length greater than 2 centimorgans corresponds to a common ancestor within approximately the past 25 generations, and these long shared haplotypes can be detected with high power using existing methods for SNP array data.

Identity by descent (IBD) detection methods can also be applied to sequence data, and the detected IBD can be used to phase rare variants and to detect genotype errors. In theory, the complete marker coverage in sequence data will maximize power to detect IBD. However in practice, alignment and sequencing artifacts, reduced phase accuracy at low frequency markers, and recent mutations pose challenges to IBD detection in sequence data. We will present and evaluate several methods for improving IBD detection in sequence data using 1000 Genomes Project and simulated data.

1538W

Robust similarity regression for population substructure in rare variant aggregation analyses. *J. Tzeng[1,2], C. Smith[2].* 1) Department of Statistics, North Carolina State University, Raleigh, NC., Select a Country; 2) Bioinformatics Research Center, North Carolina State University, Raleigh NC.

Population substructure, which included population stratification and cryptic relatedness as two extreme manifestations, is a major confounder in genetic association studies. Recent studies found that rare variants are more geographically localized or private to specific populations. It is indicated that the impact of population substructure is more complex and stronger on rare variants than common variants. Furthermore, the number of variants collapsed also impact the inflation level of the test statistics. When substructure is caused by discrete clusters (e.g., continental-wise stratifications), stratified permutation or incorporating principal component covariates can be effective at controlling the confounding effect. However, for regional and complex substructure, existing correction methods for single marker analysis may fail to control for confounding effects. Using the framework of similarity regression, we propose to model the whole genome sharing at the rare variants loci, a measure that is suggested to reflect the local substructure. By modeling average allele sharing level across genome, it corrects the non-zero baseline sharing induced by population substructure and explicitly controls the substructure effect on the phenotype. The proposed method is robust to a wide range of substructure induced by genealogy, from continual to regional, and from population stratification to cryptic relatedness. We evaluate the performance of the proposed framework based on grid genotype simulation, and demonstrate the robustness, validity and utility of the proposed approaches.

1539T

Detecting differential expression in splice variants in the absence of annotated isoforms using RNAseq. *N.J.I. Lewin-Koh, T. Bhangale, M. Huntley, J. Kaminker, F. Cai, M. van der Brug.* Genentech Inc, South San Francisco, CA.

Differences in isoform expression, inherited or spontaneous, lead to differences in phenotype or disease susceptibility. Identifying these differences enables us to find genes associated with disease and shed light on the underlying biological mechanisms. Current technologies allow sequencing of transcriptomes (RNAseq) and expression quantification of known isoforms. To our knowledge, methods for RNAseq do not explicitly try to quantify differential isoform expression, except in cases where a known alternate splicing event is being tested. However, evidence suggest that the differences can be subtle involving multiple events in a gene, each with a small effect on phenotype. Moreover, available methods are not applicable to novel/unknown splicing events. We present a method based on kernel regression that addresses these problems.

Our method detects overall splice-variant-level differences using RNAseq data. The method uses Kernel regression. A kernel is a matrix of pairwise functions that produces a positive semidefinite matrix of pairwise dissimilarities between the samples. Under squared error loss, kernel regression reduces to a mixed effects model, and can be fit with standard software. We use the counts of reads that map to exon junctions as random effects, modified by a kernel e. g. for a case control study, we fit a logistic mixed effects model and using a score test, assess the over/under dispersion of the coefficients from the random effects. We explore two kernels: 1. a linear kernel where all junction combinations are weighted equally, and 2. a crude distance kernel, where exon skipping is penalized by the distance in exons skipped. We compare our models among each other and against a DEXSeq model using simulation and two RNAseq data sets: 1. cancer data set with tumor and matched normal tissue within patient and 2. brain tissue datasets from Parkinson's disease patients and matched controls.

In preliminary tests with cancer data our method detected differential isoform expression in known cancer genes. Simulation and theoretical results in other contexts have shown kernel regression to be quite powerful in situations where data is sparse. The advantage is isoforms do not need to be pre-specified as in existing methods such as the TopHat/Cufflinks workflow or MISO (Mixture of Isoforms), where isoform discovery is a separate step in the pipeline.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1540F

Identification of new genetics variants for T2D in WTCCC data using genome-wide interaction analysis followed by core SNP and gene determination. *Z.-X. Zhu[1,2], Z.-H. Zhu[1], X. Tong[1], M. Liang[1], W.-C. Cui[2], M. Yang[3], L. Zhou[3], M.D. Li[2,4], J. Zhu[1].* 1) Institute of Bioinformstics, Zhejiang University, Hangzhou, Zhejiang, China; 2) State Key Laboratory for Diagnosis and Treatment of Infection Diseases, First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China; 3) Microsoft Research Asia, Beijing, China; 4) Department of Psychiatry and Neurobehavioral Sciences, University of Virginia, Charlottesville, VA, USA.

Although a significant number of genome-wide association studies have been reported for different diseases, only a relatively small number of vulnerability loci have been replicated for each disorder, and a large proportion of the loci for most complex diseases has yet to be uncovered. It has been suggested that ignorance of gene-gene interactions might have contributed greatly to such a low replication rate. Current interest and effort are focused on detecting gene-gene interactions that might fail to be identified using traditional single-locus tests. Previously, we reported a model-free Generalized Multifactor Dimensionality Reduction (GMDR) approach for detecting SNP-SNP interactions in both dichotomous and quantitative phenotypes while allowing covariate justification and handling of unbalanced data. However, the computational burden and less efficient implementation of the original programs make them impossible to use for genome-wide association studies (GWAS). In this study, we developed a graphics processing unit (GPU)-based GWAS-GMDR program, which is able not only to analyze GWAS data but also to run much faster than the earlier version of the GMDR program. We then used the GWAS-GMDR software to analyze a publicly available genome-wide dataset on type 2 diabetes (T2D) from the Wellcome Trust Case Control Consortium. Through an exhaustive search of pairwise interactions and a selected search of three- to five-way interactions conditioned on significant pair-wise results, we identified 24 core single nucleotide polymorphisms (SNPs) in 6 genes that appear to be important for T2D. Of these core SNPs, 11 have been reported to be associated with T2D, obesity, or both, providing an independent replication of previously reported SNPs. More importantly, we identified three new susceptibility genes for T2D. To make this program more useful, versions for Windows, Mac OS X, and Linux systems have been developed for free distribution to the scientific community.

1541W

Imputing Genotypes in Large Pedigrees: A Comparison between GIGI and BEAGLE. *C.Y.K. Cheung, E. Wijsman.* Dept Biostatistics, Univ Washington, Seattle, WA.

Imputation of dense genotypes facilitates identification of rare risk variants. GIGI (Genotype Imputation Given Inheritance) is a new genotype imputation approach that can handle large pedigrees. While GIGI uses correlation from identity-by-descent in pedigrees, BEAGLE uses correlation from linkage disequilibrium in the population. Here we used a 95-member pedigree to compare the imputation quality from GIGI and BEAGLE when BEAGLE was used while ignoring the pedigree structure. In a 50 cM region, 60 subjects were observed for 323 SNPs. In an initial analysis, we kept complete data on only 13 subjects and masked all but 35 evenly spaced SNPs on 47 other subjects. Then, we imputed missing genotypes on 288 other SNPs using GIGI and BEAGLE. In BEAGLE, we also evaluated the effects of adding 202 outside reference subjects from 3 other pedigrees of similar ethnic background, as well as providing denser SNPs in the reference panel in a “Leave-One-Out” analysis. In the “Leave-One-Out” analysis, we sequentially omitted one SNP from the 47 subjects who received incomplete data, imputed missing genotypes of this omitted SNP, and repeated for each SNP. We called genotypes using the most probable configuration and compared the accuracy of imputation in all genotypes and in a subgroup considering only genotypes having at least 1 rare allele from SNPs with Minor Allele Frequency of less than 0.05. In calling the overall genotypes, GIGI was more accurate than BEAGLE under the initial setup (79.7% vs. 70.2%) but was less accurate when the outside reference subjects and denser SNPs were included (79.7% vs. 95.4%). However, in calling rare genotypes, GIGI was more accurate under all settings (62.9% vs. 4.5–28.1%) and also detected an extra 46.2% of the rare alleles in the 35 completely un-typed relatives, which BEAGLE could not do because it could not impute genotypes on completely unobserved subjects. These results demonstrate advantages in use of large pedigrees for identifying rare risk variants through explicitly modeling the co-transmission of alleles to multiple affected individuals. Supported by NIH GM046255.

1542T

A Hidden Markov Model for Coalescent-Based Mapping of Complex Trait Loci from Sequencing Data in Large-Scale Case-Control Studies. *Z. Geng[1], P. Scheet[3], S. Zoellner[1, 2].* 1) Dept Biostatistics, Univ Michigan, Ann Arbor, MI; 2) Dept Psychiatry, Univ Michigan, Ann Arbor, MI; 3) Dept Epidemology, UT MD Anderson Cancer Center, Houston, TX.

Association mapping based on linkage disequilibrium (LD) is widely used to localize positions of disease variants. However, due to the complicated genetic dependence structure, apply LD mapping to complex diseases is challenging. Potentially, with the hope of modeling the evolutionary process that produces our sequencing data, coalescent-based approaches may extract more information to improve our mapping. Such methods, by tracing all sampled alleles to a single ancestral copy known as the most recent common ancestor (MRCA), provide the genealogy at all sites in the region we have sequenced. Hence, we can model the probability of carrying risk variants at all loci jointly, and obtain confidence intervals where true risk variants are most likely to occur. Such methods may provide extra benefits when we have low coverage sequences — its Bayesian framework is more tolerant to each single site's being at low quality. However, existing coalescent-based methods typically suffer from two major challenges: correctly modeling LD, and efficiently scaling to large sample sizes. Indeed, we propose a novel approach which may overcome the two difficulties simultaneously. First, we use hidden Markov models (HMM) to improve LD modeling. Such careful modeling of LD allows us to integrate over phase uncertainty such that our results do not depend on pre-phase accuracy. In simulation studies, our HMM-enhanced method shows better specificity to avoid false positives caused by incorrectly modeling the dependence structure of markers. Second, we improve our model efficiency to handle large-scale sequencing data. Since the space of all coalescent histories is of high dimension, sampling from this space is computationally extremely intensive. Instead, we first identify several clusters of closely related haplotypes from the present sequence data and estimate the ancestral sequence of each cluster. As a result, we can ignore the genealogy within each cluster, and the sampling space reduces to the pre-determined number of ancestral sequences. In the end, we are able to model phenotypes and localize risk variants, based on both the cluster pattern of our sample and the coalescent history between ancient clusters. In summary, we have developed a novel approach to estimate the coalescent genealogy throughout sequenced regions. In fine mapping of complex trait loci, our method provides higher specificity, and is applicable for large-scale case-control studies using sequencing data.

1543F

Genetic Simulation Resources (GSR): A website for the registration and discovery of genetic data simulators. *B. Peng[1], B. Racine[2], H. Chen[3], L. Mechanic[3], L. Clarke[2], E. Gillanders[3], E. Feuer[3].* 1) Dept of Genetics, Univ Texas MD Anderson CA Ctr, Houston, TX; 2) Cornerstone Systems NorthWest Inc; 3) Division of Cancer Control and Population Sciences, National Cancer Institute.

Many genetic simulation programs have been developed to simulate evolutionary processes under realistic ecological and genetic scenarios and generate genetic data resulting from such processes. Such simulations have been used, for example, to predict properties of populations retrospectively or prospectively according to mathematically intractable genetic models, and to assist the validation, statistical inference and power analysis of a variety of statistical models. In a 2010 NIH meeting (Mechanics, et al. 2012), a catalogue of simulation programs was identified as a tool which could help advance the development of genetic epidemiology analytic tools. However, due to the differences in type of genetic data of interest, simulation methods, evolutionary features, input and output formats, terminologies and assumptions for different applications, choosing the right tool for a particular study can be a troublesome process that usually involves searching, downloading and testing many different tools. Here we present Genetic Simulation Resources (GSR) (http://popmodel.nci.nih.gov/geneticsimulation), a web service provided by the National Cancer Institute (NCI) that aims to help researchers compare and choose the right simulation tools for their studies. This service allows authors of simulation software to register their applications and describe them with well-defined attributes, and users to search and compare simulators according to specified features. Activities such as citations, updates from authors, number of visitors, and user feedbacks from a planned extension to GSR will be tracked and provide updated information about the status of the applications to help users weed through a large number applications to select the most appropriate program to support the research questions.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1544W

Rare-variant tests in stratified populations. *M. Schmidt, D. Kinnamon, E. Martin.* Hussman Inst Human Genomic, Univ Miami, Miami, FL.

Population stratification is well-known to cause spurious associations with individual common variants in case-control studies. In sequence data with rare variants, this effect may be particularly magnified by gene-based tests that combine information across multiple variants due to the accumulation of population-specific variants. We recently suggested that gene-based tests that combine non-negative single-variant test statistics using permutation inference provide robust and efficient alternatives to approaches that combine minor allele counts or that require complete genotype data (Kinnamon et al. PLoS One, 2012). Though originally proposed for homogeneous populations, our approach can be modified to incorporate information about population structure with an appropriate choice of test statistic and permutation procedure. Specifically, we have implemented a modified approach for gene-based association testing in unrelated cases and controls that first assigns individuals to a specific subpopulation based on genotypes from a set of genome-wide markers using a clustering algorithm such as Ward's method. These clusters are then used as strata in calculating a generalized Cochran-Mantel-Haenszel chi-square statistic using all available genotype data for each variant within the gene. Like the original approach, the statistics can be combined across the gene by taking either the maximum statistic or sum of statistics. A p-value for the combined statistic is obtained by permuting affection status within each cluster, which maintains between-marker correlations and does not rely on asymptotic approximations. Using computer simulations, we demonstrate that the approach maintains the correct type I error in stratified populations. We compare the approach to the alternative Sequence Kernel Association Test (Wu et al. AJHG, 2010) in the presence of varying levels of stratification and missing data. Our results suggest that the modified approach provides a valid and robust method for analysis of rare variants in sequence data in the presence of population stratification. The method is implemented in a C++ computer program RVStrat that takes advantage of multi-processor machines.

1545T

Software pipeline to detect SNPs and call their genotypes. *M. Trost, H.M. Kang, G. Jun, G.R. Abecasis.* Center for Statistical Genetics, Department of Biostatistics, University of Michigan, Ann Arbor, MI.

As the volume of next-generation sequencing data increases, the need for high quality, high performance data processing pipelines that can convert short read sequence data into genotypes becomes more pressing. Converting short read sequence data into high-quality genotypes is a multi-step process that requires alignment, alignment refinement, identification and filtering of variant calls across many samples, and haplotype aware genotype calling. Each of these steps can require multiple compute and disk intensive tasks, so that analysis of even a few hundred samples can require many terabytes of intermediate storage, months of compute time and coordination of thousands of computing jobs.

Our UMAKE pipeline automates variant calling for whole genome or targeted exome sequencing projects including many samples. The process of converting aligned sequence data (typically stored in a BAM file) into a list of variant sites and genotypes (typically stored in a VCF file) is decomposed into many small tasks which are then executed in a coordinated fashion. UMAKE is configurable, so that alternate analysis tools that use standard formats can be used as replacements for our default set of components. In addition, UMAKE is designed to allow incremental analysis where a project is gradually updated to incorporate new samples.

UMAKE is currently used for variant calling in a variety of projects, ranging from the 1000 Genomes Project to the NHLBI Exome Sequencing Project, and totaling >15,000 whole genome or whole exome sequenced samples and we provide examples of the computing requirements and call quality for several of these projects. UMAKE and related documentation are freely available at: http://genome.sph.umich.edu/wiki/UMAKE.

1546F

Bidimensional scale analysis: A novel approach to genographic analysis. *Q. Huang, Y. Wu, Z. Li, X. Liu, W. Xie.* Instutite of Life Sciences, Nanjing, Jiangsu, China.

Human leukocyte antigen (HLA) is the most complex and highest polymorphic genetic system and is one of the first genetic markers used in the study of population relationship. Many studies had shown diversity and similarity of global populations using phylogenetic tree and multidimensional scale (MDS) analysis, but traditional MDS analysis only presents clustered information, without showing more genographic information. Here we report a novel bidimensional scale (BDS) analysis for genographic analysis to study difference of population's geographical distribution. Method: Datasets were compiled from published studies from 1990 to 2012 and the International Histocompatibility Workshops, containing locus A, B and DRB1 high-resolution allele frequency. These data represent approximately 339,900 chromatids and 95 populations from five continents (Africa, Europe, Asia, Australia and America) and Pacific islands. The reconstructed datasets of allele frequency were comparable after correcting variety of errors in original data, normalizing different statistical methods, and redesigning ambiguous alleles. Phylogenetic trees of HLA-A, -B, -DRB1 and -ABD (combined by locus A, B, and DRB1) were constructed by four commonly used genetic distance algorithms. The distances computed by these algorithms were analyzed by the BDS algorithms, and the results were dot-plotted in two-dimensional figure. Results: Most of studied populations in each continent were clustered together according to their geographical distribution in each phylogenetic tree, but some neighbor populations were separated by other populations. The location of populations in dot-plotted figure of BDS was basically consistent with that of populations sampled from. African, European, Amerindian and native Australian were located at four quadrants. East Asian was located in middle of the dot-plotted figure. Moreover, North African was plotted between European and South African, West Asian was among European, North African and East Asian. Overall, the figure of BDS plotted by worldwide population resembled the sampling map. The figure plotted by novel BDS method contained directed information of geographical distribution of the population. Conclusion: Our results suggest that the BDS analysis method proved to generate more genographic significance than the traditional MDS analysis, and it can be broadly used in the field of genographic analysis.

1547W

Genotype probability distributions on complex pedigrees, with applications to relationship testing. *M. Vigeland[1], T. Egeland[2].* 1) Dept of Medical genetics, Oslo University Hospital, P.O. Box 4956 Nydalen, 0424 Oslo, Norway; 2) Dept of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, P.O. Box 5003, NO-1432 Aas, Norway.

This work is motivated by power considerations in connection with relationship estimation applications. For example, paternity cases are well studied in forensics but are not straightforward if the alleged father is related to the mother. Other areas, including linkage and association studies, also provide examples where relationship testing is important. We focus on cases where all or most individuals have not yet been typed and aim at calculating the prior probability that a specific relationship will be rejected or excluded. We present software (implemented as an R package) for computing joint genotype distributions for any set of pedigree members. The program handles a wide range of pedigrees, including complex pedigrees with inbreeding. Furthermore, it is possible to condition on known genotypes, including partial genotypes with one missing allele. In addition to exact probabilities, the program also provides fast simulation of marker genotypes, again with the possibility of conditioning on known alleles/genotypes. These simulations are useful in cases with many alleles and many unknown individuals (making exact distributions unfeasible). Furthermore, simulations are used to compute confidence or prediction intervals in our applications. As an application we show how to measure the ability of a battery of tests to detect inconsistencies in claimed family relationships. We present examples for autosomal as well as X-linked markers and some formulae to validate the results. Calculations for a previously reported immigration case are extended to account for possible inbreeding and cases where some genotypes are known.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1548T

Trait specific genetic relatedness matrices for heritability estimates using linear mixed models. *J. Mefford, J. Witte.* Department of Epidemiology and Biostatistics, UCSF, San Francisco, CA.

Trait heritability estimated from genome-wide SNP datasets using the linear mixed model (LMM) framework estimates the variance in the phenotype of interest that is explained by causal variants or quantitative trait loci (QTLs). Genetic data is introduced into the LMM estimators through a genetic relatedness matrix (GRM) that represents how the individuals in the dataset are related (correlated) at the causal genetic variants for the phenotype. Since the causal variants are not known, a GRM can be estimated using genome-wide SNP observations, relying on the correlation of the genotyped SNPs with the QTLs due to linkage disequilibrium (LD). In the development of the GCTA software for heritability estimation using the LMM framework [Yang et al (2010) Nature Genetics], the developers run simulations to explore the estimation of a GRM using tagging SNPs having incomplete LD with causal variants. Their simulations explore different numbers of causal and tagging SNPs for the calculation of ideal GRMs (estimated from causal SNPs) and approximating GRMs (estimated from tagging SNPs). Those simulations had thousands of QTLs randomly distributed on the genome. Our study further explores the estimation of GRMs. In particular, we try to address the following questions: (1) How is the estimation of a GRM affected by the inclusion of SNPs that are not in LD with QTLs? (2) Can we improve the approximation to an ideal GRM by developing trait-specific approximations to the GRM? In our simulations, causal variants are concentrated in particular chromosomal regions rather than randomly spread across the genome. The greater the fraction of the genome that is unlinked to the causal SNPs, the poorer the estimates of the GRM using whole genome data. Trait-specific GRM estimates are constructed by several methods. SNPs are selected for inclusion in the estimate of the GRM or their contributions are weighted by using the results of single-SNP tests for association with the phenotype. Alternatively, multi-SNP tests (minP, SKAT) are used to identify genes and regions that are associated with the phenotype, and all SNPs within those regions are uses for estimation of the GRM. There is a tradeoff between improvements to the GRM by excluding parts of the genome that are unrelated to the QTL, and damage to the GRM by rejection of relevant SNPs and genes in the GRM construction. Resampling methods are used to assess this tradeoff.

1549F

The confounding effect of cryptic relatedness on cohort studies for environmental risks. *K. SHIBATA, G. Tamiya, M. Ueki, T. Nakamura, H. Narimatsu, I. Kubota, Y. Ueno, T. Kato, H. Yamashita, A. Fukao, T. Kayama, Yamagata University Genomic Cohort Consortium.* Yamagata University, Yamagata, Japan.

In general, cryptic relatedness in epidemiological association studies, have been disregarded. However, cryptic relatedness might be cofounding environmental factors. Especially, in a regional cohort studies, the effect of cryptic relatedness could be large. Here, we investigated the confounding effect of cryptic relatedness on the association between phenotype and environmental risk factors. We selected blood pressure as the quantitative phenotypes and body mass index (BMI), smoking, and alcohol consumption as environmental risk factors. In practice, using the genome-wide 657,366 SNP data in 1,622 subjects in Yamagata, we detected 20–30% cryptic relatedness of epidemiological sampling by software program PLINK. We compared the results of association tests between hypertensive disease groups with and without cryptic relatedness and detected significantly differences. In addition, we show that genetic adjustments can be useful to complement confounding environmental risk factors. Further, we examine mixture model to detect association in the case of a population with confounding effect of cryptic relatedness. We suggest that haplotypic variance contrast measure is informative to detect association in the case of cohort studies with confounding effect of cryptic relatedness.

1550W

Genome wide mutation-rate map for the analysis of recurrent de-novo mutations. *P. Polak[1,2], S.R. Sunyaev[1,2].* 1) Brigham and Women's Hospital and Harvard Medical School, Boston, MA; 2) Broad Institute, Cambridge, MA.

The fast advances in sequencing technology enable identification of de novo mutations in human pedigrees. This opens the perspective to analyze the contribution of de novo mutations to sporadic cases of oligogenic and polygenic diseases. The main challenge in the analysis is to pinpoint individual genes harboring de novo mutations causally involved in the phenotype. This is achieved by finding genes with recurrent de novo mutations. The analysis is greatly complicated by the heterogeneity of mutation rate along the human genome. Genes exhibiting higher rates of spontaneous mutations would be falsely associated with phenotypes. Currently, mutation rate heterogeneity is being ignored. Comparative genomics has identified mutation rate heterogeneity at different scales. It was shown that this heterogeneity is partially explained by regional differences in DNA methylation and replication timing. We propose a new statistical approach to gene mapping based on recurrent de novo mutations. Our method appropriately takes into account mutation rate variation along the human genome. Mutation rate variation model incorporates sequence context and uses comparative genomics and functional genomics data to estimate gene-specific mutation rate. This model allowed us to estimate expected rates of de novo mutations per gene in all functional categories (e.g. missense and nonsense mutations). Our method ranks genes harboring de novo mutations with the help of Z-scores relative to expectation.

1551T

JBASE: A Bayesian Mixture Model for Joint Analysis of Sub-phenotypes and Epistasis. *R. Colak[1,2,3], T. Kim[1,2], H. Kazan[1,2], P.M. Kim[1,2], A. Goldenberg[1,2,3].* 1) Department of Computer Science, University of Toronto, Toronto, Ontario, Canada; 2) Donnelly Center for Cellular and Bimolecular Research, University of Toronto, Toronto, Ontario, Canada; 3) Genetics & Genome Biology Program, The Hospital for Sick Children (SickKids), Toronto, Ontario, Canada.

Genome Wide Association Studies (GWAS) have emerged as the preeminent tool for mapping variants that give rise to a specific disease/trait. So far, more than >1200 loci harboring variants associated with >160 diseases/traits have been discovered, yet the explained phenotypic variability is still far from the theoretical inheritability estimates [1]. This so called missing heritability problem presents a big challenge for personalized medicine and is at the center of extensive discussion in the statistical genetics community. Among the many potential reasons for missing heritability are two major ones: First, there are interactions between markers, a phenomenon known as epistasis, which has been addressed in a number of newer approaches [2]. Second, a factor that is not commonly addressed but potentially even more critical is the heterogeneity in the phenotypic landscape of a disease, addressed via sub-phenotyping [3]. To date, there are no methods that take into account both of these important issues. Here, we propose a Bayesian mixture model that simultaneously accounts for epistasis and sub-phenotyping. We explicitly model the hidden phenotypic and genotypic differences within the given population. Using MCMC, we simultaneously identify subpopulations with homogenous phenotypes along with their associated epistatic and marginally acting markers. We performed extensive simulations using several well-known disease models and their combinations, while taking into account the observed odd ratios and MAF distribution of the published GWAS results [2]. We compare six different methods - from classical GWAS approaches to those specifically developed to detect sub-phenotypes. Power and Type-1 Error analysis showed that our proposed method (JBASE) outperformed all competitors across the full spectrum of MAF, odd ratio and disease heterogeneity combinations. Most notably, when associations are weak, i.e. small odd ratios, and sub-phenotypic differences exist, our method is the only method that can recover true associations. References: [1] Zuk et al. (2012). The mystery of missing heritability: Genetic interactions create phantom heritability, Proc Natl Acad Sci., 109(4):1193-8. [2] Zhang et al. (2007) Bayesian inference of epistatic interactions in case-control studies. Nature Genet., 39(9),1167–1173. [3] Hauser et al (2004). Ordered Subset Analysis in Genetic Linkage Mapping of Complex Traits. Genetic Epidemiology, 27: 53–63.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1552F

IMProve: An R package for parallelizing 1000 Genomes based genotype imputation, statistical analyses, and extraction of independent hit loci. *T.A. Johnson[1], Y. Nakamura[2], M. Kubo[3], T. Tsunoda[1].* 1) Lab Med Informatics, RIKEN Yokohama Inst/CGM, Yokohama, Kanagawa, Japan; 2) Human Genome Center, Inst of Medical Science, University of Tokyo, Shirokanedai, Tokyo, Japan; 3) Research Group for Genotyping, RIKEN Yokohama Inst/CGM, Yokohama, Kanagawa, Japan.

Imputation is an important part of the modern analytical pipeline for Genome-wide Association Studies (GWAS). The 1000 Genomes Project (1000G) reference haplotypes, which possess close to complete ascertainment of common variant sites, have dramatically improved coverage and quality compared to previous HapMap datasets. This has increased power for fine-mapping disease and quantitative trait loci but also increased the number of loci requiring analysis; the current 1000G dataset has ≈11 million loci after filtering for MAF≥0.01 in East Asians. Especially with modern GWAS datasets' increasing sample sizes, this places a burden on computational infrastructure and necessitates methods to parallelize the process of imputation and data analysis. Also, with numerous loci potentially in linkage disequilibrium (LD) with the most significant SNPs, the extraction of independent hits and determination of boundaries of linked variants is a main consideration for post-imputation analysis and locus prioritization for validation studies. To facilitate these processes, we have developed an R package named IMProve, which includes methods to: 1) divide GWAS datasets into manageable chunks for phasing/imputation, 2) import and split 1000G vcf reference haplotype files, 3) export into SHAPEIT, Impute2, or MACH/minimac formats for phasing/imputation, 4) submit chunks to a compute cluster for phasing, imputation, or statistical analyses, 5) import imputation output into Bioconductor's snpStats package's compressed format, 6) summary and statistical analysis using snpStats' built-in functions, and 7) step-wise top SNP adjustment procedure (TOPSNP) for extraction and definition of independent hits. We have applied these methods to several quantitative trait and case-control analyses in Japanese samples and illustrate our methods using a height QTL analysis in ≈27k BioBank Japan samples genotyped on the Illumina 610k array and imputed with Impute 2.2.4 and 1000G Phase 1 Ver. 3. ≈8.1 million loci remained after quality control filters (MAF≥0.01, info≥0.4) and 1,776 loci had $p \geq 5 \times 10^{-8}$. The TOPSNP procedure reduced those loci to 26 independent hits, with each hit having 25±31 linked SNPs with $p \leq 5 \times 10^{-8}$ and 193±141 (median±MAD) nominally significant loci ($p \leq 10^{-4}$). Along with a normal and a single page “expanded” version of Manhattan plots, IMProve produces a novel “fireworks” plot that highlights independent hits along with the unadjusted and adjusted -log10(p-values) of linked loci.

1553W

An automatic toolbox for genome-wide association analyses. *J. Luan, S.J. Sharp, N.J. Wareham, J.H. Zhao.* MRC Epidemiology Unit, Institute of Metabolic Science, Cambridge, CB2 0QQ, United Kingdom.

Background: Numerous research consortia have been established in recent years to share data from GWASs, primarily to provide sample sizes much higher than can be achieved in any one study. Members of each consortium are expected to provide results using their own GWAS data based on a standard analysis plan developed to address a particular research question. Different analysis plans vary in terms of their requirements with respect to outcomes, covariates, interactions, subgroup analyses and methods of analysis. To meet all these different needs in a timely manner while minimising the potential for error, we have developed a toolbox of programs which automate as much of the work as possible for GWAS data from the Fenland and EPIC-Norfolk studies (www.mrc-epid.cam.ac.uk/Research/Studies), taking advantage of all available computer resources. Methods: We have developed generic programs which take advantage of the functionality of different software packages, including Stata, SNPTEST, QuickTest and ProbABEL. We also make use of Linux clusters whenever available. The application of these programs to data from a new study is straightforward, and typically uses imputed data from either HapMap or the 1000 Genomes Project. Results: A typical genome-wide association analysis using SNPTEST or QuickTest or ProbABEL is implemented with a single Stata command allowing multiple analyses to be performed in one Stata session, with types of outcomes and their transformations, the inclusion of covariates, etc, specified as options. The results can be provided as ASCII or Stata files which can be formatted to meet the specific requirements of the consortium. This work has not only improved efficiency and reproducibility for specific consortia-based analyses, but also been used in other large projects including InterAct (www.inter-act.eu) and an eQTL study of brain gene expression. One analysis of the eQTL data involved over 1000 gene expression levels each as a whole-chromosome analysis, to be used for result look-up for a genome-wide association study, and the analyses were implemented through a macro (each corresponding to an expression level and chromosome) in a batch file. Conclusion: These programs have been instrumental for analyses of our own data and can easily be adapted to analyse data from other studies. They have the potential to be used as benchmark for programs developed in other software packages such as R (http://www.r-project.org).

1554T

Variant association tool: A pipeline to perform quality control and association analysis of sequence and exome chip data. *G. Wang[1], B. Peng[2], SM. Leal[1].* 1) Baylor College of Medicine, Houston, TX; 2) The University of Texas MD Anderson Cancer Center, Houston, TX.

Currently there is great interest in detecting associations of complex human traits with rare SNVs. Association analysis of rare variants differs from those applied to common variants, because methods used to analyze common variants are underpowered to detect rare variant associations. To increase power, many rare variant association methods have been developed which aggregate variants across a genetic region (usually a gene). Before testing for rare variant associations, quality control which is specific to sequence data must be performed; where variant sites and calls are removed that do not meet specific criteria. e.g. read depth, etc. It is also necessary to annotate variants before performing association analysis to determine which variants meet inclusion criteria. To facilitate the analysis of rare variants we developed variant association tools (VAT), a tool-set that implements best practice for rare variant association studies. Highlights of the pipeline include variant site/call level quality control, phenotype/genotype based sample selections, variant annotation and selection of gene regions for analysis. Summary statistics can be generated, e.g. MAF, HWE, Ti/Tv ratios, etc. Within this pipeline a large number of rare variant association tests have been implemented to analyze both qualitative and quantitative traits including CMC (Li & Leal 2008), WSS (Madsen & Browning 2009), GRANVIL (Morris & Zeggini 2010), VT (Price et al. 2010), SKAT (Wu et al. 2011), etc. Our association testing framework is regression based which readily allows for flexible construction of association models with multiple covariates, weighting (i.e. based on frequency or predicted functionality), interactions terms (i.e. gene x gene & gene x environmental) and models for pathway analysis. VAT is capable of rapidly scanning through the entire exome using multi-processes computation, adaptive permutation approaches and conducting multiple association tests simultaneously. In addition to outputting results in tabular format they can be graphical viewed as QQ or Manhattan plots. Additionally it provides a programming interface to allow for user implementation of novel association methods. The VAT pipeline can be applied to sequence data, imputed sequence data and exome chip array data. This pipeline should be extremely beneficial in performing association analysis on the large amounts of sequence and exome chip data which is currently being generated for complex traits.

1555F

MaCH-Admix: Genotype Imputation for Admixed Populations. *E.Y. Liu[1], M. Li[2], W. Wang[1], Y. Li[1,3].* 1) Dept of Computer Science, University of North Carolina at Chapel Hill, Chapel Hill, NC; 2) Dept of Biostatistics and Epidemiology, University of Pennsylvania School of Medicine, Philadelphia, PA; 3) Dept of Genetics and Dept of Biostatistics, University of North Carolina, Chapel Hill, NC Hill.

Imputation in admixed populations is an important problem but challenging due to the complex linkage disequilibrium (LD) pattern as a result of recent admixture. The emergence of very large reference panels such as that from the 1000 Genomes Project enables more accurate imputation in general, and in particular for admixed populations and for uncommon (minor allele frequency, MAF, < 5%) genetic variants. To efficiently benefit from these large reference panels, one key issue to consider is on how to traverse the reference space harboring the most probability mass with minimum computational efforts. In modern genotype imputation framework, this corresponds to the selection of effective reference panels. In this work, we evaluate a number of methods for effective reference panel construction inside a hidden Markov model (HMM) and specific to each target individual. These methods fall into two categories: identity-by-state (IBS) based and ancestry-weighted approach. We have implemented all the methods in our software package MaCH-Admix. We evaluated the performance on individuals from recently admixed populations, including African Americans and Hispanic Americans. Our target samples include 8421 African Americans and 3587 Hispanic Americans from the Women's Health Initiative (WHI), which allow assessment of imputation quality for uncommon variants. Our experiments include both large and small reference panels; large, medium, and small target samples; and in genome regions of varying levels of LD. We also include BEAGLE and IMPUTE2 for comparison. Experiment results suggest that our novel piecewise IBS method yields consistently higher imputation quality than other methods/software. The advantage is particularly noteworthy among uncommon variants where we observe up to 5.1% information gain with the difference being highly significant (Wilcoxon signed rank test P-value < 0.0001). Our work is the first that considers various sensible approaches for imputation in admixed populations and presents a comprehensive comparison.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1556W

Two roads both taken: walking the phenotypic and genotypic paths to disease gene implication. *A. Javed, P. Ng.* Genome Institute of Singapore, A*STAR, Singapore.

Next Generation Sequencing has proven to be an invaluable search tool for pinpointing causal variants for Mendelian inherited rare disorders. Despite the constant flux in sequencing technology, the broad steps in downstream analysis of such studies show remarkable consistency. Our contribution to this analysis is twofold. Firstly, we quantify the performance of deleteriousness predictors commonly used to sift out the culprit variant. This is done both individually and in conjunction using a naïve Bayesian approach. Secondly, we use available knowledge of the phenotype to set the prior expectation of involvement of each gene. These two paths, followed independently, are combined within a Bayesian framework where the phenotype dictates the priors and the genetic data defines the observation. We estimate deleterious impact of coding and regulatory variants and benchmark these predictions using Human Gene Mutation Database as a positive set; and for each class of variants an appropriate neutral set. The continuum of predicted probabilities, after applying appropriate thresholds, yields an accuracy of 80% (or better) at the expense of 3% (or less) false positives across the board. In particular for non-synonymous mutations we combine SIFT and PolyPhen-2 predictions to achieve 91% accuracy with less than 2.1% false positives; which compares favorably to MutationTaster at 87% accuracy and 4% false positives using the same data. The predictions are pooled within each gene, extending the boundaries to incorporate cis regulatory interactions. Prior knowledge of phenotype-gene involvement comes in two flavors. Direct knowledge of it being implicated for the same or similar phenotype; or guilt by association due to known interaction with the fore mentioned. We use significance of mapped phenotype ontology to establish informativeness of direct knowledge; and a random walk of protein interaction networks to quantify association. Each step of our loosely coupled pipeline is evaluated first individually; and then in totality. It is currently being used to study a shared rare disease inflicting three unrelated families. The designed software will be made available for public use. To our knowledge, this is the first comprehensive analysis to systematically quantify, and recommend, deleteriousness predictors for both coding and regulatory mutations. Our objective of travelling the phenotypic path less travelled by makes all the difference in causal variant prediction.

1557T

Inferring Highly Polymorphic Gene Alleles Using Sequence Data and SNP Genotypes in the Han Chinese Population. *S. Chang[1], C. Fann[2], A. Hsieh[2], C. Shiu[2].* 1) Clinical Informatics and Medical Statistics Research Center, Chang Gung University, Taoyuan County, Taiwan; 2) Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan.

To dissect fully the complication of human leukocyte antigen (HLA) genetic variations, direct typing and deep re-sequencing will be of great help to reveal the disease-specific and shared risk factors involved in multifactorial, complex diseases or traits. However, high resolution HLA typing remains a laborious and expensive task due to the complexity and ambiguity of the sequence variations and thus prohibits clinicians and scientists from widespread application of HLA alleles in practice and research. Recent studies have demonstrated the usefulness of SNPs on or near the HLA loci for predicting or inferring classical HLA alleles. To date, different types of such methodologies are a few but not many. We seek to develop an ethnic-specific predictive model of highly polymorphic gene alleles, e.g., the class I and II HLA genes. Building upon available phased and unphased SNP genotype data obtained from family trios and unrelated individuals in a Han Chinese population whose HLA allele information are known, our methodology will be able to identify a small set of most informative SNPs and to accurately predict HLA alleles in a general population. By cataloging the prediction condition and dissecting the performance for a Han Chinese population, we will gain more insights into the HLA genetic variations across different ethnic populations. Besides, with the construction of a reference panel with dense SNP marker data and known HLA allele genotypes for inferring HLA alleles in Han Chinese populations, many HLA associated studies in the same population can be realized and thus may lead to the construction of allele prediction models of various associated diseases. The inferred HLA alleles can aid in mapping causal or functional variants across the MHC region, increase the power for disease or phenotype association studies, and may facilitate meta-analysis of whole-genome association data and help dissect the genetic basis of complex diseases and traits.

1558F

The Design of Custom-made Genotyping Arrays to Study Asthma in Populations of African Descent. *J. Gao[1], Y. Hu[2], J.S. Lee[3], N. Pearson[3], G. Abecasis[4], I. Ruczinski[5], T.H. Beaty[6], Z. Qin[1,2], R.A. Mathias[7], K.C. Barnes[7].* 1) Center for Comprehensive Informatics, Emory University, Atlanta, GA; 2) Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, GA; 3) Knome Inc., Cambridge, MA; 4) Department of Biostatistics, School of Public Health, University of Michigan, Ann Arbor, MI; 5) Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD; 6) Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD; 7) Division of Allergy and Clinical Immunology, School of Medicine, Johns Hopkins University, Baltimore, MD.

Asthma is a complex disease, and its increased prevalence and severity in populations of African ancestry are partially attributed to genetic ancestry. While genome-wide association studies (GWAS) on populations of European ancestry have discovered strong associations between genetic factors and asthma susceptibility and mortality, few studies have focused on populations of African descent, and even fewer have found genome-wide-significant associations. It is unclear to what extent the coverage of current state-of-the-art and commercially available genotyping chips explain the paucity of GWAS findings in populations of African ancestry for asthma. To address the challenge, we established the Consortium on Asthma among African-ancestry Populations in the Americas (CAAPA) to perform whole genome sequencing in 1,000 individuals of African ancestry (50% asthmatic), and leverage discoveries in public catalogs including the 1,000 Genomes Project, to find novel single nucleotide polymorphisms (SNPs) in African and African admixed populations, in order to develop a custom, African-ancestry gene-centric 200K SNP genotyping array ("African Power Chip"). In a preliminary analysis, we examined whole genome sequencing (WGS) data from a small group of individuals of African descent and compared this to the latest genotyping chip (HumanOmni5-Quad chip), which contains the largest list of SNPs commercially available to date. We found that among all bi-allelic SNPs that we identified, only a small proportion (~20%) is included in the HumanOmni5-Quad chip. We next attempted genotype imputation for other SNPs using Hapmap phase III data as the reference panel. We implemented the minor allele frequency- (MAF-) adjusted R^2 criteria to identify SNPs that can be well imputed. Under this criterion, we found only an additional 3% of the SNPs could be accurately imputed. Our results indicate that the current genotyping chip, aided with imputation, is highly inadequate to cover the majority of SNPs found in samples of African descent. This suggests that a specially designed genotyping panel is needed, in order to efficiently screen genetic variation, and discover genes conferring risk to asthma and other diseases, in these populations. We believe that the new chip will complement current commercially available GWAS chips, for which common and rare variants are not adequately tagged by the existing SNPs, and thereby facilitate GWAS studies on populations of African descent.

1559W

Detecting rare variants for both quantitative and qualitative traits using pedigree data. *W. Guo, Y. Y. Shugart.* Division of Intramural Division Program, National Institute of Mental Health, National Institute of Health, USA.

With the advent of sequencing technology opening up a new era of personal genome sequencing, huge amounts of rare variants data have suddenly become available to researchers seeking genetic variants related to human complex disorders. There is an urgent need of developing novel statistical methods to analyze rare variants in a statistically powerful manner. While a number of statistical tests have already been developed to analyze collapsed rare variants identified by association tests in case-control studies, to date, only two FBAT tests-for-rare (described in the updated FBAT version v2.0.4) have applied collapsing methods analogously in family-based designs. To further research in this area, we introduced three new beta-determined weights-based tests for detecting rare variants for quantitative traits in nuclear families [Guo & Shugart, 2012]. In addition to evaluating the performance of these new methods, we also evaluate the power and validity of the two FBAT tests-for-rare, using extensive simulations of situations with and without linkage disequilibrium. Results from these simulations suggested that the four tests using beta-determined weights out-perform the two collapsing methods used in FBAT (-v0 and -v1). Four new beta-determined weight methods of statistical analysis in a computer program have been submitted to the Comprehensive R Archive Network (CRAN) for general use. We now propose to extend our methods to analyzing the qualitative traits by using logistic regression model and modify the script to handle data structure of general pedigrees for both quantitative and qualitative traits.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1560T

IsoDOT detects differential RNA-isoform usage between two or more samples with high sensitivity and specificity. *W. Sun.* Biostatistics, University of North Carolina, Chapel Hill, NC.

The study of RNA isoform expression is of great importance for understanding gene expression regulation and the molecular basis of complex phenotypes, such as human diseases. Using RNA-seq data, we have developed a statistical framework and software for transcriptome reconstruction, isoform abundance estimation, and Differential isOform usage Testing (DOT). Here differential isoform usage refers to the changes of RNA-isoform expression relative to the total expression of the corresponding gene. We model the numbers of RNA-seq fragments covering any exon set by a negative binomial distribution, and achieve accurate isoform abundance estimation by a penalized regression approach. Differential isoform usage of a gene is assessed by a likelihood ratio statistic. Both simulation and real data studies demonstrate that IsoDOT achieves accurate isoform abundance estimation and delivers greatly improved sensitivity and specificity to detect differential isoform usage.

1561F

Identification of Loss of Heterozygosity from Unpaired Tumor Samples Using Next-Generation Sequencing. *J. Lee, N. Pearson.* Knome Inc., Cambridge, MA.

Loss of Heterozygosity (LOH) is one key event observed during cancer development. Inference of LOH regions is typically drawn by sequencing or genotyping tumor and matching normal DNA samples and locating chromosomal regions where heterozygous markers in a normal sample are shown to be homozygous in the matching tumor sample. Since a pair of normal and tumor sample often may not be available, several methods have been developed for SNP genotyping data for an unpaired tumor sample. However, normal sample contamination can complicate the detection of LOH region based on those methods. In addition to that, the next-generation sequencing technologies are being actively utilized in cancer research. So we developed a model for the detection of LOH regions based on a hidden Markov model (HMM) that takes into account the estimated tumor purity in a given tumor sample. We investigated 10 impure tumor samples for patients with chronic lymphocytic leukemia (CLL) of which the range of the estimated percentages of tumor cells is 50–94%. We identified long LOH segments in chromosome 6q and 21. We also spotted shorter segments shared by many of those samples in 21q11 and 13q14. We implemented the algorithm in an R package, knomeLOH.

1562W

A robust and well-defined method for Single Nucleotide Variant and Indel detection. *N. Tuzov.* Partek Incorporated, Saint Louis, MO.

A number of present methods for SNV and Indel detection try to incorporate the information from annotations such as base quality scores supplied by the sequencer and alignment quality scores supplied by the aligner of choice. In the currently available algorithms, the outcome is sensitive to a large number of optional parameters that must be set either by the user or by the software developer. Such latitude provides a rather questionable way out for the cases when the algorithm is not stable enough to deliver any sort of answer. While the pool of annotations to choose from may be large, the user is given insufficient means to determine which particular annotations, if any, contribute to better SNV and Indel detection. In addition, the introduction of integer-valued and categorical annotations is problematic or impossible. To address these shortcomings, we develop a supervised learning method for SNV and Indel detection that (i) trains the model based on available SNP databases, such as dbSNP, (ii) is fast and robust, (iii) accepts both quantitative and categorical annotations, such as transition / transversion factor, (iv) allows the researcher to build the model consisting of annotations that contribute the most to better SNV and Indel detection and (v) has few optional algorithm parameters, reducing the impact of subjective interference on the outcome and resulting in a simple user interface. We demonstrate the performance of our method on experimental data using both quantitative and categorical annotations. The results show a practically significant improvement in SNV and Indel calling compared to the baseline case with no annotations, and an important contribution of categorical annotations that are not accepted by the currently available SNV and Indel detection software.

1563T

DETECTANCE: Graphical software for study design and planning, based on evaluation of the detectance distributions for probands conditional on the phenotypes of relatives. *T. Hiekkalinna[1, 2], P. Norrgrann[2], J.H. Lee[8, 9, 10], K.M. Weiss[7], M. Perola[1, 2], J.D. Terwilliger[1, 2, 3, 4, 5, 6].* 1) Chronic Disease Prevention, Public Health Genomics Unit, National Institute for Health and Welfare, Helsinki, Finland; 2) Institute for Molecular Medicine Finland, Helsinki, Finland; 3) Genetics and Development, Columbia University, New York, NY; 4) Psychiatry, Columbia University, New York, NY; 5) Columbia Genome Center, Columbia University, New York, NY; 6) Division of Medical Genetics, New York State Psychiatric Institute, New York, NY; 7) Anthropology, Pennsylvania State University, State College, PA; 8) Epidemiology, Columbia University, New York, NY; 9) Sergievsky Center, Columbia University, New York, NY; 10) Taub Institute, Columbia University, New York, NY.

Medical genetics researchers most often think about effect size measured as the probability of a given phenotypic outcome conditional on genotype. However, most genetic epidemiology studies start from a collection of individuals or families ascertained conditional on phenotype, with the hope that the phenotypes of individuals in these families predict the underlying risk genotypes, which is the fundamental basis of mapping. To this end, it is important to consider the ramifications of an etiological model in terms of penetrances [P(Ph |G, E)] and ascertainment scheme on the detectance P(G, E | Ph; Ascertainment). As a trivial example, if we wanted to determine the probability that a randomly ascertained individual had genotype X at locus 1 conditional on observed phenotype Ph, the detectance would be simply

$P(G_1=X|Ph) = \sum_{G2}...\sum_{GN}\sum_{E1}...\sum_{EM} P(Ph|G_1=X,G_2,...,G_N,E_1,...,E_M)P(G_1 = X, G_2,...,G_N,E_1,...,E_M)/\phi$,

where ϕ is the probability of the phenotype in the population (i.e. the same sum over all possible genotypes for G_1). The detectance represents what you would therefore expect to see in a dataset conditional on phenotype.

DETECTANCE currently works on models with a small number of multiallelic genes influencing a qualitative or quantitative trait, and computes the detectance for a proband conditional on his phenotype and the phenotypes of others in his nuclear family. In real time it graphically displays the change in detectance as parameters of the etiological model, phenotypes, or linkage/LD among the genes change through the use of sliding bars. This visually makes the conditions in which family-based sampling or sampling on extreme phenotypes is much more powerful than cross-sectional association studies obvious. DETECTANCE software is freely available from the authors.

1564F

Genetic Instrumental Variable Studies of the Effects of Maternal Risk Factors on Oral Clefts. *G.L. Wehby[1], L.M. Moreno[2], P. Romitti[2], K. Christensen[4], L. DeRoo[6], A. Wilcox[6], R. Munger[3], R. Lie[5], J.C. Murray[2].* 1) Health Management and Policy, University of Iowa, Iowa City, IA; 2) University of Iowa; 3) Utah State University; 4) University of Southern Denmark; 5) University of Bergen; 6) National Institute of Environmental Health Sciences.

Several studies have suggested that maternal prenatal smoking, alcohol use, and obesity may each play a role in the etiology of cleft lip and/or cleft palate (CL-P). However, these studies do not account for unobserved or unmeasured confounders that influence maternal selection into these risk factors but are also related to CL-P risk. It is very difficult to observe and directly control for all relevant confounders in available data samples. Ignoring unobservable confounders may significantly bias the estimated risk factor effects and their interactions with genetic risk factors. The objective of this study is to identify the causal effects of maternal prenatal smoking, alcohol use, and obesity on CL-P risks using a genetic instrumental variable (IV) model (or Mendelian Randomization) to account for unobserved confounders. The IV model uses genetic variants (called instruments) that are: 1) strongly correlated with the risk factors of interest; 2) otherwise unrelated to CL-P through unobservable confounders. Under these two conditions, the model exploits variation in the risk factors that is predicted by instruments but is unrelated to unobservable confounders. Using appropriate instruments, the IV model provides unbiased estimates of the causal effects of these risk factors on CL-P. In addition, the IV model can be applied to obtain unbiased estimates of the interaction effects between these risk factors and genetic factors that modify their effects. The study includes a unique and large international consortium of population-based samples of 3800 mothers of children with CL-P and 5000 mothers of unaffected children. We are currently genotyping the top 50 candidate SNPs identified from the literature to be evaluated as instruments for prenatal smoking, alcohol, and obesity. We will use the SNPs that fit the above two IV conditions as instruments to estimate the average effects of these risk factors on CL-P as well as the interactive effects between these risk factors and 40 SNPs that include the top variants identified so far to affect CL-P risk and influence detoxification and metabolism. The presentation will report on the methodology and analysis results. This is the first large-scale study to estimate the causal effects of major maternal risk factors on CL-P. These estimates are essential for unraveling the etiology of CL-P and developing cost-effective programs for CL-P prevention and counseling at-risk women.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1565W

The Prevalence of X-linked Hypohidrotic Ectodermal Dysplasia (XLHED) in Denmark. *M. Nguyen-Nielsen[1], S. Skovbo[1], D. Svaneby[2], L. Pedersen[1], J. Fryzek[1,3].* 1) Department of Clinical Epidemiology, Aarhus University Hospital, Aarhus, Denmark; 2) Department of Clinical Genetics, Aarhus University Hospital (Skejby), Aarhus, Denmark; 3) Epidstat Institute, Ann Arbor, MI, USA.

X-linked hypohidrotic ectodermal dysplasia (XLHED, MIM 305100) is a genetic disease caused by a defect in the ectodysplasin-A gene (EDA, MIM 300451). The cardinal clinical features associated with XLHED are hypohidrosis (absence of sweat glands), thin/sparse hair, and abnormal teeth. Children with XLHED have been shown to have an increased risk of death from hyperpyrexia due to the inability to sweat and thereby regulate core body temperatures. XLHED is the most common form of hypohidrotic ectodermal dysplasia (HED); however, no population-based prevalence estimates are available. We aimed to: 1) estimate the prevalence of XLHED in the general Danish population per January 1, 2011; 2) determine the age group in which diagnosis most frequently occurs; and 3) quantify the most frequent clinical feature of XLHED. Subjects and Methods: Denmark (population 5.5 million) has an extensive collection of medical registries and each resident has a unique personal identification number which allows for unambiguous data-linkage across multiple registries. We conducted a nationwide cross-sectional study using registry data (the Danish National Patient Register and the Civilian Registration System) and data from clinical departments (Clinical Genetics, Dermatology, Pediatrics and Specialized Dentistry) in order to identify all XLHED cases in Denmark diagnosed between 1995–2010. We categorized XLHED patients into three groups: 1) molecularly-confirmed XLHED; 2) clinically-diagnosed HED and 3) probable HED (registered with a constellation of clinical features on the basis on a clinical algorithm that we designed). We calculated prevalence estimates on the basis of these groups and stratified by age and gender. We also investigated patients' medical history and quantified the occurrence of cardinal features. Results: We identified 90 molecularly-confirmed XLHED, 146 clinically-diagnosed and 988 possible HED cases (total n= 1224). The prevalence ranged from 1.6–22 per 100,000 across the three groups. Diagnosis occurred most frequently between the ages 11–18. Teeth abnormalities occurred in 79 percent of all cases and 52 percent of the molecularly-confirmed. Conclusion: We present the first ever population-based prevalence estimates of XLHED and show that the prevalence may be higher than previously estimated. Diagnosis of XLHED is most often delayed until late childhood. Teeth abnormalities were the most frequent clinical marker of XLHED.

1566T

The PhenX Toolkit: standard measures facilitate cross-study analyses. *C.M. Hamilton[1], W. Huggins[1], H. Pan[1], D.B. Hancock[1], J.G. Pratt[1], J.A. Hammond[1], T. Hendershot[1], D.R. Maiese[1], K.A. Tryka[2], K. Sher[3], K. Conway[4], M. Scott[5], W.R. Harlan[6], J. Haines[7], L.C. Strader[1], H.A. Junkins[8], E.M. Ramos[8].* 1) RTI International, Research Triangle Park, NC; 2) National Center for Biotechnology Information, National Library of Medicine, Bethesda, MD; 3) University of Missouri, Columbia, MO; 4) National Institute on Drug Abuse, Rockville, MD; 5) National Institute on Alcohol Abuse and Alcoholism, Bethesda, MD; 6) Retired, Associate Director for Disease Prevention, Office of the Director, National Institutes of Health, Bethesda, MD; 7) Center for Human Genetics Research, Vanderbilt University, Nashville, TN; 8) National Human Genome Research Institute, Bethesda, MD.

The PhenX (consensus measures for Phenotypes and eXposures) Toolkit (https://www.phenxtoolkit.org/) is an online resource that catalogs broadly validated and well-established measures of phenotypes and exposures for use in genomic and other types of studies involving human subjects. The PhenX Toolkit currently includes 339 measures covering a broad scope of 21 research domains including Demographics, Anthropometrics, Cancer, Nutrition, Environmental Exposures, Neurology and Social Environments, and six specialty areas related to Substance Abuse and Addiction (SAA). Since its initial release in 2009, the Toolkit has logged 300,000 visits, and currently has 800 Registered Users. Toolkit visitors come from every state in the USA and 148 countries. Investigators can find measures of interest by browsing domains, collections or measures, or by searching using the Smart Query Tool. For each measure, the Toolkit provides a description of the measure, the rationale for its inclusion, detailed protocol(s) for collecting the data, and supporting documentation. The goal of the SAA project was to create six “Specialty” collections and one “Core” collection of SAA related measures. The SAA project used a streamlined version of the established PhenX consensus process to select 44 new measures which were added to the Toolkit in February, 2012. Another project, PhenX RISING (Real world, Implementation, SharING), brings together seven investigators who were awarded funds to incorporate PhenX measures into existing, population-based genomic studies. Over the course of the project, these early adopters are evaluating PhenX measures and recommending improvements to the Toolkit. The Toolkit also provides tools to help investigators integrate PhenX measures into their study design. That is, the Toolkit provides custom data collection worksheets to support data collection and custom data dictionaries to facilitate data submission to the database of Genotypes and Phenotypes (dbGaP). In an effort to link PhenX measures with data that is already in dbGaP, PhenX measures were mapped to 16 dbGaP studies. Of the 822 non-administrative variables included in these dbGaP studies, 79% could be mapped to PhenX measures. This mapping will help investigators identify studies in dbGaP that may be well-suited for various cross-study analyses. Implications for integration of PhenX measures in dbGaP will be presented. Funding was provided by NHGRI 5U01HG004597 and 3U01HG004597-03S3.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1567F

Comparison of Genome-Wide association studies for smoking behaviour between self-reported and biomarker measures. *A.I. Stiby[1], N. Timpson[2], M. Hickman[1], D. Evans[2], J. Kemp[2], G. Davey Smith[2], S. Ring[2], B. St Pourcain[2], J. Henderson[1], J. Macleod[1].* 1) School of Social and Community Medicine, The University of Bristol, Bristol, United Kingdom; 2) MRC CAiTE Centre, School of Social and Community Medicine, University of Bristol, Bristol, United Kingdom.

The validity of self-reported smoking is questionable as smokers tend to underestimate the amount they smoke and adolescents overestimate their cigarette consumption due to patterns in their social group. Cotinine is the principle metabolite of nicotine and is used as an objective assessment of cigarette smoking. Work here compares the results of genomewide association (GWA) analysis performed on self-reported (SR) smoking versus cotinine measures in individuals from the Avon-Longitudinal Study of Parents and Children (ALSPAC). It has been established that there are considerable power gains to be afforded using cotinine as a biomarker for smoking behaviour, though this has yet to be examined in the GWA setting. Genetic data from 9,912 participants were genotyped using the Illumina HumanHap550 quad genome-wide SNP genotyping platform by 23andMe. We used 4122 individuals with self-reported smoking data, of which 723 (17.5%) were smokers. 2843 of these also had data on cotinine levels, of which 260 (9.1%) were above the cut off for smokers (7 ng/ml of blood). Of note and highlighting the value of these data, 268 (9.7%) individuals who describe themselves as smokers were found to be non-smokers by cotinine assay and 20 (0.7%) cotinine-assessed smokers described themselves as non-smokers. Cotinine was associated with SR smoking status, although SR smokers had higher circulatory levels of cotinine than the normal cut-off between the biologically assessed non-smokers and smokers (39.35 ng/ml of blood, $p < 0.001$). GWA results showed a clear contrast between SR frequency of smoking within smokers and results of objectively assessed levels of cotinine which were above the active smoking cut-off. Whilst the former was underpowered to deliver even established signals for smoking behaviours, the analysis of cotinine gave multiple suggestive hits, the highest of which was rs12607939 ($p=9.53 \times 10^{-8}$), this was not maintained in SR smoking ($p = 6.12 \times 10^{-1}$). Two genes were found to exhibit suggestive evidence of association in both the cotinine and self reported GWA analysis (*DAB1* and *BTBD3*), although most hits were from the analysis of cotinine levels in smokers. Our results suggest problems with bias in the self-reported smoking data due to adolescent over-estimation of the amount they smoke; furthermore it suggests that cotinine is giving a more precise measure than SR smoking status in GWA studies.

1568W

Design of Sequencing Studies in Family Samples, With Application To The Study of Substance Use Disorders. *S. Feng[1], S. Vrieze[2], B. Tarrier[1], J. Bragg-Gresham[1], X. Zhan[1], C. Sidore[1,3], M. McGue[2], W. Iacono[2], G. Abecasis[1].* 1) Biostatistics Dept, Center for Statistical Genetics, University of Michigan, Ann Arbor, MI; 2) Psychology Dept, University of Minnesota, Minneapolis, MN; 3) Dipartimento di Scienze Biomediche, Università di Sassari, Sassari, Italy.

Next-generation sequencing is enabling complex trait studies to move beyond common variants and also examine the contribution of rare variants. Given the large per sample costs associated with sequencing experiments, an important step in these experiments is the selection of individuals likely to carry trait associated variants.

Here, we propose a likelihood based method for selecting individuals for sequencing which takes into account phenotypic information, individual specific covariates, and any available genotypes. Our method is applicable when information on related individuals is available and can prioritize families and individuals likely to carry trait associated rare variants. Using simulations, we show that our method is effective at identifying sets of individuals enriched for rare trait-associated alleles. Furthermore, we show that the resulting samples provide greater power for association analysis than samples selected using extreme phenotype-based strategies.

We apply our method to select 1,000 individuals likely to carry rare variants associated with a substance use phenotype measured in 4,934 individuals from the Minnesota twin and family study registry. We simulated rare variants explaining 0.1–1.0% of the trait variance. At 0.1% frequency, we expect about 10 copies of the variant to be present in the sample of 4,934 individuals. If these variants account for 1% of trait heritability (an effect size of 2.2 standard deviations per allele), simulations predict that using our method to prioritize 1,000 individuals for sequencing, would result in nearly all of these 10 copies (9.4 on average) being carried by one of the sequenced individuals. Simulations of downstream association analyses predict ~86% power (from permutation tests at $\alpha = 10^{-3}$) to detect trait associations for the above rare variants in samples selected using our method, and ~81% power from samples selected using extreme phenotype-based strategy (again, at $\alpha = 10^{-3}$). The power to detect association afforded by sequencing the original sample of 4,934 individuals would be ~93%. Our results show that with limited resources, using family information, phenotypes and covariates to prioritize individuals for sequencing can facilitate discovery of rare trait associated alleles.

1569T

Family-based designs for sequencing studies. *D. Thomas.* Preventive Med, Univ Southern California, Los Angeles, CA.

The cost of next-generation sequencing is approaching that of early genome-wide SNP genotyping panels, but still out of reach for the large-scale epidemiologic studies with the tens of thousands of subjects that will be required for testing association with disease. The many millions of rare variants that will be discovered in a genomewide scan also pose a major challenge for distinguishing causal from non-causal variants. Family-based designs for sequencing substudies provide a means to identify novel variants and prioritize them for their likelihood of causality by exploiting their co-segregation with disease (Petersen et al, AJHG 1998;62:1516-24 PMID 9585599). Other advantages include the possibility of enriching for family history, improved imputation (Li et al, Annu Rev Genomics Hum Genet 2009;10:387-406 PMIC 19715440), and efficient two-step analyses that exploit between- and within-family comparisons (Wason & Dudbridge, AJHG 2012;90:760-733 PMID 22560088). Using simulation, we compare simple rule-based, likelihood ratio, Bayes factor, and score statistics in terms of their ability to distinguish causal from non-causal variants. We find that the Bayes factor criterion of Petersen et al provides the best discrimination, but a novel generalization of the score test of Ionita-Laza & Ottman (AJHG 2011;189:1061-8 PMID 21840850) that combines the phenotype information from all pedigree members together with the available sequence data on the sequenced subset is a close contender, is model free, and computationally simpler. Two-phase designs using a family study for discovery and prioritization followed by a large case-control or family-based association test consistently provide better power for discovery, prioritization, and subsequent association testing than single-phase case-control association tests because of the considerable reduction in the multiple comparisons penalty. The approach is illustrated using two substudies within the Colorectal Cancer Family Registries, one large pedigree with 50 members genotyped at GWAS SNPs, one with whole exome data on a subset of families from a previous linkage scan.

1570F

From a large-scale discovery study to a replication study. *C. Kuo, D. Zaykin.* Biostatistics Branch, National Institute of Environmental Health Sciences, Durham, NC.

Thousands of susceptibility loci have been identified to date through large-sacle association studies. However, only a fraction of top hits are being successfully validated in replication studies. Replication is indispensable in that it allows researchers to eliminate study-specific biases, refine association signals, and generalize findings to other samples and populations. Usually, the most significant results are followed up, and a result is declared a replication if the statistical significance is reached and if the effect direction in the discovery and replication studies match. We propose a strategy for planning a replication study based on integrating the ranking false discovery rate, which is the proportion of false effects among the top follow-up loci (Kuo and Zaykin 2011, Genetics, 189:329–340) and our new developments for estimation of the number of true effects and their effect size distribution. We give guidelines for the number of follow-up loci and the optimal sample size for the replication stage.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1571W

Optimal Sample Selection for Large-Scale Next Generation Sequencing Experiments. *G.W. Beecham[1], R. Rajbhandary[1], E.R. Martin[1], L.S. Wang[2], K. Lunetta[6], R. Mayeux[4], J.L. Haines[5], L. Farrer[3], M. Pericak-Vance[1], G. Schellenberg[2], The Alzheimer Disease Genetics Consortium.* 1) Hussman Institute for Human Genomics, University of Miami, Miami, FL; 2) School of Medicine, University of Pennsylvania, Philadelphia, PA, USA; 3) School of Medicine, Boston University, Boston, MA, USA; 4) Taub Institute of Research on Alzheimer's disease, Columbia University, New York, NY, USA; 5) Center for Human Genetics Research, Vanderbilt University, Nashville, TN, USA; 6) Boston University School of Public Health, Boston University, Boston, MA, USA.

Though next generation sequencing (NGS) costs continue to drop, they remain too high for comprehensive sequencing of all available samples in all but the smallest projects. As such, it is often necessary to select a subset of samples for large-scale NGS experiments. As part of a large NGS experiment for Late-Onset Alzheimer disease (LOAD) we have evaluated a variety of sample selection schemes and their impact on statistical power to find rare risk variants and rare protective variants.

We have evaluated designs based on random selection, extreme-phenotype, “super-controls” (low probability of conversion to AD), and protective effects. In these designs we have investigated the impact of misclassification rates (rate at which controls progress to case status), genetic heterogeneity (a case's disease status is essentially caused by other risk factors), and risk exposure (the amount of risk to which a control has been exposed; important in detecting protective variants). A large sample of 10,000 cases and 10,000 controls from the Alzheimer Disease Genetics Consortium were used to evaluate the different designs using real data. Baseline risk was taken from known sex, age, and apolipoprotein e genotype specific incidence measures and was modified using known common risk loci discovered in genome-wide association studies. Designs were evaluated by comparing the effective risk ratios after accounting for misclassification, heterogeneity, and exposure to that of an ideal study population.

While it is commonly held that extreme phenotype designs are more powerful when sample sizes are limited, we found that it greatly depends on the rates of misclassification. In LOAD the extreme phenotypes are also correlated with higher rates of misclassification, thereby limiting the power of the extreme phenotype designs. Additionally, some protective designs are poorly powered to detect association at risk loci, but do not offer significant gains in power at protective loci either.

As with any power calculation, the optimal study design will ultimately depend on the number of samples and specific rates of incidence, misclassification, heterogeneity, and exposure for the particular disease. For LOAD, a combination of the extreme-phenotype design and “super-control” design appear to be optimal in most projects.

1572T

Meta-analysis of ITGAM coding variant, rs1143679 (R77H) in systemic lupus erythematosus (SLE) cases and controls. *X. Kim-Howard[1], V. Pradhan[2], K.H. Chua[3], A. Maiti[1], K. Ghosh[2], S. Nath[1].* 1) Oklahoma Med Res Foundation, Oklahoma City, OK; 2) National Institute of Immunohaematology, Indian Council of Medical Research, Mumbai, India; 3) Department of Medical Microbiology, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia.

Background: We identified ITGAM (CD11b) as a novel SLE susceptibility gene. To date, ITGAM variant rs1143679 (R77H) is one of the few SLE susceptibility loci which consistently replicated in multiple populations with European, African, Hispanic, and Asian ancestries. Methods: Meta-analyses were performed using independent published data for ITGAM SNP rs1143679. This analysis included 26,603 individuals (10,992 cases and 15,611 controls) from 17 countries across North American, Europe, Asia and the Indian subcontinent. Breslow-Day tests were performed to assess the homogeneity of odd rations (OR) across populations and to estimate the overall strength of association between rs1143679 with SLE. Since ORs were homogeneous then a fixed model was appropriate in the meta-analysis, performed with R-Meta, CatMap software, and PLINK. Results: Meta-analysis of published data from multiple populations suggest that rs1143679 is strongly associated with SLE across diverse ancestries. A meta-analysis using published data from multiple European-derived, African-derived, Asian, and Hispanic populations greatly reinforces the ITGAM-SLE association. As expected, minor allele frequencies for rs1143679 varied across ethnic populations, from 0.51% in Chinese controls to 24% in Portuguese controls. Test for homogeneity was not significant (Q-statistic, PHET = 0.30) demonstrating that odds ratios were homogenous across populations. Results from the meta-analysis showed strong association between this SNP and SLE (Pmeta = 1.52×10-94, OR (95% CI) = 1.81 (1.71–1.92)). For all population specific analyses rs1143679 was strongly associated with SLE (European: P = 2.0×10-30, OR = 1.80; European-American: P = 1.8×10-35, OR = 1.79; African-American: P = 3.8×10-12, OR = 1.64; Latin-American: P = 1.8×10-13, OR = 1.97; Asian: P = 3.4×10-11, OR = 2.56). Conclusion: Our meta-analysis further confirmed the exon-3 coding SNP, rs1143679 (R77H) is highly associated with SLE across multiple populations across from North American, Europe, Asia and the Indian subcontinent.

1573F

Understanding missing data in a large STR dataset of over 100,000 Brazilian individuals genotyped with PowerPlex 16 kit. *V.R.C. AGUIAR[1], R.V. ROHLFS[2], K.E. LOHMUELLER[2], A.M. CASTRO[3], F.S.V. MALTA[3], A.C.S. FERREIRA[3], I.D. LOURO[1], R. NIELSEN[2].* 1) UNIVERSIDADE FEDERAL DO ESPIRITO SANTO, VITORIA, ES, BRAZIL; 2) UNIVERSITY OF CALIFORNIA BERKELEY, BERKELEY, CA, US; 3) LABORATORIO HERMES PARDINI, BELO HORIZONTE, MG, BRAZIL.

Short tandem repeat (STR) or microsatellite genotyping can be an error-prone process. This is especially challenging because these errors can have diverse, complex and cryptic origins, such as mutations in primer-biding sites, low quality or quantity of DNA template, or human errors in allele scoring.

All these errors can create bias in the estimates of heterozygosity, population size and structure, migration rates, kinship and parentage. It is therefore important in forensic and population genetic analyses to consider the possibility of genotyping errors, so that the reliability of the collected data is not compromised.

In this study, we investigate a large dataset of over 100,000 individuals whose DNA samples were submitted for paternity testing. These individuals were genotyped at 15 STR loci using the standard PowerPlex16 kit, manufactured by Promega Corporation. Preliminary results show that the amount of missing data per locus is very heterogeneous across loci, and there is some evidence of a correlation between missing data and Hardy-Weinberg departures at particular loci.We shall further investigate if the distribution of missing data across loci is correlated with marker length or fluorescent dye type. Finally, we will test whether the amount of missing data varies across individuals and whether it is correlated with an excess of homozygous genotypes per individual.

Our study will help illuminate the most common sources of errors in STR datasets of this kind.

1574W

Population based study of permanent teeth agenesis in Japanese schoolchildren. *J. Machida[1, 4], T. Nishiyama[2], M. Kamamoto[3, 4], S. Yamaguchi[4], M. Kimura[4], A. Shibata[4], K. Yamamoto[1, 4], S. Makino[1, 4], H. Miyachi[4], K. Shimozato[4], Y. Tokita[5].* 1) Oral and maxillofacial surgery, Toyota memorial hospital, Toyota city, Aichi, Japan; 2) Clinical Trial Management Center, Nagoya City University Hospital, Nagoya, Japan; 3) Oral and Maxillofacial Surgery, Himeji Red Cross Hospital, Himeji, Hyogo, Japan; 4) Maxillofacial Surgery, Aichi-Gakuin University, Nagoya, Aichi, Japan; 5) Perinatology, Institute for Developmental Research, Aichi Human Service Center, Kasugai, Aichi, Japan.

Tooth agenesis is one of the most common congenital anomaly in human and it is characterized by developmental absence of teeth. Clinical and epidemiologic studies of defined geographic populations can serve as a means of establishing data for genetic counseling and as a first step in identification of causes. In this study, we carried out clinical and epidemiologic study at Aichi Prefecture, Japan. 4088 Medical records of 15 years old schoolchild from Toyota city were reviewed and prevalence rates calculated. In addition, medical and dental records of 68 unrelated patients with tooth agenesis of Aichi-Gakuin University were reviewed to obtain sibling recurrence risk. The prevalence rate of congenital missing teeth was 7.64 %(95% CI: 6.82%; 8.52%). The sibling recurrence risk was 0.29 (95% CI: 0.20; 0.40) overall. One, two, three to four, and more than 4 missing teeth count 0.26 (95% CI: 0.09; 0.51), 0.25 (95% CI: 0.07; 0.52), 0.21 (95% CI: 0.05; 0.51), 0.31 (95% CI: 0.17; 0.49) respectively. The study is population based and assure high data quality. Our findings on the intra oral distribution of agenesis of permanent teeth in children may help us better understand the etiology of agenesis.

1575T

Are heritability estimates of binary traits meaningful? *P.H. Benchek, N.J. Morris.* Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, OH.

It is commonly acknowledged that estimates of heritability from classical twin studies have many potential shortcomings. Despite this, in the post-GWAS era, heritability estimates have come to be a continual source of interest and controversy due to the so-called “missing heritability.” While the heritability estimates of quantitative traits are certainly subject to a number of biases, in this paper we will argue that the standard statistical approach to estimating the heritability of a binary trait relies on some additional untestable assumptions which, if violated, can lead to badly biased estimates. The ACE liability threshold model, assumes at its heart that each individual has an underlying liability or propensity to get the binary trait (e.g., disease), and that this unobservable liability is multivariate normally distributed (i.e., for a given twin pair). We investigate a number of different scenarios which violate this assumption, such as the existence of a major gene and the existence of a dichotomous exposure. We found in each case that substantial asymptotic biases can occur, which no increase in the sample size can remove. Thus, the meaningfulness of heritability estimates under the ACE liability model is questionable.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1576F

Genetic ancestry and mammographic density among US Caucasians. *J.L. Caswell[1], K. Kerlikowske[1,2], J. Shepherd[3], S.R. Cummings[2,4], D. Hu[1], S. Huntsman[1], E. Ziv[1].* 1) Department of Medicine, UCSF, San Francisco, CA; 2) Department of Epidemiology and Biostatistics, UCSF, San Francisco, CA; 3) Department of Radiology, UCSF, San Francisco, CA; 4) California Pacific Medical Center Research Institute, San Francisco, CA.

Background: Mammographic density (MD) is one of the strongest risk factors for breast cancer and is known to be approximately 60 percent heritable. Genome wide association studies (GWAS) have identified several variants associated with MD, but only a small fraction of the heritability has been explained. We performed a GWAS of women at the extremes of MD. Here we report an unexpected association between genetic ancestry and MD that has important implications for gene discovery for this trait. *Methods:* We used samples from women who participated in the California Pacific Medical Center Research Institute Cohort. MD was measured as percent fibroglandular volume using the single-energy X-ray absorptiometry technique and linked to blood samples via the San Francisco Mammographic Registry, which also collected data on ethnicity and breast cancer risk factors. We selected 1000 self-identified Caucasian women from the top and bottom quintiles of age and BMI-adjusted MD. After removing samples with low quality DNA, we were left with 478 women in the highest MD quintile and 472 in the lowest. We genotyped samples using the Illumina 1M platform at CIDR. We incorporated publicly available genotype information from European and Middle Eastern populations into principal component analysis (PCA) to infer the ancestry of the women in our study. *Results:* PCA identified a subpopulation of women defined the by 1st principal component representing approximately 15 percent of the sample. Women in this group were more likely to have high MD ($p = 0.0025$). Incorporating European and Middle Eastern populations into PCA identified these women as having Ashkenazi Jewish ancestry. Using multivariate regression to correct for epidemiologic factors associated with MD, including age at parity and use of postmenopausal hormone therapy, did not attenuate the association. *Conclusion:* Women of Ashkenazi Jewish descent are more likely to have high MD. Previous studies have demonstrated that BRCA1 and BRCA2 mutations do not affect MD and are therefore unlikely to explain our results. Since Ashkenazi Jews are a founder population, they may have a unique set of variants that increase MD. A better understanding of the genetic determinants of MD may open new avenues for breast cancer prevention and treatment.

1577W

Robust partitioning of local heritability at associated GWAS loci. *A. Gusev[1], G. Bhatia[2], B. Pasaniuc[1], N. Zaitlen[1], A.L. Price[1].* 1) Epidemiology, Harvard School of Public Health, Boston, MA; 2) Harvard-MIT Division of Health, Science and Technology, Boston, MA.

Recent work quantifying the heritability of complex traits accounted for by dense marker data has spurred interest in obtaining robust and biologically meaningful results from partitioned heritability. In particular, estimates of observed heritability (h^2) from one or more specific loci can yield additional insights that are fundamental to understanding the potential impact of discovered loci as well as designing follow-up and fine-mapping studies. Here, we explore the efficacy of estimating local h^2 from loci suspected to be harboring causal variants, as are typically identified in large-scale association studies. A major hurdle for this type of analysis is the presence of LD between markers, which has been known to deflate h^2 when causal variants are poorly tagged. Moreover, even when all causal variants are observed in the data, we find that the h^2 can still be biased due to differentially structured LD between allele frequencies. For example, in a simulation on 5000 individuals from the WTCCC the estimate of genome-wide h^2 is deflated by 40% if a majority of the observed causal variants are uncommon (where no such deflation is present when markers are uncorrelated). In a scenario where putative causal regions are known but the causal SNPs unobserved, local h^2 over the regions can recover 95% of the true heritability from common variants and 50% from low-frequency variants. After applying a regression-based technique to account for structured LD, the latter estimate increases to 70%. When the putative loci are previously unknown but genome-wide significant associations account for at least 5% of the phenotype variance, we can query the regions around these hits for additional missing heritability. Where the underlying causal SNP is uncommon (common), local h^2 can detect statistically significantly more heritability than traditional stepwise-conditional association 50% (28%) of the time. This value increases to 95% (67%) for uncommon (common) SNPs when multiple unobserved variants are causal at the locus. Measuring h^2 at GWAS loci in three real diseases where modest heritability is explained (Crohn's Disease, Rheumatoid Arthritis, and Type 1 Diabetes), all yield significant increases in observed-scale h^2 by a factor of 1.3, 1.5, and 2.2 respectively over the h^2 estimated from GWAS hits alone. These findings support the use of local h^2 to interrogate discovered loci for additional heritability and motivate subsequent fine-mapping studies.

1578T

Multiple mediation analysis in case-control studies: application to detecting joint mediating effects of smoking and chronic obstructive pulmonary disease on the association between CHRNA5-A3 genetic locus and lung cancer risk. *J. Wang, S. Shete.* Dept Biostatistics, UT MD Anderson Cancer Ctr, Houston, TX.

Mediation model is a statistical approach to explore the direct and indirect effects between an independent variable and a dependent variable by including another mediator (or mediators). Recently, there has been of interest using mediation analysis to dissect the direct and indirect effects of genetic variants on complex diseases using case-control studies. However, the case-control study design could cause bias in the estimates of the genetic variant-mediator associations, and in turn, lead to biased estimates for the specific indirect effects and percent mediated. We investigated a multiple mediation model involving a three-path mediating effect through two mediators using case-control study data. We proposed an approach to correct the biased coefficients and provided the accurate estimates for the specific indirect effects and percent mediated. Our approach can also be used when the original case-control study is frequency-matched on one of the mediators. We employed the bootstrapping approach to assess the significance of indirect effects. We conducted simulation studies to investigate the performance of the proposed approach and showed that it provides more accurate estimates for the indirect effects as well as the percent mediated compared with the standard regression approach. We applied the proposed approach to the multiple mediation study of simultaneous mediating effects of smoking and chronic obstructive pulmonary disease (COPD) on the association between CHRNA5-A3 genetic locus and lung cancer risk using lung cancer case-control study. The results of this analysis showed that the genetic variant influences the lung cancer risk indirectly through all three different pathways. The percent mediated was 18.3% through smoking alone, 30.2% through COPD alone and 20.6% through the three-way path including both smoking and COPD, and the total genetic variant-lung cancer association explained by the two mediators, smoking and COPD, was 69.1%.

1579F

Bayesian-Optimized Linkage Disequilibrium improves gene-based association tests. *H. Huang[1, 2, 4], S. Zahid[1], P. Chanda[1, 2], D.E. Arking[3], J.S. Bader[1,2].* 1) Department of Biomedical Engineering, Johns Hopkins University, Baltimore MD; 2) High Throughput Biology Center, Johns Hopkins University School of Medicine, Baltimore MD; 3) McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore MD; 4) Current address: Analytic and Translational Genetics Unit, Department of Medicine, Massachusetts General Hospital, Boston MA and Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge MA.

Genome-wide association studies have revealed that multiple associations can occur independently within a gene or locus. When individual genes or loci are enriched for independent association signals, multivariate gene-based tests that gather information across loci defined by gene annotations can have improved signal-to-noise and better power than univariate tests. Multivariate tests require knowledge of linkage disequilibrium (LD) within a locus. This information is typically unavailable for large-scale meta-analyses, and thus many gene-based tests employ LD from a reference population. It is plausible to assume that choosing a reference population whose LD resembles the study population will improve the performance of multivariate tests. Here we present Bayesian-Optimized Linkage Disequilibrium (BOLD), a method that infers LD for a study population given its marginal allele frequencies and a set of reference chromosomes. We first assign non-negative weights to chromosomes from a reference population so that the weighted sum reproduces the marginal allele frequencies for the study population. The weighted chromosomes are then used to impute the LD for the study population. The objective function for the non-negative regression is motivated by a Bayesian formulation for the probability of the observed marginal allele frequencies given the reference chromosomes and the weights. Non-negative regression is a computationally efficient convex optimization that can be performed rapidly. This approach is motivated in part by previous studies demonstrating the possibility of ascertaining whether an individual genotype has contributed to a study where marginal allele frequencies have been reported. Results demonstrate that BOLD significantly improves the imputed LD values. Furthermore, if the study population is a single ethnicity and the reference population contains a mix of ethnicities, the method is effective in giving higher weights to the reference chromosomes with matched ethnicities, and smaller or zero weight to other ethnicities. This addresses the current need for methods to select appropriate reference panels. The LD matrices imputed by BOLD are then compared with standard methods of uniform weighting, and with the ideal case of access to the true LD. In summary, we have shown that BOLD is able to infer LD successfully for a study population, and thus can increase power for gene-based tests conducted in large-scale meta-analyses.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1580W

Statistical models and inference methods for the in vivo DNA methylation process. *A.Q. Fu[1], D.P. Genereux[2], R. Stoeger[3], A.F. Burden[2], C.D. Laird[2], M. Stephens[1,4].* 1) Department of Human Genetics, University of Chicago, Chicago, IL; 2) Department of Biology, University of Washington, Seattle, WA; 3) School of Biosciences, University of Nottingham, Sutton Bonington Campus, Leicestershire, United Kingdom; 4) Department of Statistics, University of Chicago, Chicago, IL.

Single-molecule double-strand DNA methylation patterns collected in vivo using hairpin-bisulfite PCR provide an excellent opportunity to study the DNA methylation process and the properties of DNA methyltransferases in vivo. Each double-stranded methylation pattern contains the binary methylation states of multiple CpG sites on the parent and daughter strands of the DNA molecule from a single cell, although the data do not contain direct information on which strand is the parent strand and which is the daughter. We developed statistical models and Bayesian inference methods to analyze these data, quantifying the transmission fidelity of the DNA methylation process over cell division, as well as the processivity and substrate preference of DNA methyltransferases that are involved in the methylation process. Specifically, our Independent Events Model (IEM) estimates CpG site-specific rates of maintenance and de novo methylation events, as well as the site-site variation in the rates, while allowing for departure from the steady-state. Our Hidden Markov model (HMM) describes the behavior of the three primary DNA methyltransferases identified in mammals: DNMT1, DNMT3A and DNMT3B. The HMM captures the clustering of methylation states across CpG sites in our data, and yields statistical estimates of parameters that quantify enzyme processivity and substrate preference. Both models use multi-site information, infer parent/daughter strand identity, account for experimental errors, and estimate what activities likely gave rise to the observed methylation state at each CpG site. We applied these methods to our data collected from the promoter regions of two densely methylated inactive X-linked loci (FMR1 and G6PD), and that of a moderately methylated autosomal locus (LEP), in humans. Among the key results, applying the IEM to the FMR1 locus reveals a maintenance rate close to, but lower than, 1 per site per cell division, with little site-site variation. It also suggests very low rates of de novo methylation on both strands and more variable de novo rates on the parent strand. Results on DNMT1 from our HMM at all three loci are consistent with existing in vitro results, but suggest longer association tracts (i.e., stronger processivity), and much shorter non-association tracts for humans in vivo than for mice in vitro. The R code for these methods is publicly available.

1581T

A comparison of statistical methods for the analysis of methylome data. *M. LeBlanc[1,2], B. Kulle[1,2].* 1) Epi-Gen, Institute of Clinical Medicine, Akershus University Hospital, University of Oslo and Akershus University Hospital, Oslo, Norway; 2) Department of Biostatistics, University of Oslo, Oslo, Norway.

A major epigenetic modification is the methylation of cytosines (C), most commonly in cytosine-guanine dinucleotides (CpGs) (Rakyan et al. 2011). The methylation profile has been shown to be associated to different complex diseases such as cancer (Portela and Esteller 2010). Due to technical advances, methylation data is increasingly available, but presently there is no consensus on how to analyse this type of data. One interesting problem is the identification and analysis of DNA methylation features based on dependencies in the data. Using a case-control dataset typed using the Roche NimbleGen Human DNA Methylation 2.1M Deluxe Promoter Array, we compare statistical methods for the identification and analysis of DNA methylation features. A key issue is the calculation of p-values for regions in dense methylation data. We compare different techniques based on functional data analysis and permutation methods. Further, we consider how the definition of regions of interest itself influences the results.

1582F

Designing GWAS arrays for efficient imputation-based coverage. *Y. Zhan, Y. Lu, T. Webster, J. Schmidt.* Affymetrix, Inc. Santa Clara, CA.

Imputation has become increasingly popular as a standard step of GWAS data analysis. As a result, imputation-based coverage, instead of coverage based on pairwise tagging only, is becoming a more relevant metric for the value of a GWAS array with respect to the statistical power it offers for genome-wide association tests. However, traditionally whole-genome genotyping arrays have been designed via a tagging marker selection strategy, aiming for good pairwise tagging results, which do not align very well with the goal of having efficient imputation-based coverage. Taking advantage of the genotype data generated by the 1000 Genomes Project, the latest development of imputation tools, and our flexible in-house greedy algorithm for selecting markers based on LD, we came up with a GWAS array design strategy that iterates between an LD-based marker selection step and an imputation-based coverage evaluation step. The coverage evaluation step identifies and removes variants from the target set of markers when their genotypes are imputed well enough based on already selected markers, enabling the selection of markers more valuable for imputation-based coverage in the next iteration. This strategy generates markers for designing genotyping arrays that are better optimized for imputation-based coverage when compared with a set of markers selected with the pairwise-tagging-based selection strategy. Our results show that the new marker selection method makes it possible to design powerful yet efficient GWAS arrays with better imputation-based coverage compared to existing arrays that contain much larger sets of markers.

1583W

Does a higher BMI directly influence levels of physical activity in children? Mendelian Randomization using genotypic scores. *R.C. Richmond[1,2], G. Davey Smith[1,2], D.M. Evans[1,2], B. St Pourcain[1], J.P. Kemp[1,2], A.R. Ness[3], S.M. Ring[1], K. Tilling[1], N.J. Timpson[1,2].* 1) School of Social and Community Medicine, University of Bristol, Bristol, United Kingdom; 2) MRC Centre for Causal Analysis in Translational Epidemiology, University of Bristol, Bristol, UK; 3) School of Oral and Dental Sciences, University of Bristol, Bristol, UK.

Declining levels of physical activity may be partly responsible for rising rates of childhood obesity and many studies have investigated this relationship. However, far fewer studies have considered reverse causation in this association, where a higher body mass index (BMI) leads to a direct reduction in physical activity levels. This was explored in the Avon Study of Parents and Children (ALSPAC), for which there is data on BMI and objectively assessed activity levels (uni-axial actigraph) for children at age 11, as well as genome-wide genotypic data. Using instrumental variable (IV) methods, we aimed to estimate the strength of the unconfounded and unbiased effect of BMI on physical activity. In this context, genetic instruments were used to proxy BMI in a Mendelian Randomization (MR) framework. These analyses were undertaken using established 32 genetic correlates of BMI singularly and combined in a weighted allele score, constructed from directly genotyped SNPs in the ALSPAC sample. The weighted allele score for each individual was derived as a sum of the putative risk alleles at each locus multiplied by their regression coefficients from genome-wide meta-analysis. Using MR, a 3.3 kg/m2 (one standard deviation) increase in BMI was seen to reduce total physical activity by 30.0 counts per minute (95% CI= -60.7,-0.7; P= 0.045), where the mean total physical activity was 607.2 counts per minute. The level of moderate or vigorous activity per day decreased by approximately 2.6 minutes (95% CI= -5.2,-0.1; P= 0.044), where the mean daily level was 20.0 minutes. In addition, sedentary activity was increased by 13.0 minutes (95% CI= 1.1,24.6; P= 0.031), where the mean daily level was 427.3 minutes. Results should be treated with caution until it is has been possible to statistically exclude activity as a mechanistic route to BMI allele score effect, however both crude adjustment for activity and use of randomly drawn combinations of BMI genotypes to form sub scores for sensitivity analyses showed consistent results. Identified associations suggest a causal, independent effect of BMI on activity levels and results have obvious public health implications in terms of tailoring physical activity interventions to children who are already overweight, and ensuring young children establish an active lifestyle before signs of adiposity appear.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1584T

Role of *GATA5*, *FBN1* and *TGFBR2* mutation in patients with bicuspid aortic valve. *N. Abdulkareem, J.A. Aragon-Martin, G. Arno, V. Ramachandran, A. Child, M. Jahangiri.* Cardiac and Vascular Sciences, St. George's University of London, London, United Kingdom.

Background: Bicuspid aortic valve (BAV) is the commonest cardiac defect found in 1–2% of the general population. However, the developmental basis of BAV and its associated complications is poorly understood. We aimed to examine the role of *GATA5, FBN1* and *TGFBR2* in patients with BAV. **Methods:** A consecutive series of 35 UK Caucasian patients were included in our study. BAV was confirmed at surgery. Patients with Marfan syndrome or other diagnosed connective tissue disorders were excluded. Genomic DNA was extracted from venous blood, and PCR was performed to amplify all the coding exons including the intron/exon boundaries of *FBN1, TGFBR2* and *GATA5*. Bi-directional gene sequencing was carried out and the sequences were compared with reference sequence (NM_000138, NM_001024847 and NM_080473) to check for sequence alterations. **Results:** Out of 35 BAV patients, twenty one patients (60%) had concomitant aortic root aneurysm. One patient (2.85%) carried a heterozygous missense mutation in *GATA5* exon-2 (c.698T>C), resulting in amino acid substitution of leucine with proline at 233 position (p.L233P). This change was not present in 100 controls. One heterozygous polymorphism was found in *GATA5* exon-3 (c.8A>G) resulting in amino acid substitution of glutamine with arginine at 3rd position (p.Q3R). This change was found in 2/100 controls of heterozygous state. According to UCSC Genome browser this polymorphism has a frequency >1%. No pathological mutations were found in *FBN1* and *TGFBR2* in BAV cohort. **Conclusion:** For the first time we demonstrate the association of *GATA5* mutation to human phenotype and BAV. The missense mutation p.L233P is located at the DNA binding site and probably affects the function of GATA5 protein. *GATA5* plays an important role in valvulogenesis already demonstrated in *GATA5* mouse model studies.

1585T

Re-sequencing of 3699 individuals reveals novel low-frequency missense polymorphism in *SCN10A* associated with PR interval that alters channel activation. *J.A. Brody[1], V.S. Macri[2], D.E. Arking[3], X. Yin[4,12], C. Liu[4,5], A.C. Morrison[6], A. Alonso[7], J.C. Bis[1], S.R. Heckbert[1,8], T. Lumley[9], C. Sitlani[1], L.A. Cupples[4,5], S. Pulit[2,10], C. Newton-Cheh[2,10], C.J. O'Donnell[4,11], E.J. Benjamin[4,12], D. Muzny[13], R. Gibbs[13], R. Jackson[14], J.W. Magnani[4,15], S.S. Rich[16], B.M. Psaty[1,8,17], E. Boerwinkle[6,13], P.T. Ellinor[2,11], N. Sotoodehnia[1], Cohorts for Heart and Aging Research in Genetic Epidemiology(CHARGE), NHLBI Exome Sequencing Project.* 1) Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, WA, USA; 2) Cardiovascular Research Center, Massachusetts General Hospital, Charlestown, MA, USA; 3) McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA; 4) National Heart, Lung, and Blood Institute and Boston University's Framingham Heart Study, Framingham, MA, USA; 5) Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA; 6) Human Genetics Center, University of Texas Health Science Center at Houston, TX, USA; 7) Division of Epidemiology and Community Health, University of Minnesota, Minneapolis, MN, USA; 8) Department of Epidemiology, University of Washington, Seattle, WA, USA; 9) Department of Statistics, University of Auckland, Auckland, New Zealand; 10) Broad Institute, Cambridge, MA, USA; 11) Cardiology Division, Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA; 12) Boston University Schools of Medicine and Public Health, MA, USA; 13) Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX, USA; 14) Center for Clinical and Translational Science and Division of Endocrinology, Diabetes and Metabolism, The Ohio State University, OH, USA; 15) Section of Cardiovascular Medicine, Boston University School of Medicine, Boston, MA, USA; 16) Center for Public Health Genomics, University of Virginia, VA, USA; 17) Group Health Research Institute, Group Health Cooperative, Seattle, WA, USA.

Background: PR interval from an electrocardiogram reflects cardiac atrioventricular conduction. *SCN10A*, encoding the Nav1.8 sodium channel, was the top locus in several large GWAS of the PR interval, but genetic variation at this locus has not been evaluated systematically. **Methods**: As part of CHARGE-S, a multi-cohort targeted sequencing effort, the 120 kb region containing the *SCN10A* gene was sequenced in 3699 participants of European descent. Adjusting for age, sex, height, BMI, and cohort specific variables, variants with at least 1% minor allele frequency (MAF) were evaluated individually using standard regression methods, while rare variants (MAF ≤ 1%) were tested using Sequence Kernel Association Test. Results were then meta-analyzed across cohorts. We sought to replicate our findings in an independent sample of 526 European descent samples from the Exome Sequencing Project (ESP). The functional impact of missense variants in *SCN10A* was examined by cellular electrophysiology. **Results**: A total of 3209 SNPs were identified, including 6 rare nonsense variants and 128 missense variants. Among the missense SNPs, the MAF was inversely correlated with the functional prediction scores PhyloP (p=1E-5) and Polyphen2 (p=0.015), but no significant association of PR with rare variation (≤1% MAF) was detected. As reported previously, rs6795970 (V1073A; 58% allele frequency) was associated with shorter PR interval (-3.9 ms, p=9E-12). By contrast, rs73062575 (P1045T; 2.5% allele frequency), a previously unrecognized missense variant that is uncorrelated with rs6795970 (r2 < 0.05), was associated with longer PR interval (5.6 ms, p=5E-4). The associations between PR interval and both rs73062575 and rs6795970 were independently replicated in ESP (p=8E-4 and 5E-4, respectively). Both SNPs were introduced into *SCN10A*, expressed in a neuroblastoma 2A cell line and examined by patch-clamp electrophysiology. Compared with wild-type *SCN10A* (4.1 ± 1.7 mV, n=4), both SNPs resulted in a marked change in the midpoint of channel activation. Rs73062575, associated with PR interval prolongation, resulted in decreased channel activation (13.5 ± 2.9 mV, n=5). Conversely, rs6795970, which is associated with shorter PR interval, resulted in increased channel activation (-3.0 ± 1.5 mV, n=4). **Conclusions**: Sequencing of the *SCN10A* locus for the PR interval has led to the identification and replication of a novel missense SNP missed by GWAS that alters *SCN10A* channel activation.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1586T

A functional SNP at APLN gene confers a risk to arterial stiffness in Chinese women. *Y.C Liao[1,2,3], S.W Huang[3], Y.S Wang[3], K.C Chen[3], S.H.H Juo[3,4].* 1) Section of Neurology, Taichung Veterans General Hospital, Taichung, Taiwan; 2) Department of Neurology, National Yang-Ming University, School of Medicine, Taipei, Taiwan; 3) Department of Medical Genetics, College of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan;; 4) Department of Medical Research, Kaohsiung Medical University Hospital, Kaohsiung, Taiwan;.

Background: Apelin (APLN) has been showed to be a potent vasodilator and cardiac inotropic agent. We hypothesized that single nucleotide polymorphisms (SNPs) at APLN gene are associated with susceptibility of arterial stiffness. **Methods and Results**: Stiffness parameters including beta (β), elasticity modulus (Ep) and pulse wave velocity (PWV) were obtained by carotid ultrasonography from the general population. Four tagging SNPs of APLN were initially genotyped in 799 subjects. Student's t test and regression analysis were used to evaluate the genetic effect. Men and women were analyzed separately since APLN was located at the X chromosome. The SNP rs3115757 showed significant associations with the stiffness parameters in women (p = 0.008, 0.032, and 0.036 for β, Ep, and PWV respectively), but not in men (p = 0.13–014). Another group of 937 subjects was genotyped for rs3115757 to validate the results. Combing the screening and validating data, rs3115757 was significantly associated with β, Ep, and PWV in women (p = 0.0003–0.0013). After adjustment of age, hypertension and body mass index (BMI), the SNP remained to be significantly related with stiffness parameters (p = 0.025–0.046). For both screening and validation data, rs3115757 was not related to arterial stiffness in men. In our previous study, the expression levels of APLN varied among human adipocytes carrying different rs3115757 genotypes. The influence of rs3115757 on vasculature needs further investigation. **Conclusions**: APLN polymorphism rs3115757 was significantly associated with carotid artery stiffness in Chinese women.

1587T

Impact of inflammatory gene polymorphisms on left ventricular dysfunction in coronary artery disease patients. *A. Mishra, A. Srivastava, T. Mittal, N. Garg, B. Mittal.* Sanjay Gandhi Post Graduate Institute of Medical Sciences (SGPGIMS), Lucknow, India.

Rationale: Inflammation affects almost all heart failure (HF) patients, provoke a number of deleterious effects on the heart, most notable being left ventricular dysfunction (LVD). However, cause and effect has so far not been proven, but the inhibition of inflammation in HF is expected to be beneficial. A promoter polymorphism of NFKB1 gene (encodes p50 subunit) results in lower protein levels of NFkB p50 subunit and recent evidences suggest that the NFkB p50 dimer has anti-inflammatory effects. The active NFkB transcription factor promotes the expression of over 150 target genes including IL6 and TNF-α. Therefore, the aim of the present study was to assess the association of NFKB1, IL6 and TNF-α gene polymorphisms with LVD in CAD patients in North Indian population. Methods and Results: The present study included a total of 830 subjects (600 CAD patients and 230 healthy controls) and was carried out in two independent cohorts. The primary cohort included 310 consecutive patients with angiographically confirmed CAD. Out of them, 95 with reduced left ventricle ejection fraction (LVEF≤45%) were categorized having LVD. LVEF was calculated quantitatively by echocardiography, just before angiography procedure, using Simpson's method. The NFKB1 -94 ATTG ins/del (rs28362491), IL6 -174 G/C (rs1800795) and TNF-α -308 G/A (rs1800629) polymorphisms were genotyped by PCR/ARMS-PCR methods. Our results showed that NFKB1 ATTG1/ATTG1 i.e. deletion genotype was significantly associated with LVD (primary cohort: p value=0.035; OR=2.44, replication cohort: p=0.040; OR= 2.38). Also, ATTG1/ATTG1 genotype was found to be associated with LV End diastole (p value=0.013), End systole (p value=0.011) dimensions, LV mass (p value=0.024), mean LVEF (p value=0.001) and Myocardial infarction (p value=0.038) but not with risk factors like diabetes mellitus, hypertension and smoking. In contrast, IL6 -174 G/C and TNF-α -308 G/A polymorphisms were not found to be associated with either CAD or LVD both at genotypic and allelic levels. The above findings were validated in an additional 290 cases with similar clinical characteristics. Conclusion: Our data suggest that NFKB1 -94 ATTG ins/del polymorphism plays significant role in conferring susceptibility of LVD and ATTG1/ATTG1 genotype may modulate disease severity by increasing ventricular remodeling and worsening LV function. Financial support from DBT (India) and ICMR.

1588T

Identification Of A New Chromosomal Locus for A Mutation Causing Left Ventricular Non-Compaction With Ventricular Tachycardia Cardio-Pathology. *E. Muhammad[1], A. Levitas[2], V. Sheffield[3], R. Parvari[1,4].* 1) Developmental Genetics, Ben-Gurion University, Beer-Sheva, Israel; 2) Soroka Medcal Center, Ben Gurion University of the Negev, Beer Sheva; 3) University of Iowa, Iowa City, IA, United States; 4) National Institute of Biotechnology in the Negev, Beer Sheva, Israel.

Structural and functional disorders of the heart are important causes of morbidity and mortality. Three Bedouin patients at ages 14–18 years of a single large consanguineous Bedouin family presented with Left ventricular non-compaction and sustained Ventricular Tachycardia. They were evaluated by Echocardiography which showed severe left side enlargement, severely depressed left ventricular (LV) function with focal lacunas in the LV free wall, but normal origin of the coronary artery. The 2 older patients were treated with recurrent electrical cardio version and intravenous administration of amiodarone and on discharge have undergone implantable cardioverter-defibrillator (ICD) implantation. The recessive pattern of inheritance in the consanguineous family suggested homozygosity of the mutation inherited from a common founder. We have performed homozygosity mapping using the affymetrix SNP 5 array on the patients and found a single large chromosomal block of homozygosity on chromosome 11q24 shared by all of them. Verification of this block with VNTR markers in all available family members proved linkage. The Lod score analysis, under the model of a recessive trait with full penetrance, using the PedTool server, was 2.46 for 2 point with marker D11S4958 and 4.59 for the multipoint analysis. This chromosomal locus does not contain any of the known genes causing this disease; efforts are now being done to identify the mutated gene.

1589T

Sequencing candidate genes associated with conotruncal heart defects. *K. Osoegawa[1], M. Ladner[1], K. Schultz[1], C. Parodi[1], N. Mohammed[1], D. Noonan[1], G. Shaw[2], E. Trachtenberg[1], E. Lammer[1].* 1) CHORI, Oakland, CA; 2) Pediatrics, Stanford University, Palo Alto, CA.

Congenital heart defects are one of the most common anatomical groupings of human birth defects, yet little is known about their etiologies. Conotruncal defects are an important pathogenetic subset of congenital heart defects, comprising nearly 20 percent of the total. DNA from a population-based sample of 389 California infants born during 1999–2004 with conotruncal defects were used for sequencing candidate genes for conotruncal defects. We developed an efficient and cost effective strategy for sequencing candidate genes associated with conotruncal defects. We selected FGF8, FOXH1, NKX2-5, GATA4, GATA5, GATA6, ISL1, ZFPM2, MEF2C and CRKL that are either expressed in secondary heart field, for which mutations have been reported in a few infants with conotruncal defects, or identified via array comparative genomic hybridization (array-CGH). Ninety-six DNA fragments or “amplicons” containing exons from these genes were amplified using Fluidigm's Access Array System. The system enabled us to amplify up to 48 samples, each with a unique multiplex identifier, x 48 amplicons (2304 independent PCRs) in a single chip. The amplicons were pooled, and sequenced using the Roche 454 GS FLX system. Resulting DNA sequences were aligned against reference human DNA sequences, and potential novel variations or single nucleotide polymorphisms (SNPs) identified within protein coding sequences using SeqNext software (JSI Medical Systems). SNPs were then confirmed by the Sanger DNA sequencing method. We eliminated rare SNPs that could be found in the 1000 Genome project and NHLBI Exome Sequencing Project (ESP) databases. We identified four subjects with novel mutations in GATA4 (H167Y), ZFPM2 (G164R and S210R) and GATA5 (L394P). In addition to the initial 10 candidate genes, we initiated sequencing of FAK, MAPK1, SOX7, SNAI2 and BMPR1A using the same study population and strategy. These genes were selected based on the identification of chromosomal microdeletions/duplications using array-CGH and gene prioritization.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1590T

Significant Association of Two Genetic Variants in the ENPEP Gene with Lone Atrial Fibrillation in Koreans. *D. Shin[1,3], H. Hwang[2], A. Park[3], N. Son[3,4], E. Shin[5], J. Lee[5], H. Pak[2], M. Lee[2], Y. Jang[2,3].* 1) Cardiovascular Research Institute, Yonsei Univrsity College of Medicine, Seoul, South Korea; 2) Division of Cardiology, Yonsei University Health System; 3) Cardiovascular Genome Center; 4) Department of Biostatistics, Yonsei University College of Medicine; 5) DNA Link Inc, Seoul, South Korea.

Atrial fibrillation (AF) is one of the most common cardiac arrhythmias by chaotic electrical activity of the atrial and is associated with increasing disability and mortality. Up to date, several evidences indicate the phenotypes of lone AF are more likely to occur in patient with a genetic susceptibility. The ENPEP gene has a crucial function in cardiac development and may be one underlying mechanism for AF development. The ENPEP gene encodes an aminopeptidase A (APA) ectopeptidase, which is responsible for the breakdown of angiotensin II in the vascular endothelium and a 20 exon gene located at human chromosome 4q25. We examined a case-control study for a total of 754 Korean subjects, comprising of 354 cases with AF and 400 controls. Genotyping for variants was carried out by direct DNA sequencing and TaqMan® assay. We examined 13 tagging single nucleotide polymorphisms (tSNPs) in the ENPEP gene. We observed the significant difference in the allele and genotype distributions of rs669262 between patients with lone AF and controls (P=0.011). Logistic regression analysis represented that tSNPs, rs669262 and rs2029073, confer the risk of lone AF. The significant associations were observed under a dominant model for rs669262 with an increasing risk of AF (OR, 1.769; 95% CI, 1.174-2.666; P=0.006). In addition, ENPEP rs2029073 showed significant association in Korean AF condition under a recessive model with an increasing risk (OR, 1.583; 95% CI, 1.077-2.327; P=0.019). We provide evidence that ENPEP novel variants may contribute to predispose Korean individuals to the risk of lone AF.

1591T

Molecular Investigation in Sudden Unexplained Death Syndromes in New York City. *Y. Tang[1], D. Wang[1], L. Eng[1], K. Shah[1], S. Um[1], B. Zhou[1], M. Prinz[1], B. Sampson[2].* 1) Forensic Biol, Office Chief Med Examiner NYC, New York, NY; 2) Forensic Pathology, Office Chief Med Examiner NYC, New York, NY.

Objective: Sudden unexplained death syndrome **(SUDS)** is sudden death in apparently healthy individuals which remains unexplained after a thorough investigation including a complete autopsy, examination of the death scene and review of clinical history. It is estimated that more than 8,500 such deaths occur in the United States every year. Currently, molecular genetic testing is the only means to diagnose certain cardiac arrhythmias that can result in SUDS, such as long QT syndrome. We performed molecular studies on 323 SUDS cases. **Methods and Results:** 323 SUDS cases, which were investigated from June 2001 to December 2011 at the City of New York Office of Chief Medical Examiner (NYC OCME), were included. Forensic examinations were conducted, including a full autopsy, toxicological testing, histological examination of major organs, review of available clinical history, and a thorough scene investigation. Molecular studies were performed by the genetics laboratory within OCME, accredited by College of American Pathologists (CAP). Six major cardiac channel genes (KCNQ1, KCNH2, SCN5A, KCNE1, KCNE2 and RyR2) previously reported to have disease-causing sequence variants in SUDS were analyzed. Entire coding regions were tested for KCNQ1, KCNH2, KCNE1, KCNE2, and selected exons were tested for SCN5A (exons 12 to 28) and RyR2 (exon 8, 14, 15, 44–47, 49, 88–93, 95–97, 100–105). Baseline characteristics of SUDS were compared amongst ethnic groups (Blacks, Hispanics, Whites, and Asian), by gender, and between various age groups (<1 year old, 1–5 years old, 6–18 years old, and >18 years old). Sequence variants were classified by relevance to clinical significance in the functional domain of each gene by age group. Approximately 15% of SUDS cases may be explained by the presence of possible disease-causing sequence variants in the decedents. **Conclusion:** Molecular testing has proven to be a valuable tool in the forensic investigation of SUDS. Expanded genome-wide testing is expected to increase our ability to elucidate the cause of death in the future.

1592T

Associations of NINJ2 sequence variation with incident ischemic stroke: analysis of common and rare variants in the Cohorts for Heart and Aging in Genomic Epidemiology (CHARGE) Consortium. *J.C. Bis[1], M. Fornage[2,3], A.L. DeStefano[4], J. Brody[1], M.A. Ikram[5,6], C. van Duijn[5], R. Gibbs[7], J.G. Reid[7], E. Boerwinkle[2,3,7], W.T. Longstreth[1,8], T.H. Mosley[9], S. Seshadri[10] on behalf of the CHARGE Neurology Working Group.* 1) Cardiovascular Health Research Unit and Department of Medicine, University of Washington, Seattle, WA, USA; 2) 2. Brown Foundation Institute of Molecular Medicine, Division of Epidemiology, School of Public Health, University of Texas Health Science Center at Houston, Houston, TX, USA; 3) Human Genetics Center, Division of Epidemiology, School of Public Health, University of Texas Health Science Center at Houston, Houston, TX, USA; 4) Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA; 5) Department of Epidemiology; Erasmus MC Rotterdam, The Netherlands; 6) Departments of Radiology, Neurology; Erasmus MC Rotterdam, The Netherlands; 7) Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX, USA; 8) Department of Neurology, University of Washington, Seattle, WA, USA; 9) Department of Medicine (Geriatrics), University of Mississippi Medical Center, Jackson, MS, USA; 10) Department of Neurology, Boston University School of Medicine, Boston, MA, USA.

Stroke is the leading neurologic cause of death and disability. Despite studies suggesting a substantial genetic component, few genes underlying this risk in the general population have been identified. In a previous genome-wide association meta-analysis in the setting of the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium, we identified and replicated associations of single-nucleotide polymorphisms adjacent to the *NINJ2* gene with incident stroke ($P<5\times10^{-8}$). We therefore sequenced a 200kb window spanning *NINJ2* (12p13) among 3,979 European ancestry participants from the Atherosclerosis Risk in Communities Study, the Cardiovascular Health Study, and the Framingham Heart Study that included a random sample as well as a subset of 200 individuals who had had an ischemic stroke of atherothrombotic origin. We applied two methods of analysis for this region. Common variants, those with minor allele frequency (MAF) ≥ 1%, were analyzed independently with a fixed effect meta-analysis on each variant. Rare variants (MAF < 1%) were meta-analyzed in aggregate using the T1 count statistic, defined as the number of loci in the target window with MAF less than 1% at which a person has at least one rare allele. All associations of genetic variation with ischemic stroke were estimated using Cox proportional hazards models with adjustment for age and sex. While this targeted sample provided evidence of association for the original GWAS signals (rs11833579, p=0.0005; rs12425791, p=0.006), the most significant association in the meta-analysis of 290 common variants was a novel intronic SNP (MAF=0.012, p=0.0003) in the *NINJ2* gene, providing fine-mapping evidence that the responsible gene within the broad GWAS peak may be *NINJ2*. Further analysis of the 2,117 rare variants showed modest evidence that, in aggregate, rare variation in this gene also contributed to stroke risk (p=0.03). Sequencing of the *NINJ2* regions within the CHARGE cohorts demonstrated evidence of association between common and rare sequence variants and ischemic stroke. While additional investigation is needed to understand the role of these newly discovered variants, these findings may help clarify the genetic architecture of a region previously implicated in the incidence of stroke.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1593T

CCR5, TNF/LTA, CCL2, IL12B, IL10, MAL/TIRAP, ACTC1 genes are associated to the development of chronic Chagas' disease cardiomyopathy. *C. CHEVILLARD[1], A.F. FRADE[2], P. TEIXEIRA[2], B. IANNI[3], C. WIDE PISSETTI[4], A. FRAGATA[5], M. HIRATA[5], M. SAMPAIO[5], B. SABA[5], F. DIAS[6], E. DONADI[6], V. RODRIGUES JR[4], A. PEREIRA[3], J. KALIL[2], E. CUNHA-NETO[2].* 1) INSERM, U906, Faculté de médecine, Marseille, France; 2) Heart institute (Incor), University de Sao Paulo, Av. Dr. Enéas de Carvalho Aguiar, 05403-000 Sao Paulo, Brazil; 3) Hospital das Clinicas, Av. Dr. Enéas de Carvalho Aguiar, 255 Cerqueira César, 05403-000 Sao Paulo, Brazil; 4) Laboratory of Immunology, Universidade Federal do Triângulo Mineiro, 30 Frei Paulino, 38025-180 Uberaba, Brazil; 5) The Instituto Dante Pazzanese, Avenida Dr. Dante Pazzanese, 500 Vila Mariana, 04012-909 Sao Paulo, Brazil; 6) Faculdade de Ribeirao Preto, Av. Bandeirantes, 3900 - Monte Alegre, 14049-900 Ribeirao Preto, Brazil.

Chagas disease, due to Trypanosoma cruzi, occurs exclusively in the Americas, particularly in poor, rural areas of Mexico, Central America, and South America. An estimated 300,000 new cases and 50,000 fatilities occur per year. Chronic Chagas disease cardiomyopathy (CCC) is an inflammatory cardiomyopathy that affects approximately 30% of infected individuals. Approximately 1/3 of patients developing CCC present a particularly lethal form of dilated cardiomyopathy, with short survival. Familial aggregation of CCC suggests that there might be a genetic component to disease susceptibility. Here, we conduct a study on several main candidate genes on a large Brazilian population. This study was done on 315 CCC cases and 118 asymptomatic subjects considered as “controls”. Significant associations were detected for CCR5 (rs3176763: p=0.006 OR=1.79; rs11575815: p=0.030 OR=1.41), TNF/LTA (rs2844484: p=0.007 OR=1.51; rs909253: p=0.035; OR=1.50), CCL2 (rs909253: p=0.037 OR=1.56; rs2530797: p=0.028 OR=1.28), IL12B (rs1003199: p=0.008 OR=1.49; rs919766: p=0.007 OR=1.45), IL10 (rs3024496: p=0.020 OR=1.47), MAL/TIRAP (rs8177376: p=0.004 OR=1.42) genes. These associations are consistent with the fact patients who develop Chagas cardiomyopathy display an exacerbated Th1 immune response and a reduced numbers of IL-10-producing and CD4+CD25+FOXP3+ regulatory T cells. Cardiac muscle alpha Actin 1 is a protein that in humans is encoded by the ACTC1 gene. On heart biopsies from cases and controls, proteomic analysis, performed by 2-D DIGE experiments, has shown that the ACTC1 proteins are under represented into patient heart biopsies compared to control heart biopsies. This result was confirmed by western blot analysis. Some significant associations were detected with one polymorphism located into the promoter (rs640249: p=0.006 OR=1.59). Based on linkage disequilibrium map, a promoter haplotype of susceptibility was identified (rs640249-rs641563). Some functional studies done on primary human cardiomyocytes and an imputation analysis are ongoing in order to identify the functional variants. Authors hypothesized that ACTC1 mutations affecting sarcomere contraction leading to cardiomyopathy and that mutations affecting force transmission from the sarcomere to the surrounding syncytium lead to dilated cardiomyopathy.

1594T

Analysis of Candidate Genes Associated with Intermediate Traits Identifies Novel Variants Associated with Coronary Artery Disease Presence and Severity. *D.M. Craig[1], E.R. Hauser[1], E. Burns[1], M. Chryst-Ladd[1], C. Haynes[1], W.E. Kraus[1], I.C. Siegler[2], B.H. Brummett[2], R.C. Becker[2], R.B. Williams[2], S.H. Shah[1].* 1) Medicine, Duke Medical Center, Durham, NC; 2) Psychiatry, Duke Medical Center, Durham NC.

Background. Cardiovascular disease (CVD) is heritable but the genetic architecture remains incompletely understood. Genetic variants have been associated with endophenotypes (i.e. diabetes, hypertension) that are themselves CVD risk factors. We hypothesized that these genetic variants are also associated with CVD. Methods. A sequential cohort of CATHGEN Biorepository patients referred for cardiac catheterization to assess ischemic heart disease was identified. Eleven single nucleotide polymorphisms (SNPs) were genotyped using the Taqman platform. CVD was defined by three phenotypes: coronary artery disease (CAD, defined as 1+ epicardial coronary vessel ≥75% stenosis), degree of CAD (number of diseased epicardial vessels), and time to incident CVD events. Association between SNP and CVD was tested using logistic regression with a basic (age, sex) and multivariable model (age, sex, body mass, hypertension, diabetes, hyperlipidemia, ejection fraction, smoking) stratified by race (N=5092 Caucasian, N=1254 African-Americans) using dominant and additive genetic models. Time to incident CVD was modeled using Cox proportional hazards. Results. All SNPs were in Hardy-Weinberg equilibrium. In Caucasians, the T allele of rs2943641 (near insulin receptor substrate 1 [IRS1] gene) was associated with degree of CAD (basic model odds ratio [OR] 0.91, 95% CI 0.83, 0.99; p=0.04), with attenuation after adjustment in the multivariable model (OR 0.91, 95% CI 0.83, 1.01; p=0.06). Rs2943641 was not associated with CAD as a binary trait, and no other SNPs were associated with any CVD phenotype in Caucasians. In African Americans, there was association between the C allele of rs2131127 (in angiotensin II receptor, type 1 [AGTR1]) and CAD (basic OR 1.25, 95% CI 0.99, 1.58; p=0.07; multivariable OR 1.33, 95% CI 1.02, 1.72; p=0.03) and trend for association with degree of CAD (basic OR 1.11, 95% CI 0.91, 1.34; p=0.3; multivariable OR 1.19, 95% CI 0.97, 1.47; p=0.09). No SNPs were associated with CVD event times. Conclusions. We found evidence for modest association of a SNP near IRS1, previously associated with type II diabetes, and for a SNP in AGTR1 with CVD phenotypes. This latter finding is consistent with a previous study linking the rs2131127 TT genotype to lower blood pressure reactivity to stress and platelet aggregation. Genetic variants associated with CVD-related endophenotypes may serve as CVD risk markers, but with only modest effects.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1595T

The Shared Allelic Architecture of Adiponectin Levels and Coronary Artery Disease. *Z. Dastani[1], T. Johnson[2,3,4], F. Kronenberg[5], C. Nelson[6,7], T.L. Assimes[8], W. März[9], J.B. Richards[10,11], CARDIoGRAM Consortium, ADIPOGen Consortium.* 1) CanadaDepartment of Epidemiology, Biostatistics and Occupational Health. Jewish General Hospital, Lady Davis Institute, McGill University Montreal, Quebec H3T 1E2, Canada; 2) Department of Medical Genetics, University of Lausanne, 1005 Lausanne, Switzerland; 3) University Institute of Social and Preventative Medicine, Centre Hospitalier Universitaire Vaudois (CHUV) and University of Lausanne, Lausanne, Switzerland; 4) Swiss Institute of Bioinformatics, Lausanne, Switzerland; 5) Division of Genetic Epidemiology, Innsbruck Medical University, 6020 Innsbruck, Austria; 6) Department of Cardiovascular Sciences, University of Leicester, Leicester LE3 9QP, UK; 7) National Institute for Health Research (NIHR) Leicester Cardiovascular Disease Biomedical Research Unit, Glenfield Hospital, Leicester LE3 9QP, UK; 8) Department of Medicine, Stanford University School of Medicine, Stanford, CA; 9) Synlab Center of Laboratory Diagnostics Heidelberg, Eppelheim, Germany; 10) Department of Twin Research and Genetic Epidemiology, King's College London, London SE1 7EH, UK; 11) Departments of Medicine, Human Genetics, Epidemiology and Biostatistics. Lady Davis Institute, Jewish General Hospital, McGill University Montreal, Quebec H3T 1E2, Canada.

A large body of epidemiologic, clinical and experimental data strongly suggests an association between excess adiposity and CAD. Low levels of adiponectin, a hormone secreted only from adipocytes, has been associated with an increased risk of CAD in observational studies. However, these associations cannot clarify whether this relationship is causal or due to a shared set of causal factors or even a consequence of reverse causality meaning that CAD secondarily results in low adiponectin levels. Genetic influences account for a large proportion of the variation in adiponectin levels and several genome-wide association studies have identified numerous common variants across the genome that influence adiponectin levels. These variants provide a valuable tool to examine the genetic relationship between adiponectin and CAD, enabling insights into the causality of this relationship. Using 145 genome-wide significant SNPs for adiponectin levels from the ADIPOGen consortium (n=49,891) we aimed to test whether adiponectin-decreasing alleles influenced risk of CAD in the CADRIoGRAM consortium (n=82,000). In single-SNP analysis, 5 variants among 145 SNPs were associated with increased risk of CAD after correcting for multiple testing (P < 4.4×10-4). Using independent SNPs from145 SNPs, we applied a multi-SNP genotypic risk score to test whether adiponectin levels and CAD have a shared genetic etiology, we found that adiponectin-decreasing alleles increased risk of CAD (P = 5.4×10-7). These findings demonstrate that adiponectin levels and CAD have a common allelic architecture and provide rationale to undertake a Mendelian randomization to understand if low adiponectin levels cause CAD.

1596T

Investigation of mutations in exons 19 and 22 MYH7 gene in HCM patients in Chaharmahal Va Bakhtiyari province, Iran. *S. Heydari Sodejani[1], R. Pourahmad[1], A. Khaledifar[2], M. Hashemzadeh[3], F. Azadegan[4], N. Bagheri[5], S. Badfar[6].* 1) Department of Genetics, shahrekord university, Shahrekord, chaharmahal va bakhtiyari, Iran; 2) Department of Cardiology, School of Medicine, Sharekord University of Medical Sciences, chaharmahal va bakhtiyari, Iran; 3) Department of Genetics, Molecular and Cellular Research Center, Sharekord University of Medical Sciences, chaharmahal va bakhtiyari, Iran; 4) Department of Genetics,School of Basic Sciences, Zabol University, Sistan Va Balochestan, Iran; 5) Department of Immunology, Sharekord University of Medical Sciences, chaharmahal va bakhtiyari, Iran; 6) Echocardiography Unit, Hajar Hospital of Shahrekord, chaharmahal va bakhtiyari, Iran.

Abstract: Background and Aim: Hypertrophic cardiomyopathy (HCM) is the most common kind of Mendelian inherited heart disease, affects 0.2% of the global population. HCM is also the most common cause of sudden cardiac death in individuals younger than 35 years old. To date more than 900 individual mutations has been identified in over 20 genes, such as MYH7, MYBPC3, and TNNT2. Interestingly, most of these genes encode sarcomeric proteins. In the present study, we investigated the possible presence of mutations in the 19 and 22th exons of the MYH7 gen, reported to accommodate the maximum mutations, in a cohort of 30 patients with HCM in Chaharmahal Va Bakhtiyari province. Materials and Methods: DNA was extracted using standard phenol- chloroform method and then was investigated by PCR- SSCP procedure. Finally, the suspected cases were selected for the direct sequencing and the results were analysed using chromas, BLAST and Clustal W online servers. Results: There is no mutation in the 19th exon but were found tow mutation in the 22th exon that including: a substitution G>A 10824 and an insertion Ins 10864G. Discussion and Conclusion: Two of probands had nucleotide changes. One of them was a 60 years old woman with family history of HCM and has Ins 10864G mutation. Another is a 73 years old women without family history and showed the G>A 10824 mutation. Our data indicated that mutation in the MYH7 gene can be found in patients without a family history of HCM. This is necessary the study of other exons to better assessment.

1597T

Polymorphisms in *CHRNA3-CHRNA5-CHRNB4* are associated with body mass index and systolic blood pressure in smokers in the Northern Finland Birth Cohort 1966. *M. Kaakinen[1,2], F. Ducci[3,4], M.J. Sillanpää[1,2], E. Läärä[1], M-R. Järvelin[1,2,5,6].* 1) University of Oulu, Oulu, Finland; 2) Biocenter Oulu, Oulu, Finland; 3) King's College, London, United Kingdom; 4) St George's University of London, London, United Kingdom; 5) National Institute for Health and Welfare, Oulu, Finland; 6) Imperial College London, London, United Kingdom.

The *CHRNA5-CHRNA3-CHRNB4* gene cluster on 15q25 has consistently been associated with smoking quantity, nicotine dependence, lung cancer and chronic obstructive pulmonary disease. Evidence for the involvement of this gene cluster on regulation of cardiovascular homeostasis has also been reported. One twin study reported an association between genetic variation within *CHRNA5-CHRNA3-CHRNB4* and systolic blood pressure, whereas results from two meta-analyses indicated that polymorphisms in *CHRNA3* influence body mass index and susceptibility to develop peripheral arterial disease, but only among smokers. More research is needed to examine the potential pleiotropic effects of *CHRNA3-CHRNA5-CHRNB4* gene cluster on blood pressure and BMI, and how smoking behavior may modify these effects. We studied the associations between 18 variants in the gene cluster and systolic blood pressure (SBP), diastolic blood pressure (DBP), and body mass index (BMI) in 5402 young adults from the Northern Finland Birth Cohort 1966. The analyses were performed separately in non-smokers and smokers, and finally in the entire sample including an interaction term for smoking. Several single nucleotide polymorphisms from the gene cluster were associated with SBP and BMI, while evidence for associations with DBP was weaker. The associations were manifest only in smokers, and in such a way that the smoking-associated alleles predicted lower SBP and BMI. Each additional copy of the rs1948 G-allele and rs950776 A-allele in *CHRNB4* were associated with about 1.20 mmHg lower SBP in smokers. The variants associated with BMI included rs2036534 in *LOC123688* and rs6495309, rs1996371, rs6495314, rs4887077, rs11638372 in *CHRNB4* and had an average effect size of -0.35 kg/m^2 per an additional copy of the risk allele in smokers. The results from interaction models provided further evidence of the interactive effects observed in stratified analyses. The associations with the phenotypes in smokers were driven by different loci with low linkage disequilibrium with each other suggesting that the gene cluster regulates SBP through biological mechanisms that partially differ from those regulating BMI. Large collaborative efforts are needed to find further evidence for the associations, to test for their existence in other populations and to understand the mechanisms by which the gene cluster influences BMI and SBP.

1598T

Genetic association studies of left ventricle dysfunction in coronary artery disease patients. *B. Mittal, A. Mishra, A. Srivastava, T. Mittal, N. Garg.* Sanjay Gandhi Med Inst, Lucknow, UP, India.

Rationale: Left ventricular dysfunction (LVD) often develops in response to an ischemic insult, ventricular remodeling and sarcomeric disorganization. The structural and extracellular matrix proteins can contribute to unfavourable ventricular remodelling followed by fall in cardiac output, which further leads to activation of Renin Aldosterone Angiotensin System (RAAS). Therefore, the present study was undertaken to evaluate association of genetic variations in cardiac myosin binding protein-C (MYBPC3), RAAS, and Matrix metalloproteinases (MMPs) with LVD in coronary artery disease (CAD) patients in North Indian population. Methods and Results: The present study was carried out in two cohorts. The primary cohort included 310 consecutive patients with angiographically confirmed CAD and 230 controls. Among CAD patients, 95 with reduced left ventricle ejection fraction (LVEF≤45) were categorized as LVD. LVEF was calculated quantitatively by echocardiography, just before angiography procedure, using Simpson's method. The 25 bp deletion in MYBPC3 gene and genetic variants in RAAS [angiotensin I converting enzyme; ACE I/D (rs4340), angiotensin II type1 receptor; AT1A1166C (rs5186), aldosterone synthase; CYP11B2 T-344C (rs1799998)] and MMPs [MMP2 C-735T (rs2285053), MMP7 A-181G (rs11568818) and MMP9 R279Q (rs17576), P574R (rs2250889) and R668Q (rs17577)] were determined by PCR/ARMS-PCR/PCR-RFLP methods. Our results showed that 25bp deletion in MYBPC3 (p value <0.001; OR = 4.49), AT1 A1166C (p value=0.013; OR=3.69) and MMP9 R668Q (p value=0.009; OR=3.82) were significantly associated with LVD. Also the frequency of haplotype MMP9 279R, 574P, 668Q was significantly higher in reduced LVEF subjects (p value = 0.008; OR = 1.83). We further performed a replication study in additional 200 cases with similar clinical characteristics and results again confirmed consistent findings: 25bp deletion in MYBPC3 (p= 0.029; OR=3.30), AT1 A1166C (p value=0.020; OR=5.20) and MMP9 R668Q (p value=0.033; OR=3.59) were significantly associated with LVD. Conclusion: The 25bp deletion in MYBPC3, AT1 A1166C and MMP9 R668Q were significantly associated with LVD in CAD patients. These findings may be important for the understanding of genetics of left ventricular dysfunction which is leading cause of heart failure. Financial support from DBT (India) and ICMR.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1599T

Associations of sequenced regions SLC17A4 and PIK3CG with common carotid intima-media thickness and plaque: a common and rare variant meta-analysis in the Cohorts for Heart and Aging in Genomic Epidemiology (CHARGE) Consortium. *C.J. O'Donnell[1,2], C.C. White[1,3], N. Franceschini[4], J. Brody[5], D. Muzny[6], R. Gibbs[6], E. Boerwinkle[7], B. Psaty[5,8,9,10], K.E. North[11], L.A. Cupples[1,3], J.C. Bis[5] on behalf of the CHARGE Subclinical Atherosclerosis Working Group.* 1) National Heart, Lung and Blood Institute's Framingham Heart Study, Framingham, MA, USA; 2) Cardiology Division, Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA; 3) Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA; 4) Department of Epidemiology, University of North Carolina Chapel Hill, Chapel Hill, NC, USA; 5) Cardiovascular Health Research Unit and Department of Medicine, University of Washington, Seattle, WA, USA; 6) Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX, USA; 7) Human Genetics Center, University of Texas Health Science Center at Houston, TX, USA; 8) Department of Epidemiology, University of Washington, Seattle, WA, USA; 9) Department of Health Services, University of Washington, Seattle, WA, USA; 10) Group Health Research Institute, Group Health, Seattle, WA, USA; 11) Carolina Center for Genome Sciences, University of North Carolina, Chapel Hill, NC, USA.

Atherosclerosis, the precursor to clinical coronary heart disease and stroke, is caused by an accumulation of fatty materials in the intimal-medial layers of the arteries. Common carotid intima media thickness (cca-IMT) and plaque, observed with ultrasonography, are subclinical atherosclerosis measures that predict cardiovascular disease events. Identifying genomic locations associated with cca-IMT and plaque might shed light on the mechanism of atherosclerosis and identify people at risk of cardiovascular disease. For this project, we sought additional information about two regions that previously demonstrated evidence for association with cca-IMT (*SLC17A4*) and plaque (*PIK3CG*) in a large-scale genome-wide meta analysis. We sequenced windows of 120kb around *SLC17A4* (6p22.2) and 251kb around PIK3CG (7q22.3) among 3,688 European ancestry participants from the Atherosclerosis Risk in Communities Study, the Cardiovascular Health Study, and the Framingham Heart Study that included a random sample as well as a subset of individuals with selected for high values of cca-IMT. We applied two methods of analysis for these regions. Common variants, those with minor allele frequency (MAF) ≥ 1%, were analyzed independently with a fixed effect meta-analysis on each variant. Rare variants (MAF<1%) were meta-analyzed in aggregate using the T1 count statistic, defined as the number of loci in the target with MAF<1% at which a person has at least one rare allele. In the meta-analysis of common variants, rs2286149 showed the smallest p-value in both analyses of cca-IMT (p-value= 0.0003) and plaque (p-value=0.0003). This variant falls in the 3' UTR of *CCDC71L*, itself upstream from the targeted *PIK3CG* gene. For the region spanning SLC17A4, the strongest finding was for a newly-discovered intronic variant with cca-IMT (p=0.008), which may provide modest evidence toward fine-mapping the broad GWAS peak. Further analysis of rare variants showed no evidence of association between the either region with cca-IMT, but found suggestive association of the *PIK3CG* region with plaque (p=0.054). Sequencing of the *PIK3CG* and *SLC17A4* regions within the CHARGE cohorts shows further evidence of association between plaque and the *PIK3CG* region for both common and rare variants. While not conclusive, these findings may help to better understand the genetic architecture of two regions previously implicated in subclinical atherosclerosis.

1600T

CNVs contributing to the cause of congenital heart defects may include not only those containing candidate genes but regions more commonly associated with autism and other disorders. *D. Warburton[1,6], M. Ronemus[2], J. Kline[3,4], M. Wigler[2], V. Jobanputra[5], D. Levy[2], K. Anyane-Yeboa[6], W. Chung[6], D. Awad[7].* 1) Dept. Gen & Development, Columbia Univ, New York, NY; 2) Cold Spring Harbor Laboratory, Cold Spring, NY; 3) NY Psychiatric Institute, New York, NY; 4) Mailman School of Public Health, Sergievsky Center, Columbia University, New York, NY; 5) Dept. Pathology, Columbia University, New York, NY; 6) Dept. of Pediatrics, Columbia University, New York, NY; 7) Dept. of Medicine, Columbia University, New York, NY.

Our study aimed to determine the contribution of CNVs to congenital heart disease (CHD) of two distinct types. We recruited 73 probands and their parents (trios) with hypoplastic left heart syndrome (HLHS) and 146 trios with conotruncal anomalies (CNT). All cases of DiGeorge syndrome were omitted. Using the Nimblegen 2.1 chip for aCGH and analytical software designed at Cold Spring Harbor Laboratory (CSHL), our primary focus was to determine the frequency and gene content of de novo (DN) CNVs in both types of anomalies. All DN lesions, ranging from 7 kb to 8.6Mb, were verified in probands and parents, by FISH or qPCR. The frequency of trios with DN lesions was higher for HLHS (12.3%) than for CNT (7.4%), and higher for male probands (11.3%) than for female probands, but neither difference is statistically significant. The frequency of de novo lesions in CHD is significantly higher than that in the 700 control families using the same platform and software (2%), indicating that a high proportion of the observed DN changes are contributing to the cause of the CHD. Five DN lesions involved genes or regions known to be associated with CHD or other pathology. These include a deletion in 8p23.1 containing GATA4, the common duplications in 16p13.11 and 1q21.1, a deletion in 2q23.1 including MBD5, and a 6Mb duplication in the Prader-Willi region. The last two have not been associated so far with CHD. We also examined inherited but ultra-rare CNVs (not seen in more than one of the control trios) and examined those that (1) were also found de novo (2) contained candidate genes for CHD (3) occurred more than once (4) and had an affected sib eith the lesion. Within this group we detected a deletion of NODAL and a duplication of CFC1 (genes known to be involved in heterotaxic heart defects) and also three cases of the 1q21.1 duplication and two cases of the 15q11.2 duplications found de novo. In all 213 trios, there were only 5% where the abnormality revealed by CNV analysis is most certainly related to the cause of the CHD. However more studies of clearly defined populations may reveal new repetitive CNVs and lead to identification of new genes involved in CHD. In particular the presence in multiple cases of CHD of lesions well recognized to lead to autism and other developmental disorders (16p13.11, 15q11–12, 2q23.1, 1q21.1) suggests that these areas may affect the activity of a wide variety of developmental processes by interracting with other genetic variation.

1601T

Whole exome sequencing and hypertrophic cardiomyopathy. *F. Dewey, S. Cordero, M. Wheeler, A. Pavlovic, K. Bommakanti, S. Pan, C. Caleshu, R. Chen, M. Snyder, E. Ashley.* Stanford Univ, Stanford, CA.

Background: Hypertrophic cardiomyopathy (HCM) is a common cause of sudden death in young people. Whole exome sequencing may provide novel insight into the pathogenesis of the disorder. **Methods:** We performed whole-exome sequencing on six related subjects in whom target sequencing had not identified a pathogenic mutation: four with manifest HCM, one obligate carrier, and one unaffected adult. First, we excluded common, benign polymorphisms using a combination of publicly available controls from the 1000 genomes variation database, dbSNP, and sequence data from our laboratory. Second, we used a novel phasing algorithm to identify candidate recent ancestral haplotypes across the three generation pedigree to narrow the search space for pathogenic variants. Non-synonymous or splice disrupting single nucleotide variants (SNVs), and frameshift coding indels (NS/SS/FI) were scored for evolutionary conservation using multiple sequence alignment of 46 mammalian species. Candidate variants were further prioritized by gene expression in myocardial tissue from 436 human subjects with heart failure, 49 human subjects with cardiac hypertrophy, 75 unused, structurally normal explanted donor hearts, and in non-cardiac tissues from 1,058 human subjects. **Results:** Sequence data was successfully generated for 97.9% of targeted exonic regions at an average read depth of 134x. We identified 12,649 NS/SS/FI variants in the pedigree, and 892 of these variants were not seen in population controls. Six of these private variants arose on a recent ancestral haplotype that co-segregated with disease. Additional phenotyping of thirteen related individuals revealed two additional affected subjects; targeted sequencing demonstrated co-segregation of three of these six candidate variants with disease. One SNV occurred in a gene that was highly and uniquely expressed in human myocardial tissue compared to non-myocardial tissue, wilms tumor interacting protein (WTIP). **Conclusions:** Whole exome sequencing in families can be combined with population variation, evolutionary conservation, linkage data, and gene coexpression for pathogenic gene identification in HCM. Experimental confirmation of the functional role of this variant allele of WTIP may illuminate a novel role for non-cannonical wnt signaling in the pathogenesis of HCM.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1602T

Fine Mapping the *SCN10A* Gene Region Identifies Novel Associations with PR Interval in African Americans from an Electronic Medical Record Population. *J. Jeff[1], M. Ritchie[2], J. Denny[3,4], H. Dilks[5], C. Sutcliffe[5], M. Basford[6], D. Roden[3,6,7], D. Crawford[1].* 1) Human Gen, Vanderbilt Univ, Nashville, TN; 2) Center for Systems Genomics, Penn State University; 3) Department of Medicine, Vanderbilt Univ, Nashville, TN; 4) Department of Biomedical Informatics, Vanderbilt Univ, Nashville, TN; 5) DNA Resources Core, Vanderbilt Univ, Nashville, TN; 6) Office of Personalized Medicine, Vanderbilt Univ, Nashville, TN; 7) Department of Pharmacology, Vanderbilt Univ, Nashville, TN.

Variations in the sodium channel gene, *SCN5A*, have been extensively studied by candidate gene studies in African Americans and are associated prolonged QT syndrome and sudden cardiac death. Recently, genome wide association studies (GWAS) have identified novel loci that contribute to electrocardiographic parameters such as atrioventricular conduction (PR interval). Most GWAS have been performed in European American populations collected from population-based cohorts or surveys. *SCN10A* variants, located in a sodium channel gene cluster (5A-10A-11A) on chromosome 3, have been consistently associated with PR interval by GWAS in Europeans; however, to date these associations fail to generalize in African Americans. In a previous study, GWAS-identified *SCN10A* SNPs in European Americans were rare and not associated with PR interval in African Americans (p>0.10); however, the effect sizes trended in the same direction. To identify alleles associated with PR interval in African Americans, we performed a fine-mapping experiment with variants in the *SCN10A* region with PR interval in 1,016 African Americans, without heart disease identified using electronic medical records. We genotyped 140 SNPs within 165kb of the *SCN10A* gene encompassing *SCN5A*, *SCN10A*, and *SCN11A*. We detected a total of 22 associations with PR interval at a liberal significance threshold of p<0.05. Our most significant association, rs3922844 (*SCN5A*, β=4.33, p= 1.6 × 10-6) has been previously associated with ECG traits in African Americans. In addition to rs3922844, we replicated four other SNPs and identified six novel associations in the *SCN5A* gene. As previously mentioned, there have been no significant associations reported in African Americans with *SCN10A* and none with *SCN11A* for any population. We detected two associations with intronic SNPs in the *SCN10A* (rs6599251, β=1.86, p=0.03; rs9848976, β=-2.18, p = 0.05) and four associations in *SCN11A* with PR interval in our African American sample (rs4284929, β=-2.61, p=7.5 × 10-3; rs4514993, β=-2.60, p=0.03; rs11924818, β=-2.50, p=0.04; rs9818979, β= 1.60, p=0.04). The exact role *SCN10A* and *SCN11A* have on atrioventricular conduction is poorly understood. Given the novelty of these associations, these results set a framework for future association and molecular studies to elucidate the function of these genes and their impact on atrioventricular conduction.

1603T

Association between 4q25 variants and cardioembolic stroke in the Vienna Stroke Registry. *M.M. Luke[1], C.H. Tong[1], J.J. Catanese[1], J.J. Devlin[1], S. Greisenegger[2], C. Mannhalter[3].* 1) Cardiovascular Disease, Celera, Alameda, CA; 2) University Clinic of Neurology, Medical University of Vienna, Vienna, Austria; 3) Clinical Institute of Laboratory Medicine, Medical University of Vienna, Vienna, Austria.

Background—Two single nucleotide polymorphisms (SNPs) on chromosome 4q25 (rs2200733 and rs10033464) have previously been reported to be associated with ischemic stroke—in particular with the cardioembolic stroke (CES) subtype. Aims—We evaluated the association of rs2200733 and rs10033464, both separately and in combination, with CES. Methods—The genotypes of rs2200733 and rs10033464 were determined for 202 CES cases and 815 controls in the Vienna Stroke Registry (VSR) case-control study. We then tested the two SNPs both separately and in combination for association with CES in regression models. We also estimated the effect sizes for the two SNPs in a random effect meta-analysis that combined results of VSR and previous studies. Results—While the association between rs2200733 and CES in VSR was statistically significant (p = 0.0004, odds ratio (OR) = 2.16, 95%CI 1.41-3.32 for each risk allele), the association between rs10033464 and CES in VSR was not significant (p = 0.26, OR = 1.23, 95% CI 0.86–1.76 for each risk allele). In combined results of VSR and previous studies, the combined odds ratio per risk allele for each SNP individually was 1.53 (95%CI 1.38–1.70) for rs2200733 and 1.18 (95%CI 1.07–1.31) for rs10033464. In VSR, we considered the effect of both SNPs together and found that, compared with noncarriers (those who did not carry a risk allele of either SNP, 60% of the control group), carriers of one risk allele (of either SNP, 35% of the control group) and carriers of more than one risk allele (of either SNP, 5% of the control group) had odds ratios for CES of 1.46 (95% CI 1.05–2.03) and 2.74 (95%CI 1.57–4.76) respectively, with p = 0.0002 in a trend test. Conclusion—In the VSR, carriers of one and carriers of more than one risk alleles of the 4q25 SNP rs2200733 or rs10033464, compared with noncarriers, had increasing risk of CES.

1604T

Common genetic variants do not predict CAD in Familial Hypercholesterolemia. *E.P.A. van Iperen[1,2], S. Sivapalaratnam[3], S.M. Boekholdt[3,4], G.K. Hovingh[3], S. Maiwald[3], M.W. Tanck[2], N. Soranzo[5], J.C. Stephens[6,7], J.G. Sambrook[6,7], M. Levi[3], W.H. Ouwehand[5,6,7], J.J.P Kastelein[3], M.D. Trip[3,4], A.H. Zwinderman[2].* 1) Durrer Center for Cardiogenetic research, Amsterdam, Netherlands; 2) Department of Clinical Epidemiology, Biostatistics and Bioinformatics Academic Medical Centre, Amsterdam, the Netherlands; 3) Department of Vascular Medicine, Academic Medical Centre, Amsterdam, the Netherlands; 4) Department of Cardiology, Academic Medical Centre, Amsterdam, the Netherlands; 5) Human Genetics, Wellcome Trust Sanger Institute, Hinxton, United Kingdom; 6) Department of Haematology, University of Cambridge, Cambridge, United Kingdom; 7) National Health Service Blood and Transplant, Cambridge, United Kingdom.

Objective: In recent years, multiple loci dispersed on the genome have been shown to be associated with coronary artery disease (CAD). We investigated whether these common genetic variants also hold value for CAD prediction in a large cohort of patients with Familial Hypercholesterolemia (FH).

Methods and Results: We genotyped a total of 146 SNPs in 1701 FH patients, of whom 482 patients (28.3%) had at least one cardiovascular event during 112.943 person-years follow-up. The association of each SNP with event-free survival time was calculated with a Cox proportional hazard model. In CVD risk adjusted analysis, the lead SNP at the well-known 9p21 locus rs1333049 near CDKN2B-AS1 had a HR for CAD risk of 0.82 (95%CI 0.77–0.87; p-value 0.000945). None of the other tested CAD-associated SNPs were significantly associated with CAD risk (p > 8.9*10-5 for each).

Conclusions: Of all the loci analyzed, the 9p21 locus had the strongest negative association with CAD in this high-risk FH cohort. None of the SNPs at neither this 9p21 locus, however, nor any of the other tested CAD-associated SNPs were significantly associated with risk of CAD according to a priori defined significance threshold that took into account multiple testing.

1605T

Genetic sources of variation in proprotein convertase subtilisin/kexin type 9 (PCSK9), a protein that mediates degradation of LDL receptors and increases LDL levels. *E.A. Rosenthal[1], R. Rajagopalan[1], J. Ranchalis[1], D. Szeto[2], K. Akinsanya[2], T.P. Roddy[2], J.M. Castro-Perez[2], G. Forrest[2], G. Wolfbauer[3], J.J. Albers[3,6], J.D. Brunzell[4], A.G. Motulsky[1,5], M.J. Rieder[5], D.A. Nickerson[5], E.M. Wijsman[1,5,7], G.P. Jarvik[1,5].* 1) Dept Medical Genetics, Univ Washington School of Medicine, Seattle, WA; 2) Molecular Biomarkers Team, Merck & Co., Inc., Rahway, NJ; 3) Northwest Lipid Metabolism and Diabetes Research Laboratories, UW, Seattle, WA; 4) Div of Metabolism, Endocrinology, and Nutrition, Univ Washington School of Medicine, Seattle, WA; 5) Dept of Genome Sciences, Univ Washington School of Medicine, Seattle, WA; 6) Dept of Pathology, Univ Washington School of Medicine, Seattle, WA; 7) Dept of Biostatistics, Univ Washington School of Public Health, Seattle, WA.

High levels of LDL are associated with cardiovascular disease. Statin drugs are effective at reducing LDL by up-regulating LDL receptors (LDLR). However, this up-regulation is compromised by statin associated increases in proprotein convertase subtilisin/kexin type 9, PCSK9, a serine protease which mediates degradation of mature LDLR. Thus, statin use effects associated with increased PCSK9 protein level counters its effects on LDLR. To identify genetic regulation of PCSK9 protein levels, we searched for novel genetic variation associated with PCSK9 level, using four large families (n= 462), previously ascertained for familial combined hyperlipidemia. PCSK9 level was measured on 215 individuals using a sandwich-based antibody immunoassay. Genotypes on 330 individuals were measured on an average 5.4cM microsatellite marker map, as well as the Illumina 50K HumanCVD Beadchip. We detected linkage to chrs 1 and 20 in a single family (n=94; 42 phenotyped, 79 genotyped) using Bayesian oligogenic Markov chain Monte Carlo linkage analysis (maximum Bayes factor =47 and 88, respectively). The linkage signal at chr 1 is 3' to the structural locus, *PCSK9*, and is explained by *PCSK9* SNPs in a measured genotype analysis (including rs2495478, p<0.0001). Model based linkage analysis confirmed the linkage to chr 20 with lod=2.3, suggesting a potential genetic regulator of PCSK9 protein level in this region. To further clarify possible genes of interest under this signal, exome sequencing of 4 informative family members, based on their influence to the linkage signal at chr 20, is underway. Additionally, we found significant association of PCSK9 level with log transformed VLDL (p= 0.004) in all families, adjusting for family structure. This association with VLDL is a novel finding in humans that supports results from previous studies in mice.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1606T

Twenty-three unreported genetic associations with lipid phenotypes: a dense gene-centric meta-analysis in 66,240 individuals across 32 studies. *Y. Guo[1,2,11], F.W. Asselbergs[3,4,5,11], E.P.A. van Iperen[6,7,11], S. Sivapalaratnam[8,11], V. Tragante[3,5,11], C. Elbers[4,9], H. Hakonarson[1], B.J. Keating[1,12], F. Drenos[10,12], IBC-Lipids consortium.* 1) Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, PA, USA; 2) BGI-Shenzhen, Shenzhen, China; 3) Department of Cardiology, Division Heart and Lungs, University Medical Center Utrecht, Utrecht, The Netherlands; 4) Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, The Netherlands; 5) Department of Medical Genetics, Biomedical Genetics, University Medical Center, Utrecht, The Netherlands; 6) Durrer Center for Cardiogenetic Research, Amsterdam, The Netherlands; 7) Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands; 8) Department of Vascular Medicine, Academic Medical Center, 1105 AZ Amsterdam, The Netherlands; 9) Complex Genetics Section, Department of Medical Genetics (DBG), University Medical Center Utrecht, Utrecht, the Netherlands; 10) Centre for Cardiovascular Genetics, Institute of Cardiovascular Science, Faculty of Population Health Sciences, University College London, 5 University Street, London WC1E 6JF, UK; 11) Equal contribution; 12) Joint corresponding authors.

Genome wide association studies (GWAS) have identified many single-nucleotide polymorphisms (SNPs) underlying variations in plasma lipid levels. We set out to explore whether additional loci associated with plasma lipid phenotypes, such as high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), total cholesterol (TC) and triglycerides (TG) can be identified by a dense gene-centric approach. Our meta-analysis of32 studies in 66,240 individuals of European ancestry was based on the custom ~50,000 SNP genotyping array covering ~2,000 candidate genes (the ITMAT-Broad-CARe (IBC) array). SNP-lipid associations were replicated in a cohort comprising an additional 24,736 samples or within the Global Lipid Genetic Consortium. We identified 4, 6, 10 and 4 new SNPs in established lipid genes for HDL-C, LDL-C, TC and TG respectively. We also identified several lipid-related SNPs in previously unreported genes: *DGAT2*, *GPR109A*, *L0C338328*, *PPARG*, and *FTO* for HDL-C; *SOCS3*, *APOH*, *SPTY2D1*, *BRCA2* and *VLDLR* for LDL-C; *SOCS3*, *UGT1A1*, *BRCA2*,*UBE3B*, *FCGR2A*, *CHUK*, and *INSIG2* for TC; and *SERP1NF2*, *C4B*, *GCK*, *GATA4*, *INSR* and *LPAL2* for TG. The proportion of phenotypic variance explained in the subset of studies providing individual-level data was 9.9% for HDL-C, 9.5% for LDL-C, 10.3% for TC and 8.0% for TG. This large meta-analysis of lipid phenotypes using a dense gene-centric approach identified multiple SNPs not previously described in established lipid genes and several previously unknown loci. The explained phenotypic variance using this approach was comparable to meta-analysis of GWAS data suggesting that a focused genotyping approach can further increase the understanding of heritability of plasma lipids.

1607T

After *LDLR*, *APOB* and *PCSK9*, *APOE* is another major gene of autosomal dominant hypercholesterolemia. *M. Varret[1, 2], M. Marduel[1], K. Ouguerram[3], V. Serre[2, 4], D. Bonnefont-Rousselot[5], K.E. Berge[6], M. Devillers[4], G. Luc[7], J.-M. Lecerf[7], L. Tosolini[1], M. Abifadel[1, 8], T.P. Leren[6], J.-P. Rabès[1, 9, 10], C. Boileau[1, 9, 10].* 1) INSERM U698, Hopital Bichat, Paris, France; 2) Université Paris Denis Diderot, France; 3) INSERM U915, Institut du Thorax, Université de Nantes, France; 4) INSERM U781, Paris France; 5) 5AP-HP, CHU Pitié-Salpêtrière, Service de Biochimie Métabolique, France; 6) Medical Genetics Laboratory, Department of Medical Genetics, Oslo University Hospital Rikshospitalet, Norway; 7) Nutrition Department, Institut Pasteur de Lille, France; 8) Faculté de Pharmacie, Université St-Joseph, Beirut, Lebanon; 9) AP-HP, Hôpital A. Paré, Laboratoire de Biochimie et Génétique Moléculaire, Boulogne-Billancourt, France; 10) Université Versailles Saint-Quentin-en-Yvelines, UFR de Médecine Paris Ile-de-France Ouest, Guyancourt, France.

Autosomal Dominant Hypercholesterolemia (ADH) is characterized by high LDL levels, high risk of premature cardiovascular disease and is due to mutations in *LDLR*, *APOB* or *PCSK9* genes and the HCHOLA4 locus at 16q22.1. We found further genetic heterogeneity of ADH through a genomewide scan in a large multiplex French family with a new disease locus at 19q13.31–13.32. Whole exome sequencing for three affected family members combined with Sanger sequencing of regional and functional candidate gene were performed to identify this new ADH gene. An unique mutation in the *APOE* gene, p.Leu167del, that was not detected in over 440 control chromosomes, segregated with ADH within the French family.This finding was unexpected since apo E mutants are essentially associated with familial combined hyperlipidemia (FCHL). *APOE* p.Leu167del was not detected in over 440 French control chromosomes and two other *APOE* mutations (p.Arg235Trp, p.Arg269Gly) were found in three other unrelated ADH probands. In silico structural prediction of the mutant protein p.Leu167-del showed an alpha-helix disruption within the receptor-binding region that could affect apo E activity and its affinity to lipoproteins and/or its receptors. Two carriers of the *APOE* p.Leu167del presented with apo E levels that were normal in total serum but decreased in LDL, which could affect LDL affinity to their receptor. In vivo apo B-100 kinetics from one *APOE* p.Leu167-del carrier, showed: an increased LDL pool, which was the consequence of both an increase in LDL production rate and a decrease in LDL catabolism. These findings are in keeping with those from patients with a mutation in the *LDLR* gene. Together, these observations indicate a decreased catabolism of LDL-bearing an apo E p.Leu167del thus explaining the ADH observed for carriers of this deletion, and underscoring the hypothesis that some mutations in the *APOE* gene may be responsible of ADH and should be looked for. The *APOE* p.Leu167del mutation was associated with FCHL and, here, *bona fide* ADH. The overlapping between FCHL and ADH phenotypes was also reported for *LDLR* gene mutations. Hypertriglyceridemia can sometimes be observed in ADH subjects, mainly because of the many common genetic and environmental factors contributing to triglyceride elevation. Mutations in the *LDLR* or *APOE* gene may amplify the effect of these factors, and thus, according to the number and/or the nature of these factors, could be associated with either FCHL or ADH.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1608T

Trans-ethnic fine-mapping of lipid loci in African Americans, East Asians, and Europeans identifies population-specific signals and extensive allelic heterogeneity that increases the trait variance explained. *Y. Wu[1], L. Waite[2], A. Jackson[3], S. Buyske[4], C. Carty[5], I. Cheng[6], D. Duggan[7], L. Dumitrescu[8], C. Haiman[9], L. Hindorff[10], C. Hsiung[11], S. Hunt[12], K. Hveem[13], J. Juang[14], E. Kim[15], M. Laakso[16], I. Njølstad[17], U. Peters[5], R. Rauramaa[18,19], W. Sheu[20,21,22], Y. Sung[23], J. Tuomilehto[24,25,26,27], T. Wang[28], Y. Chen[15], C. Kooperberg[5], M. Boehnke[3], D. Crawford[8], T. Assimes[29], K. North[30], K. Mohlke[1].* 1) Department of Genetics, University of North Carolina, Chapel Hill, NC, USA; 2) HudsonAlpha Institute for Biotechnology, Huntsville, AL, USA; 3) Department of Biostatistics and Center for Statistical Genetics, University of Michigan, Ann Arbor, MI, USA; 4) Department of Statistics and Biostatistics, Rutgers University, Piscataway, NJ, USA; 5) Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, USA; 6) University of Hawaii Cancer Center, Honolulu, HI, USA; 7) Translational Genomics Research Institute, Phoenix, AZ, USA; 8) Center for Human Genetics Research, Vanderbilt University, Nashville, TN, USA; 9) Keck School of Medicine, University of Southern California, Los Angeles, CA, USA; 10) Office of Population Genomics, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD, USA; 11) Division of Biostatistics and Bioinformatics, Institute of Population Health Sciences, National Health Research Institutes, Zhunan, Taiwan; 12) Department of Internal Medicine, University of Utah, Salt Lake City, UT, USA; 13) HUNT Research Centre, Department of Public Health and General Practice, Norwegian University of Science and Technology, Levanger, Norway; 14) National Taiwan University College of Medicine and National Taiwan University Hospital, Taipei, Taiwan; 15) Department of Medical Genetics, Cedars-Sinai Medical Center, Los Angeles, CA, USA; 16) Department of Medicine, University of Eastern Finland and Kuopio University Hospital, 70210 Kuopio, Finland; 17) Department of Community Medicine and Department of Clinical Medicine, Faculty of Health Sciences, University of Tromsø, Tromsø, Norway; 18) Kuopio Research Institute of Exercise Medicine, Kuopio, Finland; 19) Department of Clinical Physiology and Nuclear Medicine, Kuopio University Hospital, Kuopio, Finland; 20) Division of Endocrine and Metabolism, Department of Internal Medicine, Taichung Veterans General Hospital, Taichung, Taiwan; 21) School of Medicine, National Yang-Ming University, Taipei, Taiwan; 22) College of Medicine, National Defense Medical Center, Taipei, Taiwan; 23) Division of Biostatistics, Washington University School of Medicine, St. Louis, MO, USA; 24) Diabetes Prevention Unit, National Institute for Health and Welfare, 00271 Helsinki, Finland; 25) South Ostrobothnia Central Hospital, 60220 Seinäjoki, Finland; 26) Red RECAVA Grupo RD06/0014/0015, Hospital Universitario La Paz, 28046 Madrid, Spain; 27) Centre for Vascular Prevention, Danube-University Krems, 3500 Krems, Austria; 28) Cardiovascular Center and Division of Cardiology, Department of Internal Medicine, National Taiwan University Hospital,Taipei City, Taiwan; 29) Department of Medicine, Stanford University School of Medicine, Stanford, CA, USA; 30) Department of Epidemiology and Carolina Center for Genome Sciences, University of North Carolina, Chapel Hill, NC, USA.

Genome-wide association studies (GWAS) have revealed ~100 loci associated with lipids, but much of the trait heritability remains unexplained. We conducted a trans-ethnic study in individuals of African American (n=6,832), East Asian (n=9,452), and European (n=10,829) ancestry to analyze high-density genotyped Metabochip SNPs at 18, 22 and 18 GWAS loci for their association with triglycerides, high-density lipoprotein cholesterol (HDL-C), and low-density lipoprotein cholesterol (LDL-C). We aimed to identify the variants with strongest association at each locus, identify allelic heterogeneity and population-specific signals, determine whether trans-ethnic high-density genotyping would refine association signals, and assess whether previously described functional variants exhibit the strongest association at those signals. Among the 58 loci, 33 exhibited evidence of association ($P<10^{-4}$) in at least one population. Stepwise conditional analyses revealed a total of 34, 9 and 21 signals at 10, 4 and 9 loci in African Americans, East Asians and Europeans, respectively. All these signals led to a 1.3- to 1.8-fold increase in the explained phenotypic variance compared to the strongest signals alone. The seven signals at *PCSK9* in African Americans include six nonsense or missense variants shown previously to have functional effects, and five of them are African American-specific. Among the seven signals at the *TOMM40-APOE-APOC4* cluster in African Americans, the trait-lowering allele (freq=0.18) of the fourth signal (rs157588) was present on haplotypes containing the trait-raising allele (freq=0.35) of the third signal (rs1038026); after accounting for linkage disequilibrium ($r^2=0.35$) with the third signal, the fourth signal increased in significance ($P_{initial}=0.72$; $P_{condition}=2.0\times10^{-7}$). Trans-ethnic analyses localized one signal at *GCKR* to a reported functional variant P446L, and narrowed other signals at *PPP1R3B* from >10 SNPs spanning 50 kb to two SNPs (rs6601299 and rs4841132; LD $r^2>0.94$) located 1 kb apart and at *LCAT* from >100 SNPs spanning 800 kb to four SNPs within a 50 kb region. Of the 27 SNPs reported previously to affect gene expression or protein function, 74–84% exhibited the strongest association at the respective signal. In summary, trans-ethnic high-density genotyping allows the identification of allelic heterogeneity and population-specific variants, and reveals candidate SNPs for functional studies.

1609T

Identification of a Major Genetic Modifier in Mouse Models of TGFβ Vasculopathies. *J. Calderon[1], H. Dietz[1,2].* 1) McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University, Baltimore, MD; 2) Howard Hughes Medical Institute.

Multiple presentations of aortic aneurysm associate with a tissue signature for high TGFβ signaling and/or phenotypic rescue in response to TGFβ antagonists. These include Marfan syndrome (MFS, mutations in FBN1), Loeys-Dietz syndrome (LDS, mutations in TGFBR1/2), and LDS-like disorders (mutations in SMAD3 or TGFB2). The fact that some mutations would intuitively lead to low TGFβ signaling has engendered controversy in the field. Many prior studies document that the C57Bl/6J (BL6) background exacerbates TGFβ deficiency phenotypes when compared to 129SvE (SV129), but the source of modification is unknown. SV129 mice heterozygous for a LDS-associated mutation in Tgfbr2 (G357W/+) show late postnatal death due to aortic dissection. When G357W was crossed onto the BL6 background, there is near-complete penetrance of perinatal death due to persistent truncus arteriosus (PTA), a form of congenital heart disease that has been extensively associated with loss of TGFβ signaling in the cardiac neural crest. A single cross of isogenic SV129 LDS mice with pure BL6 mice results in 70% incidence of PTA among G357W/+ offspring, suggesting a dominant BL6 deleterious modifier with dosage-dependent PTA. Pure SV129 G357W/+ mice were bred to F2 mice with extensive recombination between BL6 and SV129 chromosomes and late gestation fetuses were phenotyped for PTA by intravascular latex injection. A GWA analysis was performed using the Illumina Mouse Medium Density Linkage Panel to identify haplotype enrichment in PTA+ mutation carriers. Remarkably, this analysis yielded a single BL6 locus on distal chromosome 9 (LOD=9) that was responsible for deleterious modification. Notably, the rare BL6 LDS mice that escape perinatal mortality due to PTA show protection from postnatal aneurysm progression and dissection, when compared to their SV129 counterparts. These data suggest that LDS mutations have the potential to impose TGFβ deficiency phenotypes during development, but that postnatal aneurysm progression manifests a switch to a high signaling state. Parallel work in the lab has shown that BL6 mice harboring a Fbn1 mutation causing MFS are remarkably protected against postnatal aneurysm progression and dissection ($p<0.0001$), when compared to SV129; the role of the chromosome 9 locus is under investigation. Taken together, these data suggest that genetic variation that attenuates TGFβ signaling is sufficient to blunt aneurysm severity in multiple disorders.

1610T

Upregulation of Twist1 expression during acute myocardial infarction. *J.B. Nevado.* Institute of Human Genetics, University of the Philippines Manila, Manila, Metro Manila, Philippines.

Myocardial infarction (MI) is a major cause of mortality and morbidity worldwide. During acute myocardial infarction, a series of events leads to fibrosis, which, if excessive, may hinder effective remodeling. Twist1 is a transcriptional repressor that has been well implicated in tissue fibrosis, including cardiac fibrosis. Speculatively, the fibrosis occurs through collagen-producing epithelial-derived fibroblasts via epithelial to mesenchymal transition (EMT). Due to fibrotic sequelae after an acute myocardial infarction, I hypothesize that Twist1 is likely to be expressed in the injured tissues a few days after an event, possibly triggering EMT. We used a mouse model for myocardial infarction in which the left anterior descending artery was ligated. Mice were sacrificed at selected timepoints (0-, 3-, 7- and 10-day). Immunohistochemistry of myocardial samples showed increased expression of Twist1 in the periinfarct region at day 3 with tapering at day 7, and eventual loss of expression at day 10. Comparatively, uninjured heart samples, Day 0 samples and remotely located non-injured tissues did not show significantly expression of Twist1. These results suggest that Twist1 is acutely expressed after myocardial infarction, and can be considered as a candidate molecular trigger of post-MI cardiac fibrosis.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1611T

Venous malformation-causing TIE2-mutations lead to AKT-mediated downregulation of PDGFB. *M. Uebelhoer[1], M. Nätynki[2], J. Kangas[2], J. Soblet[1], A. Mendola[1], C. Godfraind[3], L.M. Boon[1, 4], L. Eklund[2], N. Limaye[1], M. Vikkula[1, 5].* 1) Human Molecular Genetics, de Duve Institute, Université catholique de Louvain, Brussels, Belgium; 2) Oulu Center for Cell-Matrix Research, Biocenter Oulu and Department of Medical Biochemistry and Molecular Biology, University of Oulu, Oulu, Finland; 3) Neuroscience Center, Laboratory of Neuropathology, Université catholique de Louvain, Brussels, Belgium; 4) Center for Vascular Anomalies, Division of Plastic Surgery, Cliniques universitaires St-Luc, Université catholique de Louvain, Brussels, Belgium; 5) Center of Human Genetics, Cliniques universitaires St-Luc, Université catholique de Louvain, Brussels, Belgium.

Mutations in the endothelial-cell tyrosine kinase receptor TIE2 cause inherited and sporadic forms of venous malformation. The most recurrent somatic mutation, L914F, and the common germline mutation, R849W, have been shown to differ in terms of phosphorylation-level, as well as sub-cellular localization and trafficking of the receptor. We now show that PI3K/AKT and STAT1 are chronically activated by both mutant forms, with L914F exerting a stronger effect. Gene expression profiling of HUVECs overexpressing the mutant or wild-type forms of TIE2 indicates that L914F strongly dysregulates genes involved in vascular development and cell migration, while R849W has weak effects, making it indistinguishable from wild-type cells in global analyses. Interestingly, we show for the first time that VM-causative mutations inhibit the transcription factor FOXO1 in an AKT-dependent manner, resulting in deficient PDGFB-production. Lack of this mural cell attractant might explain the paucity of smooth muscle cells surrounding distended vascular channels in venous malformations.

1612T

Islet1 is a direct transcriptional target of the homeodomain transcription factor Shox2 in the sinoatrial node of the developing heart. *G.A. Rappold[1], I.M. Berger[2], A. Glaser[1], C. Bacon[1], L. Li[3], K.U. Schneider[1], N. Gretz[3], H. Steinbeisser[1], W. Rottbauer[2], S. Just[2], S. Hoffmann[1].* 1) Institute of Human Genetics, IUniversity of Heidelberg, NF 366, 69120 Heidelberg, Germany; 2) Department of Intermal Medicine II, University of Ulm, Albert-Einstein-Allee 23, 89081 Ulm, Germany; 3) Medical Research Center, University of Heidelberg, Theodor-Kutzer-Ufer 1-3, 68167 Mannheim, Germany.

The heart's rhythm is initiated and regulated by a group of specialized cells in the sinoatrial node (SAN), the primary pacemaker of the heart. Abnormalities in the development of the SAN can result in irregular heart rates (arrhythmias). Although several of the critical genes important for SAN formation have been identified, our understanding of the transcriptional network controlling SAN development remains at a relatively early stage. The LIM homeodomain transcription factor Islet1 (Isl1) represents a prominent marker for cardiac progenitor cells of the second heart field and has been proposed, very similar to Shox2, to play an essential early role in the specification and patterning of the SAN. Here, we compared gene expression levels in the right atria of wildtype and *Shox2*$^{-/-}$ hearts using microarray experiments and identified *Isl1* as one of its putative target genes. The downregulation of *Isl1* expression in *Shox2*$^{-/-}$ hearts was confirmed and the affected region narrowed down to the SAN by whole mount *in situ* hybridization. Using luciferase reporter assays and EMSA studies, we identified two specific SHOX2 binding sites within intron 2 of the *ISL1* locus. We also provide functional evidence for *Isl1* as a transcriptional target of *Shox2* by rescuing the Shox2-mediated bradycardia phenotype with *Isl1* using zebrafish as a model system. Our findings demonstrate a novel epistatic relationship between *Shox2* and *Isl1* in the heart with important developmental consequences for SAN formation and heart beat.

1613T

Analysis of sub-category for 141 cases of congenital heat disease (CHD) that have been surgically corrected during neonatal period. *Q-Y. Cao[1], J-Z. Zhu[2], N. Zhong[1,3,4].* 1) Shijiazhuan Maternal Hospital, Shijiazhuan, Hebei, China; 2) Hebei Provincial Hospital, Shijiazhuan, Hebei, China; 3) Peking University Center of Medial Genetics, Beijing, China; 4) New York State Institute for Basic Research in Developmental Disabilities, Staten Island, NY.

Congenital heart disease (CHD) is a leading condition in China as the most common visible birth defect national wide. It is also the one that could be surgically corrected right after the baby is born. Currently most of CHD is recognizable by the prenatal ultrasound and the family is prepared to have the baby gone through the surgical correction. Here we have analyzed the sub-category on 141 cases of CHD that have been surgically corrected in our local hospitals. Surgical correction was performed through a National Intervention Program for Saving the Babies with CHD. Surgical records were accessed through our provincial birth defect surveillance system or through hospital documents. The most common CHD among our patients is ventricular septal defect (VSD), which accounts for 55%. The order of sub-category of VSD is VSD without other anomaly (19.9%), VSD with patent foramen ovale (7.8%) and VSD with pulmonary hypertension (7.8%), VSD with atrial septal defect (ASD, 5.2%), and other VSD with variant complicated anomalies. Tetralogy of fallot listed as the second most common CHD (17.7%). The tetralogy of fallot usually is more complicated than expected. The atrial septal defect (ASD, 11.4%) accounts for the third common CHD among the 141 cases, within which the ASD with pulmonary hypertension (4.8%) is more than ASD with pulmonary valve stenosis (1.4%) and ASD alone (0.7%). The less common CHD are patent ductus arteriosus (PDA, 9.2%), pulmonary valve stenosis (PVS, 1.4%) and complete endocardial cushion defect partial endocardial cushion defect (0.7%), aortic regurgitation with stenosis (0.7%), main artery of ectopic (0.7%) and complicated heart anomilies (0.7%). Our data has provided information about the frequency of sub-category of CHD among Chinese patients that may help clinical genetic cousuling.

1614T

Analysis of cardiomyopathy using whole genome sequencing. *E.M. McNally[1], J.R. Golbus[1], L.L. Pesce[2], D. Wolfgeher[1], L. Dellefave-Castillo[1], M.J. Puckelwartz[1].* 1) Med, Sect Cardiology, Rm G611, Univ Chicago, Chicago, IL; 2) Computation Institute, Univ Chicago, Chicago.

Background: Congestive heart failure (CHF) is a disabling and costly medical problem that currently affects more than 5 million Americans. One of the leading causes of CHF is cardiomyopathy, a disorder with a high heritable component. There are few clinical clues to guide genetic diagnosis, and therefore genetic testing for familial cardiomyopathy may survey as many as 46 responsible genes. Recently the large TTN gene, which comprises 350 exons, was identified as being the cause of nearly 25% of dilated cardiomyopathy. Even with the inclusion of the TTN gene in gene panels for cardiomyopathy, the sensitivity of testing for dilated cardiomyopathy (DCM) is less than 50%. Because of the large number of genes required to diagnosis genetic cardiomyopathy and the improvement in high throughput sequencing approaches, we piloted whole genome sequencing for 11 subjects with dilated cardiomyopathy. Methods: WGS was performed to achieve an average of 35 fold coverage using paired end 100 bp reads. We focused our analysis on 212 genes linked to both syndromic and nonsyndromic cardiomyopathy as well as genes identified from animal modeling. Variants within these genes were analyzed using a pipeline that combined multiple prediction algorithms along with frequency data from the 1000Genomes project and the NHLBI Exome Sequencing Project. This pipeline yielded 3–15 potentially pathogenic variants per individual. Variants were confirmed in the proband using traditional Sanger sequencing and then sequenced in family members, where available, for confirmation by segregation. Results: Three subjects of 11 in dilated cardiomyopathy cohort had known primary mutations, and these three mutations were detected by the analysis pipeline. In five of eight subjects for whom the primary mutation was unknown, we identified mutations that either segregated with disease, had clinical correlates and/or additional pathological correlation to provide evidence as to causality. For at least two subjects, we identified additional variants that may be acting as modifiers to influence the severity of disease. In total, we identified the likely pathological mutation in 8 of 11 (73%). Conclusion: Cardiomyopathy is a genetically heterogeneous disease that lends itself to analysis by WGS. WGS can be used to identify clinically relevant variants of both therapeutic and prognostic importance.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1615T

THE ASSOCIATION BETWEEN NOS3 GENE VARIATIONS AND BRUGADA SYNDROME. *S. Mehrtashfar[1], A. Ebrahimi[2], M. Moghadam[2], A.S. Khatir[3].* 1) Biology Dept, Guilan University, Guilan, Iran; 2) Faculty of Medicine, Shiraz University of Medical Sciences, Shiraz, Iran; 3) Sasan Hospital, Tehran, Iran.

The brugada syndrome is the main cause of sudden unexplained death syndrome (SUDS) and suffering people have an abnormal electrocardiogram (ECG) with increased risk of sudden cardiac death without cardiac background. One of the most important genes involved in cardiac disease is nitric oxide synthase 3 (NOS3) gene which its product synthesize nitric oxide from L-Arginine. Nitric oxide plays a role as a mediator in several processes such as neurotransmission that failure in this process can be a reason of heart arrhythmia. We studied two variations with more clinical significance in cardiac disease (T786C and G894T) in 25 patients with brugada syndrome and 60 people with normal ECG as control group. These variations were studied by RFLP (T786C) and ARMS PCR (G894T) methods on extracted DNA from peripheral blood cells. Results showed T allele in G894T variation in patient group was significantly higher than control group ($p<0.5$), but there was no difference in T786C variation between patients and normal people. This study introduced G894T variations as a potential marker of brugada syndrome.

1616T

Targeted sequence capture and 454 sequencing of 23 cardiomyopathy genes; implementation into diagnostics. *O. Mook[1], M. Haagmans[1], R. Lekanne Deprez[1], F. Baas[2], M. Jakobs[2], N. Hofman[1], I. Christiaans[1], M. Mannens[1].* 1) Clinical Genetics, Academic Medical Center, Amsterdam, Amsterdam, Netherlands; 2) Dept of Genome Analysis, AMC, Amsterdam, the Netherlands.

Background—Genetic evaluation of cardiomyopathies poses a challenge. Multiple genes are involved but no clear genotype-phenotype correlation has been found so far. In the past, genetic evaluation for hypertrophic (HCM) and dilated (DCM) cardiomyopathies was performed by sequential dHPLC screening followed by Sanger sequencing of a very limited number of genes. Recent developments in sequencing has increased the throughput enabling simultaneous screening of multiple genes for multiple patients in a single sequencing run. Methods and Results—In order to increase the number of genes that can be screened in a shorter time period, we enriched all exons of 23 cardiomyopathy related genes using on-array multiplexed sequence capture followed by massively parallel pyrosequencing. We show that after optimization of array based sequence capture it is feasible to reliably detect a large panel of known and unknown variants in HCM and DCM patients, whereby the unknown variants could be confirmed by Sanger sequencing. Conclusions—The rate of variant detection in both HCM and DCM patients was increased due to a larger number of genes studied. Array based target enrichment followed by Next Generation Sequencing (NGS) is ready for implementation in diagnostics.

1617T

Targeted next generation sequencing of thoracic aortic aneurysmal genes. *L. Van Laer[1], D. Proost[1], G. Vandeweyer[1], J. Saenen[2], B. Paelinck[2], C. Vrints[2], B. Loeys[1].* 1) Center of Medical Genetics, Faculty of Medicine and Health Sciences, University of Antwerp and Antwerp University Hospital, Antwerp, Belgium; 2) Department of Cardiology, Faculty of Medicine and Health Sciences, University of Antwerp and Antwerp University Hospital, Antwerp, Belgium.

Aortic aneurysm/dissection represents an important cause of death in the industrialized world. The genetic contribution to thoracic aortic aneurysm (TAA) is significant; 20% of all affected individuals have a positive family history. The latter figure is likely to be underestimated, because many family members may not be aware of the presence of an aneurysm. TAA is genetically heterogeneous; more than ten causative genes have been identified so far. As a considerable clinical overlap exists between different types of TAA caused by mutations in different genes on the one hand, and as mutations in the same gene cause a wide phenotypic variability on the other hand, a clinical diagnosis is not always straightforward and the molecular confirmation of the clinical diagnosis is important for further patient management and therapy. However, the consecutive molecular screening of all TAA genes using conventional methods is expensive and labor-intensive. To shorten the turn-around-time, to increase the informativeness (i.e. identification of modifying variants) and to reduce the overall cost of molecular testing, we are developing a TAA gene panel for next generation sequencing of twelve TAA genes (*ACTA2*, *COL3A1*, *FBLN4*, *FBN1*, *FLNA*, *MYH11*, *MYLK*, *NOTCH1*, *SMAD3*, *TGFBR1*, *TGFBR2* and *SLC2A10*). We have opted for an innovative and simple enrichment method that does not require expensive, dedicated machinery: the Haloplex technology, which makes use of dual hybridization recognition. The complexity and the required hands-on time of sample preparation are significantly reduced through the addition of the sample identification barcodes and sequencing primers during the enrichment procedure. We have designed probes for a total of 320 exons (including intron/exon boundaries) or a total target size of approximately 62 kb using the Halo Design Wizard complemented with manual adjustments. We have obtained an overall coverage by design of 99.7%. We have performed three separate 2x 150 bp paired-end runs on a Miseq sequencer, analyzing both samples with known and with unknown genetic defects. We have achieved an overall sequencing coverage of 0.2x the average read depth for 95% of the regions. We are currently optimizing the filter settings for the analysis of the variant lists in order to reach a maximum sensitivity and specificity.

1618T

The use of next generation sequencing in clinical diagnostics of familial hypercholesterolaemia. *J. Vandrovcova[1], E. Thomas[1], S. Atanur[1], P. Norsworthy[1], C. Neuwirth[2], Y. Tan[2], L. Game[2], A. Soutar[2], T. Aitman[1], Generation Scotland.* 1) Physiological Genomics and Medicine Group, MRC Clinical Sciences Centre, Imperial College London, UK; 2) MRC Clinical Sciences Centre, Imperial College London, UK.

Familial hypercholesterolaemia (FH) is a common autosomal dominant condition that is associated with increased risk of early coronary heart disease. Treatment with statins is highly effective and the identification of affected individuals is therefore crucial. Currently only about 20% of FH patients are identified in the general population in England partly because of insufficient availability of molecular testing. The purpose of this study is to determine the sensitivity and specificity of amplicon library generation and massively parallel sequencing for the identification of causal mutations, and to investigate the potential for wider population molecular testing. In pilot studies, 23 DNA samples from patients with known mutations and 24 negative samples were screened using the Fluidigm Access Array. Amplicons were designed to cover the whole LDLR gene and mutation hot spots in PCSK9 and APOB genes. After amplification and barcoding reactions, samples were pooled and sequenced using the MiSeq platform. Data were processed using common bioinformatics tools including GATK and compared to variant calls produced by MiSeq Reporter. In addition, 100 patients referred to a local lipid clinic and 400 individuals from a population cohort were included in the study. Large insertions and deletions were screened using MLPA. All amplicons were successfully amplified and > 99% of reads mapped to target regions. In the confirmatory part of the study all 20 known point mutations (17 in LDLR, 1 pY374 in PCSK9 and 2 pQ3527 in APOB) as well as 3 short indels in LDLR were correctly identified with both GATK and MiSeq Reporter. No false positives were reported when using GATK. Several new diagnoses were made in patients included in the prospective part of the study with causative variants being detected in 20–40% of patients depending on clinical phenotype. Population cohort results from 400 patients are in progress and will be presented. The established protocol is fast, with a current turnaround including data analysis of less than a week for 48 patients, and highly accurate. We believe that in combination with MLPA it represents an alternative cost-effective high-throughput method for sequencing known FH genes that could open up molecular testing and subsequent treatment to broader populations. Moreover our results suggest that similar assays targeting other common diseases with Mendelian subtypes may be warranted with this approach.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1619T

Targeted next generation sequencing as a diagnostic test in patients with cardiomyopathies. *Y.J. Vos[1], B. Sikkema-Raddatz[1], L.F. Johansson[1], E.N. de Boer[1], L.G. Boven[1], K.Y. van Spaendonck-Zwarts[1], M.P. van den Berg[2], J.P. van Tintelen[1], J.D.H. Jongbloed[1], R.J. Sinke[1].* 1) Department of Genetics, University Medical Centre Groningen, Groningen, the Netherlands; 2) Department of Cardiology, University Medical Centre Groningen, Groningen, the Netherlands.

Introduction Next generation sequencing techniques offer novel possibilities to analyse large numbers of genes in parallel. Exome sequencing (ES) can be used to sequence the coding regions of the genome. To date, ES is not developed to a level at which it provides a reliable substitute for Sanger sequencing. Both incomplete representation and coverage of some exons would certainly result in missing clinically relevant mutations. Targeted enrichment circumvents this shortcoming by selecting only those genes involved in a particular disease. We developed a kit based on Agilent Sure Select Target Enrichment for mutation detection in 48 genes that were previously proven to be involved in hereditary cardiomyopathies.

Methods For validation purposes, two series of twelve patients were analysed applying this kit and Sanger sequencing for up to six genes. Moreover, several additional series of twelve patients were analysed. Sample preparations were performed according to the manufacturer's protocol. All samples were sequenced using 151 bp paired-end reads on an Illumina MiSeq sequencer and analysed using the MiSeq Reporter pipeline and/or Next gene software.

Results The average coverage in subsequent runs varied between 170 to 250 reads and only 1.5% of the targets showed a coverage <30 reads in more than one of twelve patients within a single run. All pathogenic mutations, Unclassified Variants (UV's) and single nucleotide polymorphisms previously detected by Sanger sequencing were identified. Furthermore all novel variants (not listed in db-SNP or 1000 genome databases) that were analysed so far could be confirmed by Sanger sequencing. Our analyses resulted on average in the identification of three mutations/UV's per patient. In a significant number of patients that previously were analysed without a disease causing mutation, a (potential) pathogenic mutation was identified. *In silico* analyses, confirmation experiments by Sanger sequencing and co-segregation analyses are performed to identify the causal mutation in each patient/family.

Conclusion Our data convincingly demonstrates that targeted resequencing of a disease specific subset of genes can reliably substitute Sanger sequencing as a diagnostic test.

1620T

Population Sampling and *in vitro* Modeling of a 25bp Deletion in MYBPC3 Associated With Hypertrophic Cardiomyopathy. *A.B. Chowdry[1], M.A. Mandegar[2], G.M. Benton[1], B.T. Naughton[1], B.R. Conklin[2].* 1) 23andMe Inc., Mountain View, CA; 2) The Gladstone Institute of Cardiovascular Disease, San Francisco, CA.

Hypertrophic cardiomyopathy (HCM) is the leading cause of sudden cardiac death (SCD) in young athletes. The prevalence of HCM has been estimated to be as high as 1 in 500 within the general population. However, due to cost and sensitivity problems, it is not screened for in the United States.

The majority of mutations leading to HCM have been identified in sarcomeric proteins, mainly the cardiac isoform of myosin-binding protein C (MYBPC3). Advances in DNA analysis could allow low-cost screening for high penetrance mutations in MYBPC3 and could inform selective echocardiograms. Of particular interest is a recently discovered 25bp deletion (rs36212066), known to have a high carrier frequency in South Asians, but previously unobserved in European populations.

23andMe has a database of over 150,000 genotyped individuals. This database contains over 160 carriers of the rs36212066 deletion, some of European ancestry, suggesting that this mutation has a broader distribution than initially reported. We are also able to observe homozygotes in the database, though at a reduced frequency consistent with previous studies. The large size of the 23andMe database uniquely positioned it to detect these individuals. We have performed an initial survey of these individuals to determine self-reported HCM status, but further follow up may be necessary.

While in vitro models of HCM have been created in human induced pluripotent stem (iPS) cells, none have been made so far for MYBPC3. We are using Transcription Activator-Like Effector Nuclease (TALEN)-directed homologous recombination to make an allelic series of MYBPC3 mutations in a common genetic background, including a heterozygous knockout. We are also engineering a FLAG epitope tag to reveal key binding partners of the of the ancestral and mutant forms of MYBPC3 via immunoprecipitation from iPS-derived cardiac myocytes.

Our work on the in vitro cellular phenotypes of iPS cells carrying this mutation can eventually be extended to population phenotyping using the 23andMe database. These studies will help explain the pathology and incomplete penetrance of rs36212066 and may provide targets for pharmaceutical therapy.

1621T

Co-localization of lipid biomarker associations with gene expression across human tissues. *C. Giambartolomei[1], H-J. Westra[2], M. Kivimaki[3], M. Kumari[3], E. Schadt[4], L. Franke[2], A. Hingorani[3], V. Plagnol[1].* 1) University College London (UCL) Genetics Institute, London, United Kingdom; 2) University Medical Centre, Groningen, Netherlands; 3) University College London (UCL) Department of Epidemiology and Public Health, London, United Kingdom; 4) Mount Sinai School of Medicine, New York, USA.

The genome-wide association study design has provided many new insights into the genetic basis of cardiovascular disease and associated biomarkers, in particular lipid traits. As the majority of these genetic findings map to regulatory rather than coding regions, the next step to understand the molecular basis of these associations is to assess the genes that are regulated by associated genetic variants. To achieve this goal, we combined two expression datasets (in liver, n = 966 samples and blood, n = 1,550 samples) with the Whitehall II cohort, a large scale prospective study with deep phenotyping for multiple lipid and disease traits. We also developed a novel statistical methodology to assess the concordance between disease and expression traits. We find that a large fraction of disease/biomarker associations are consistent with causal variant shared with an expression quantitative trait locus. Our findings provide information about the causal gene in associated intervals and have direct implications for the understanding of cardiovascular traits as well as the design of drugs to target disease pathways.

1622T

Silencing of Atp2b1 increases blood pressure through vasoconstriction. *B. Oh[1], Y-B. Shin[1], S-M. Ji[1], S-Y. Park[1], H-J. Lee[1], K-W. Hong[1], J-E. lim[1], M. Lim[2], Y-H. Lee[2].* 1) Biomedical Engineering, Sch Med, Kyung Hee Univ, Seoul, South Korea; 2) Department of Physiology, College of Medicine, Yonsei University, Seoul, South Korea.

Recent GWASs have identified 20 genetic loci that regulate blood pressure, increasing our understanding of the etiology of hypertension. However, it has been difficult to define the causative genes at these loci due to a lack of functional analyses. In this study, we aimed to validate the candidate gene ATP2B1 in 12q21, variants near which have the strongest association with blood pressure in Asians and Europeans. ATP2B1 functions as a calcium pump to fine-tune calcium concentrations—necessary for repolarization following muscular contractions. We silenced Atp2b1 using an siRNA complex, injected into mouse tail veins. In treated mice, blood pressure rose and the mesenteric arteries increased in wall:lumen ratio. Moreover, the arteries showed enhanced myogenic responses to pressure, and contractile responses to phenylephrine increased compared with the control, suggesting that blood pressure is regulated by ATP2B1 through the contraction and dilation of the vessel, likely by controlling calcium concentrations in the resting state. These results support that ATP2B1 is the causative gene in the blood pressure-associated 12q21 locus and demonstrate that ATP2B1 expression in the vessel influences blood pressure.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1623T

Metabolic imprints of genetically elevated body mass index in young adults. *P. Würtz[1,2], A.J. Kangas[2], P. Soininen[2], O. Raitakari[3], T. Lehtimäki[4], MR. Järvelin[5], M. Ala-Korpela[2,5], S. Ripatti[1].* 1) Institute for Molecular Medicine, University of Helsinki, Helsinki, UH, Finland; 2) Computational Medicine, University of Oulu, Finland; 3) Cardiovascular Research Center, University of Turku, Finland; 4) Clinical Chemistry, University of Tampere, Finland; 5) Epidemiology and Biostatistics, Imperial College London, United Kingdom.

Genetically elevated body mass index has been linked with the risk for future cardiovascular disease, yet it remains incompletely understood how this association is mediated. To address the impact of obesity-related loci on the systemic metabolic profile we tested associations of a genetic risk score composed of 32 previously identified variants with established cardiovascular risk factors and the serum metabolome profile in 6,774 young adults (mean age 31 ± 3 years) from two population-based cohorts in Finland. The genetic risk score was associated with increased body mass index ($P = 1\times10^{-30}$) in these young adults and had residual effects on blood pressure, insulin resistance index and low-grade inflammation as well as increased levels of liver enzymes in the circulation ($P < 0.0005$). Genetically elevated BMI was also associated with an adverse lipoprotein subclass profile characterized by elevated VLDL and decreased large HDL lipid levels while effects on LDL levels were only significant in women. The strongest metabolite abnormalities caused by the obesity-related genetic variants were found for circulating concentrations of branched-chain and aromatic amino acids. These amino acids have recently been associated with the risk for development of type 2 diabetes and our findings suggest that obesity is directly contributing to mediate this associations already in early adulthood. In conclusion, our results in apparently healthy young adults suggest that increased body mass index due to common genetic variants has a diverse impact on the global metabolite profile and that the association of body mass index with the risk for future heart disease is at least partly transmitted through established cardiovascular risk factors. These findings assist to elucidate the role of novel genetic loci for body mass index as well as emphasize risk factor monitoring in individuals genetically susceptible to obesity.

1624T

Mendelian randomisation studies do not support a causal effect of plasma lipids on insulin sensitivity. *T. Fall[1], W. Xie[2], K. Hao[3], J. Ärnlöv[4,5], F. Abbasi[6], E.E. Schadt[3], G. Boran[7], T. Hansen[8,9], D. Greenawalt[3], J.J. Nolan[10], O. Pedersen[8,11,12,13], H. Häring[14], E. Ferrannini[15], A. Syvänen[16], T. Quertermous[6], U. Smith[17], T.L. Assimes[6], M. Laakso[18], M. Walker[19], J.W. Knowles[6], M.N. Weedon[2], E. Ingelsson[1], T.M. Frayling[2] on behalf of the GENESIS consortium.* 1) Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden; 2) Peninsula College of Medicine and Dentistry, Exeter, Devon, United Kingdom; 3) Merck Research Laboratories, Boston, MA, USA; 4) Department of Public Health and Caring Sciences, Geriatrics, Uppsala University, Uppsala, Sweden; 5) School of Health and Social Studies, Dalarna University, Falun, Sweden; 6) Department of Medicine, Stanford University School of Medicine, Stanford, California, USA; 7) Consultant Chemical Pathologist Department of Clinical Chemistry Tallaght Hospital, Tallaght, Ireland; 8) Novo Nordisk Foundation Center for Basic Metabolic Research Faculty of Health Science, University of Copenhagen, Copenhagen, Denmark; 9) Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark; 10) Steno Diabetes Center, Gentofte, Denmark; 11) Hagedorn Research Institute, Copenhagen, Denmark; 12) Institute of Biomedical Science, Faculty of Health Sciences, University of Copenhagen, Copenhagen, Denmark; 13) Faculty of Health Sciences, University of Aarhus, Aarhus, Denmark; 14) Department of Internal Medicine, Division of Endocrinology, Diabetology, Nephrology, Vascular Medicine and Clinical Chemistry, University of Tübingen, Tübingen, Germany; 15) Department of Internal Medicine, University of Pisa, Pisa, Italy; 16) Department of Medical Sciences, Molecular Medicine and Science for Life, Laboratory, Uppsala University, Uppsala, Sweden; 17) The Lundberg Laboratory for Diabetes Research, Department of Molecular and Clinical Medicine, Sahlgrenska Academy, Gothenburg, Sweden; 18) Department of Medicine, University of Eastern Finland and Kuopio University Hospital, Kuopio, Finland; 19) Institute of Cellular Medicine, Newcastle University, Newcastle upon Tyne, UK.

Background: Elevated plasma concentrations of triglycerides (TG) coupled with decreased concentrations of high-density lipoprotein cholesterol (HDL-C), are associated with decreased insulin sensitivity and type 2 diabetes. However, previous studies have provided conflicting evidence of whether these associations are causal or not. ***Objective***: We aimed to test the hypothesis that genetically-determined changes of TG, HDL-C, and low-density lipoprotein cholesterol (LDL-C) concentrations influence insulin sensitivity using a Mendelian randomisation approach. ***Research Design and Methods***: We meta-analysed four studies comprising 2,590 European non-diabetic individuals with available TG, HDL-C, LDL-C and gold-standard based measures of insulin sensitivity (through either euglycaemic-hyperinsulinaemic clamp or insulin-suppression measurements). For each lipid trait, we created individual genetic risk allele scores weighted on known effect sizes. We used these genetic risk scores to test the causal associations between lipid traits and insulin sensitivity using instrumental variable statistics. ***Results***: We observed strong associations of the genetic risk scores for TG, HDL-C, and LDL-C with the respective circulating lipid concentrations (all *P-values*$<10^{-29}$). Lipids were strongly associated with insulin sensitivity in standard linear regression (all *P-values*$<10^{-6}$). However, there was no evidence of a genetically-determined effect of the different lipids on insulin sensitivity. The estimated causal (genetically-determined) effect sizes of lipids on insulin sensitivity were close to null and, for TG and HDL-C, different to the effects observed in standard linear regression models (TG, $P_{difference} = 3.6\times10^{-5}$; HDL-C, $P_{difference} = 1.4\times10^{-4}$). ***Conclusion***: Our results provide evidence that dyslipidemia in non-diabetic subjects is not a cause of decreased insulin sensitivity.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1625T

Dissecting the genetic architecture of coronary artery disease by genome engineering. *D.J. Segal[1], M.S. Bhakta[1], K. ThevaDas[1], C. Sun[2], N.M. Grace[3], J.A. Nolta[3], A.A. Knowlton[4], D.M. Rocke[5].* 1) Genome Ctr & Biochem & Molec Med, Univ California, Davis, Davis, CA; 2) Biomedical Engineering, Univ California, Davis, Davis, CA; 3) Institute for Regenerative Cures, Univ California, Davis, Davis, CA; 4) Internal Medicine, Univ California, Davis, Davis, CA; 5) Public Health, Univ California, Davis, Davis, CA.

One of the greatest challenges facing biomedical research since the sequencing of the human genome has been to understand the role of genetic variation in human disease. Many genetic variants have been associated with common diseases. However, determining the functional consequences of these variants has been hard. Several variants are often inherited together in tightly linked blocks, making it difficult to determine which is the causative variant. People have millions of other genetic differences, making it difficult to correlate cellular phenotypes with a particular variant. Different gene sets are expressed in different cells, but it is difficult to extract disease-relevant cells from large numbers of patients. We describe a method with the potential to revolutionize the functional analysis of genetic variation, using custom nucleases to genetically modify individual variants in induced pluripotent stem cells. This process would provide unprecedented analytical power, present the first general method to determine if a variant is causative, and analyze function disease-relevant cell types. We will focus on variants at the 9p21 region of the genome that have been associated with coronary artery disease (CAD). The methods should provide a new way to unlock the wealth of data from genome wide association studies, and to probe the genetic architecture of common diseases. We will describe our improved methods for inexpensive and rapid construction of highly-active zinc finger and TALE nucleases to examine the functional role of polymorphisms at the 9p21 CAD risk locus.

1626T

Application of Mendelian Randomization on ascorbic acid and hypertension in >11,000 participants from 4 independent studies. *K.H. Wade[1,2], N.J. Timpson[1,2], N.G. Forouhi[2,3], M-J. Brion[1,2], R.M. Harbord[1,2], D.G. Cook[4], P. Johnson[5], A. McConnachie[5], R.W. Morris[7], S. Rodriguez[7], Z. Ye[3], S. Ebrahim[8], S. Padmanabhan[6], G. Watt[9], K.R. Bruckdorfer[10], N.J. Wareham[3], P.H. Whincup[4], S. Chanock[11], N. Sattar[6], D.A. Lawlor[1,2], G. Davey-Smith[1,2].* 1) MRC Centre for Causal Analyses and Translational Epidemiology, University of Bristol, Bristol, UK; 2) Department of Social and Community Medicine, University of Bristol, Bristol, UK; 3) MRC Epidemiology Unit, Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge, UK; 4) Division of Community and Health Sciences, St George's University of London, London, UK; 5) Robertson Centre for Biostatistics, University of Glasgow, Glasgow, UK; 6) British Heart Foundation Glasgow Cardiovascular Research Centre, Faculty of Medicine, University of Glasgow, Glasgow, UK; 7) Department of Primary Care & Population Health, University College London, London, UK; 8) London School of Hygiene and Tropical Medicine, London, UK; 9) General Practice and Primary Care, Division of Community Based Sciences, University of Glasgow, Glasgow, UK; 10) University College London, London, UK; 11) Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD, USA.

Observational studies show that ascorbic acid is negatively associated with blood pressure (BP). However, these studies are susceptible to confounding, bias and reverse causation. The application of Mendelian randomization (MR) using genetic variants has proved helpful in determining causal pathways of modifiable risk factors and disease related outcomes. Genetic variants in the SLC23A1 locus, encoding the sodium-dependent ascorbic acid transporter-1, have been identified and are associated with ascorbic acid. Work here aimed to assess the association between ascorbic acid and BP using SLC23A1 SNP (rs6596473) and SLC23A1 SNP (rs33972313) as instruments for ascorbic acid in 4 independent cohorts. Initial analyses were completed in the British Women's Heart and Health Study (BWHHS, n=4286) with replication in the British Regional Heart Study (n=3945), MIDSPAN (n=1477) and Ten Towns (n=1531) studies. In BWHHS, per-allele variation in rs6596473 and rs33972313 was associated with 2.69μmol/L increase (95%CI:1.22,4.16;P=0.0003) and 3.83μmol/L reduction (95%CI:-7.50,-0.17;P=0.04) in ascorbic acid respectively. Per-allele effect of rs6596473 and rs33972313 on ascorbic acid was 1.26μmol/L (95%CI:-0.23,2.74;P=0.1) and -3.99μmol/L (95%CI:-5.95,-2.03;P<0.0001) respectively, across all studies. In BWHHS, observational analysis showed that systolic BP (SBP) decreased by 0.76mmHg (95%CI:-1.47,-0.06;P=0.03) and diastolic BP (DBP) decreased by 0.08mmHg (95%CI:-0.42,0.26;P=0.64) per quartile increase in ascorbic acid. Meta analysis showed that per quartile increase in ascorbic acid was associated with -0.63mmHg (95%CI:-0.94,-0.32;P<0.0001) SBP and -0.37mmHg (95%CI:-0.84,0.09;P=0.11) DBP. Adjusting for confounders had little effect (pooled associations: -0.66mmHg SBP (95%CI:-1.05,-0.26;P=0.001) and -0.35mmHg DBP (95%CI:-0.80,0.09;P=0.12)). Instrumental variable analysis in BWHHS using a score of ascorbic acid increasing variants from both SNPs showed no evidence of increasing ascorbic acid causing a decrease in BP (2.36mmHg SBP (95%CI:-11.13,15.84;P=0.73) and 3.85mmHg DBP (95%CI:-2.96,10.67;P=0.27)). Here, MR gave evidence against the negative causal association between ascorbic acid and BP, suggesting that supplementation use may not improve BP. Due to weak instrument bias (F-statistic=10.21), further work with larger sample sizes using variants with greater per-allele effect is needed to understand the causal pathways underlying associations between ascorbic acid and BP.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1627T

Identification of a regulatory variant that binds C/EBPβ at the *GALNT2* human high-density lipoprotein cholesterol locus. *T.S. Roman[1], M.P. Fogarty[1], S. Vadlamudi[1], A.F. Marvelle[1], K.J. Gaulton[1], A.J. Gonzalez[1], Y. Li[1,2], K.L. Mohlke[1].* 1) Genetics, Univ North Carolina, Chapel Hill, Chapel Hill, NC; 2) Biostatistics, Univ North Carolina, Chapel Hill, Chapel Hill, NC.

Genome-wide association studies have identified numerous loci associated with blood lipid and cholesterol levels; however, many of the molecular and biological mechanisms underlying these association signals remain unknown. Common variants (MAF >5%) within intron 1 of *GALNT2* (encoding an N-acetylgalactosaminyltransferase) were found to be associated with HDL-C level. We sought to identify the molecular mechanism and functional variants underlying the association signal. 23 variants exhibit strong linkage disequilibrium (r^2>.8) with an HDL-C-associated SNP, and 16 of them overlap open chromatin, active histone modifications, and/or transcription factor ChIP-seq peaks. To examine the correlation of *GALNT2* HDL-C-associated variants with *GALNT2* expression, we measured allelic expression in hepatocyte RNA from 36 individuals heterozygous for an HDL-C-associated SNP. The HDL-C-increasing alleles were associated with a 7% increase in *GALNT2* intron RNA expression (P= 5.4×10^{-7}). The direction of expression effect is consistent with other *GALNT2* expression QTLs in human liver and carotid plaque. To determine which of the 23 variants play a regulatory role in *GALNT2* expression, we tested all of them in transcriptional reporter assays in HepG2 cells. A 3-variant haplotype of rs4846913, rs2144300, and rs6143660 and one additional SNP, rs2281721, showed allele-specific enhancer activity. Through site-directed mutagenesis, we found that at least two of the SNPs in the haplotype, rs4846913 and rs2144300, contribute to increased transcription. We assessed whether differential transcription factor binding influenced the allele-specific transcriptional activity. In EMSAs, the HDL-C-increasing allele of rs4846913 showed increased C/EBPβ binding. We confirmed differential binding of C/EBPβ to rs4846913 *in vivo* using allele-specific ChIP in HepG2 and Huh-7 cells. These data are consistent in terms of direction of effect; the HDL-C-increasing alleles, one of which binds C/EBPβ, are associated with increased *GALNT2* intron RNA expression and transcriptional activity. These data identify rs4846913 as a likely functional variant and provide evidence of a molecular mechanism at *GALNT2*. The data are inconsistent with the direction of effect on HDL-C level exhibited by a rare loss-of-function human variant and altered *Galnt2* liver expression in mice. Further work is needed to elucidate the mechanisms by which the *GALNT2* GWAS variants influence HDL-C level.

1628T

Maternally transmitted coronary heart disease is associated with the mitochondrial tRNAThr 15927G>A mutation. *M.X. Guan.* Genetics, Zhejiang University, Hangzhou, Zhejiang, China.

Mitochondrial DNA (mtDNA) mutations are associated with a wide spectrum of human diseases. However, the role of mitochondrial dysfunction in coronary heart disease (CHD) remains poorly understood. In this report, thirty six subjects from a Chinese family with maternally inherited (CHD) underwent clinical, genetic, molecular and biochemical evaluations. Nine of 13 members with CHD, three matrilineal relatives had CHD, hypertension and hyperlipidemia, 6 matrilineal relatives suffered from only CHD, two non-matrilineal relatives had hypertension and two non-matrilineal relatives had hypercholesterolemia. Mutational analysis of their mtDNA identified the homoplasmic 15927G>A mutation located at a highly conserved base-pairing (28C-42G) on the anticodon stem of tRNAThr. This mutation may cause structural and functional alteration of this tRNA. Approximately 80% and 39% reductions in the steady-state level and aminoacylated efficiency of tRNAThr were observed in the cells derived from 5 matrilineal relatives carrying the 15927G>A mutation, respectively. However, this mutation itself causes 44% reduced level but did not affect aminoacylated efficiency of tRNAThr, implicating defects in nuclear genes involved in tRNA metabolism. ~53% reduction in the rate of mitochondrial translation was observed in cells carrying the 15927G>A mutation. Impaired mitochondrial translation is apparently a primary contributor to the reductions in the rate of overall respiratory capacity, malate/glutamate-promoted, succinate/glycerol-3-phosphate-promoted, or N,N,N',N'-tetramethyl-p-phenylenediamine/ascorbate-promoted respirations and the increasing level of reactive oxygen species in cells carrying the 15927G>A mutation. These data provide the direct evidence that mitochondrial dysfunction caused by 15927G>A mutation is involved in CHD. Our findings may provide new insights into pathophysiology of maternally transmitted CHD.

1629T

Mitochondrial haplogroups and genomic ancestry showed association with heart failure in Brazilians. *C. Fridman[1], M.M.S.G. Cardena[1], A.K. Ribeiro-dos-Santos[2], S.E. Santos[2], J.E. Krieger[3], A.C. Pereira[3].* 1) Department of Legal Medicine, Ethics and Occupational Health, University of São Paulo, São Paulo, Brazil; 2) Laboratory of Human and Medical Genetics, Federal University of Pará, Belém, Pará, Brazil; 3) Department of Cardiology, Laboratory of Genetics and Molecular Cardiology, Heart Institute, Medical School, University of São Paulo, São Paulo, Brazil.

Brazilian population is one of the most heterogeneous and it is interesting to analyze genetic variations which influence diseases development. In 2007, cardiovascular diseases (CD) accounted for the third leading cause of hospitalization in the public health system; among all possible causes heart failure (HF) was the most frequent. Studies have been suggested that mitochondrial haplogroups (mtDNA Hg) or genomic ancestry may contribute to development of some types of CD. Herein, we evaluated a possible association of mtDNA Hg and genomic ancestry in 511 patients with HF compared with 188 healthy individuals. mtDNA was sequenced for control region using BygDye Terminator v3.1 and polymorphisms were determined comparing the sequences with the Cambridge Reference Sequence (rCRS). A set of 48 Ancestry Informative Markers (AIMs) INDELs type from autosomal chromosomes was used to measure proportions of three different ancestries (African, European and Amerindian) and was analyzed via multiplex PCR. Population structure and estimation of global ancestry proportions analyzes were done using Structure v.2.2; statistical analysis was done with SPSS 14.0. Within patients group, 46% showed African Hg, 28% showed Ameridian/Asian Hg and 26% showed European Hg. In relation to AIMS, the means of European ancestry was 57%±22%, Amerindian ancesty was 14%±10%, and African ancestry was 28%±21%. For control group, 36% showed African Hg, 36% presented Ameridian/Asian Hg and 28% showed European Hg. The AIMS showed the mean of European ancestry was 61%±19%, Amerindian ancestry was 14%±9% and African ancestry was 24%±18%. These results showed association for African haplogroup (OR1,55 [95%I.C(1,09-2,18) p=0,013]) and African ancestry (OR 2,78 [95%I.C(1,18-6,54) p=0,019) with risk to development HF. Interestingly, the Amerindian haplogroup (OR 0,69 [95%I.C(0,48-0,98) p=0,042]) and European ancestry (OR0,43 [95%I.C(0,19-0,97) p=0,043]) conferred protection to HF. The risk factor showed for African haplogroup and ancestry is consistent with clinical practice where there is a higher prevalence and worse prognosis of CD in black patients. The protection observed for Amerindian haplogroup and European ancestry is a new data and could be attributed to environmental factors and/or lifestyle. In the next stage of this study we will analyze separately four different etiologies of HF find within our sample to check the results in patient's substructure. Supported: FAPESP, LIM40-HC.

1630T

RNAi-based functional profiling of loci from blood lipid genome-wide association studies. *H. Runz[1,3], P. Blattmann[2,3], Ch. Schuberth[1,3], R. Pepperkok[2,3].* 1) Institute of Human Genetics, University of Heidelberg, Heidelberg, Germany; 2) European Molecular Biological Laboratories (EMBL), Heidelberg, Germany; 3) Molecular Medicine Partnership Unit (MMPU), Heidelberg, Germany.

Genome-wide association studies (GWAS) have proven powerful to unravel genomic loci associated with common traits and complex human disease. However, GWAS only rarely reveal information on the exact genetic elements and pathogenic events underlying an association. In order to extract functional information from genomic data, strategies for an objective follow-up of GWAS-findings on a phenotypic level are required. Here we address these limitations by applying RNAi to systematically analyze >100 candidate genes within 55 loci identified by GWAS as associated with blood lipid levels, coronary artery disease and/or myocardial infarction for a function in regulating cholesterol levels in cells. The genes were knocked-down with siRNAs and the consequences on cellular free cholesterol (FC) and the efficiency of LDL-internalization into cells were quantified using automated microscopy and multiparametric image analysis. We will show evidence that loss-of-function of a surprisingly high number of the trait-associated genes affected LDL-uptake, FC, or both. For several genes without previously known lipid-regulatory roles the functional effects upon gene knockdown and overexpression were analyzed and in several instances closely correlated with altered LDL-receptor levels. By providing strong evidence for disease-relevant functions of lipid trait-associated genes our study demonstrates that quantitative, cell-based RNAi is a scalable strategy for a systematic, unbiased detection of functional effectors within GWAS loci.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1631T

A fine mapping of the genetic variation influencing the ratio of alternatively spliced *CD40* transcripts and conferring susceptibility to Kawasaki disease. *Y. Onouchi[1,4,17], K. Ozaki[1], M. Terai[2], H. Hamada[2], H. Suzuki[3], T. Suenaga[3], Y. Suzuki[4], K. Yasukawa[2], R. Ebata[5], T. Saji[6,17], Y. Kemmotsu[6], K. Ouchi[7], F. Kishi[8], T. Yoshikawa[9], T. Nagai[10], K. Hamamoto[11], Y. Sato[12], K. Sasago[4], A. Takahashi[13], M. Kubo[14], T. Tsunoda[15], A. Hata[4], Y. Nakamura[16], T. Tanaka[1].* 1) Lab Cardiovascular Disease, Ctr Genomic Med RIKEN, Yokohama Kanagawa, Japan; 2) Dept. Pediatrics, Tokyo Women's Medical University, Yachiyo Medical Center, Yachiyo, Chiba, Japan; 3) Dept. Pediatrics, Wakayama Medical University, Wakayama, Wakayama, Japan; 4) Dept. Public Health, Chiba University Graduate School of Medicine, Chiba, Chiba, Japan; 5) Dept. Pediatrics, Chiba University Graduate School of Medicine, Chiba, Japan; 6) Dept. Pediatrics, Toho University School of Medicine, Tokyo, Japan; 7) Dept. Pediatrics, Kawasaki Medical School, Kurashiki, Okayama, Japan; 8) Dept. Molecular Genetics, Kawasaki Medical School, Kurashiki, Okayama, Japan; 9) Dept. Pediatrics, Fujita Health University, Toyoake, Aichi, Japan; 10) Dept. Pediatrics, Dokkyo Medical University Koshigaya Hospital, Koshigaya, Saitama, Japan; 11) Dept. Occupational Therapy, International University of health and welfare, Fukuoka, Fukuoka, Japan; 12) Dept. Pediatrics, Fuji Heavy Industry LTD. Health Insurance Society General Ohta Hospital, Ohta, Gunma, Japan; 13) Lab. Statistical Analysis, Center for Genomic Medicine, RIKEN, Yokohama, Kanagawa, Japan; 14) Lab. Genotyping Development, Center for Genomic Medicine, RIKEN, Yokohama, Kanagawa, Japan; 15) Lab. Medical Informatics, Center for Genomic Medicine, RIKEN, Yokohama, Kanagawa, Japan; 16) Lab. Molecular Medicine, Human Genome Center, Institute of Medical Science, University of Tokyo, Tokyo, Japan; 17) Japan Kawasaki Disease Genome Consortium.

Kawasaki disease (KD) is an acute systemic vasculitis syndrome of infants and young children. Recently we carried out a genome-wide association study (GWAS) and one of the association peaks was seen on rs4813003, a SNP located 4.9kb downstream of *CD40* gene. It is known that there is a functional SNP of CD40 gene. The SNP (rs1883832) is located within the Kozac sequence and alters translation efficiency of CD40 protein. As GWASs for other adult autoimmune diseases revealed associations peaked at SNPs linked with rs1883832, it has been considered that modulated activity of CD40 protein underlies the pathogenesis of these diseases. To see whether rs1883832, which was not studied in our GWAS, is the responsible variant of *CD40* gene region also for KD, we assessed the association of the SNP in our Japanese KD cases (n = 1182) and controls (n = 4326). Pairwise LD analysis revealed that there is a moderate LD between rs1883832 and rs4813003 (D' = 0.86, r^2 = 0.55). Associations in logistic regression analyses were more significant for rs4813003 (OR = 1.33, 95%CI = 1.20–1.47, $P = 4.1 \times 10^{-8}$) than that for rs1883832 (OR = 1.27, 95%CI = 1.15–1.40, $P = 1.5 \times 10^{-6}$). When these 2 SNPs were studied by a conditional analysis, *P* value for rs4813003 remained significant (P = 0.003) but that for rs1883832 diminished (P = 0.25). These data indicated that the association in this region was driven by rs4813003 (or its proxy SNPs) and the responsible variant must be involved in the pathogenesis of KD by a mechanism other than alteration of protein translation efficiency. Then next, we conducted quantitative PCR to compare *CD40* transcript level as well as ratio of two major isoforms generated by alternative splicing using B-lymphoblastoid cell lines of different genotypes (n = 139). The number of risk alleles at rs4813003 was not correlated with gross *CD40* mRNA expression (P = 0.81) but significantly correlated with reduced ratio of the shorter isoform ($P = 4.9\times10^{-9}$) which is translated into a truncated protein. Re-sequencing study performed in search for the variation responsible to the regulation of the exon6 inclusion/exclusion identified two plausible candidate SNPs which are in high LD (r^2 = 0.96 ~ 0.98) with rs4813003 and located within intron 5. Further investigation is warranted to specify the responsible variant and to elucidate the mechanism by which the responsible variant is involved in the pathogenesis of KD.

1632T

The AA genotype of Tumor Necrosis Factor Induced Protein 2 (TNFAIP2) rs8126 polymorphism is associated with increased mortality in septic shock. *SA. Thair[1], JH. Boyd[1], TA. Nakada[2], JA. Russell[1], KR. Walley[1].* 1) Medicine, University of British Columbia, Critical Care Research Laboratories, Vancouver, British Columbia, Canada; 2) Chiba University Graduate School of Medicine, Department of Emergency and Critical Care Medicine, Chiba, Japan.

One of the central mediators of the pathophysiology of septic shock is the transcription factor nuclear factor κB (NF-κB). TNFα induced NF-κB signaling is one of the earliest signaling pathways up regulated in septic shock and yet the genetics of this pathway are not fully understood. We tested the hypothesis that genetic variation in inflammatory mediators of the TNFα induced NF-κB pathway are associated with increased 28-day mortality in patients who have septic shock. HeLa cells were treated with or without TNFα for 4 hours. Gene expression was interrogated using the Illumina Human HT-12 array. Data was normalized and analyzed with FlexArray (v.1.4.1). tagSNPs were selected using a linkage disequilibrium based method (r^2 threshold =0.65, MAF >5%) and were genotyped in all patients of the Vassopressin and Septic Shock Trial (VASST) cohort. We interrogated 68 SNPs in the region +/− 50,000bp of TNFAIP2 in 530 Caucasian VASST patients using an Armitage trend test. HeLa cells were then transfected with either a TNFAIP2 overexpression or control vector for 48 hours. Protein phosphorylation was then measured in cell lysate (KINEXUS KPSS 1.3 Assay). Gene expression in HeLa cells and a literature search for replication of TNFα induced gene expression identified three genes common to all data sets: TNFAIP2, NFKBIA and NFKB1. Genotyping of 13 tagSNPs in the three genes revealed TNFAIP2 rs8176373 to be highly associated with mortality in the VASST cohort of septic shock patients (p=0.0019, Bonferroni corrected p=0.03). Of the additional 68 SNPs nearby, rs8126 was most highly associated with mortality (p= 0.007). Septic shock patients with the GG (homozygous minor) genotype of rs8126 had increased 28-day mortality (relative risk 0.73 (95% CI 0.59–0.82, p=0.0065)). At 24 hours of septic shock, VASST patients of the GG/ AG genotypes have increased plasma CXCL10 compared to those of the AA genotype (n=618, p=0.014). In vitro studies of the over expression of TNFAIP2 in HeLa cells resulted in significant decreases in phosphorylation of MAPK signaling pathway members, of note, phosphorylation of Raf-1 S259 (132% decrease compared to control). We conclude that the G allele of TNFAIP2 rs8126 is associated with increased septic shock mortality, possibly due to altered TNFAIP2 function and/ or increased pro-inflammatory cytokine CXCL10 production. As well TNFAIP2 may influence phosphorylation of the MAPK signaling pathway.

1633T

Significance of the NODAL p.H165R variant in complex cardiovascular malformations: A report of two lethal cases with an asymptomatic parent. *Z. Ammous[1,2], A. Rajadhyaksha[2], P. Jayakar[2].* 1) University of Miami, Miami, FL; 2) Miami Children's Hospital, Miami, FL.

Heterotaxy is a heterogeneous genetic disorder. Mutations in 7 genes (NODAL, ZIC3, CFC1, FOXH1, LEFTY2, GDF1, ACVR2B) involved in NODAL signaling have been proposed to cause heterotaxy and/or congenital heart defects (CHD). These genes play an essential role in establishing left-right patterning during organogenesis. Heterotaxy visceral 5, an autosomal dominant disorder with variable penetrance, is caused by loss of function mutations in the TGFβ-related ligand NODAL gene. Missense, nonsense, in-frame insertion/deletions and splice site mutations have been reported while certain polymorphisms (including H165R) were found to be functionally abnormal but their significance in humans is not completely understood. We report the case of a full term baby girl who presented with respiratory distress shortly after birth and was found to have complex CHD (dextrocardia and total anomalous pulmonary venous return). Ultrasounds revealed heterotaxy with asplenia. She had a complicated course requiring multiple surgeries and eventually death at the age of 4 months due to sepsis and multi-organ failure. Karyotype and chromosomal microarray analysis were normal. A Heterotaxy panel sent to Cincinnati Children's Heart Institute revealed the patient to be homozygous for a common polymorphism, c.494A>G (H165R) in exon 2 of the NODAL gene. Family history was significant for a paternal half-brother who died at the age of 7months with Heterotaxy and asplenia but no molecular testing was available at that time. Paternal targeted mutation analysis confirmed the father, who has been asymptomatic so far, was carrying the same variant in cis configuration. The H165R polymorphism is frequently identified in the population in its heterozygous form and does not cause disease. However, it has been shown by functional analyses to reduce NODAL activity and therefore may modify the risk for CHD or serve as a susceptibility allele in combination with other genetic or environmental factors. The risk of being homozygous for this variant is not known. It has been noted in a severe case of Holoprocencephaly and CHD where the patient was homozygous for H165R in cis position. We suggest that being homozygous for the variant in cis configuration in addition to other modifier genes could attribute to the severe phenotype in our patient and her sibling. The interpretation of this variant may change in the future if similar cases are reported or if these modifying genes could be identified.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1634T

Parent of origin effect in HDL levels. *J.S. Lilley[1], I.M. Predazzi[2], W.S. Bush[2], M.F. Linton[3], S.M. Williams[2], S. Fazio[3].* 1) Pediatric Endocrinology, Vanderbilt University, Nashville, TN; 2) Center for Human Genetics Research, Vanderbilt University, Nashville, TN; 3) Division of Cardiology, Vanderbilt University, Nashville, TN.

Even though heritability estimates for lipid traits are high (40–50% for HDL), genetic contributions to lipid expression are complex and are not fully understood. Part of this variability may be due to epigenetic influences such as genomic imprinting; maternal hypercholesterolemia is thought to influence in offspring cardiovascular outcomes even in adult life. Such epigenetic factors can be observed as parent of origin effects within families. To explore this mechanism in lipid traits, we examined natural lipid variation that can be explained by parent of origin effects in the general population. We performed linear regressions in the second and third generations from the Framingham Cohort (n= 9,219 participants) to compare lipid trait correlations between mothers and daughter, mothers and sons, fathers and daughters, and fathers and sons with the goal of identifying sex-specific transmission effects. Models were adjusted for BMI, age, smoking status, treatment, and menopausal status of both parent and offspring. We performed a likelihood ratio test to compare a reduced model (containing offspring and parent covariates only) to a full model that includes parent HDL as a risk factor. As anticipated, we observe that a significant amount of offspring HDL is explained by parent HDL levels. However, we also note that maternal HDL levels explain significantly more variance in daughter HDL compared to other models of other parent/offspring pairs, with the full model showing R2=0.19 versus a reduced model with R2=0.12 (chi2 = 104.9 and p=1.285e-24). By contrast, the relationship between maternal HDL explains only 2% of the son's HDL (LR= 19.74, p<8.872e-6). To investigate the source of this transmission effect, we conducted a transmission disequilibrium test between mothers and daughters only. Variants within several genes show significant transmission distortion (p<1e-6). SLC16A7, previously associated with HDL in genome wide case-control association studies; SH3GL3, previously reported as associated with height and BMI; GLCCI1, induced by glucocorticoids; CSF1, a cytokine that controls the proliferation and differentiation of macrophages; and PSMD1, involved in peptide and MHC peptides processing. Biologically, these results suggest that maternal factors have a strong influence on offspring HDL and that this influence is sex-specific. These findings indicate that for females, maternal lipid levels are an important factor for predicting HDL levels.

1635T

Loci on chromosomes 2q12, 6p21 and 13q14 modulate the effect of smoking on carotid intima-media thickness in whites, blacks and Hispanics. *C. Dong[1], L. Wang[2], A. Beecham[2], D. Cabral[1], C.B. Wright[1,3], S.H. Blanton[2], T. Rundek[1,3], H. Zhao[4], R.L. Sacco[1,2,3].* 1) Dept. of Neurology, Univ. of Miami, Miami, FL; 2) Dept. of Human Genetics, Univ. of Miami, Miami, FL; 3) Dept. of Epidemiology and Public Health, Univ. of Miami, Miami, FL; 4) Dept. of Epidemiology and Public Health, Yale Univ. New Haven, CT.

Carotid artery intima-media thickness (IMT) is a recognized marker of atherosclerosis and smoking is a well-known vascular risk factor, but whether genetic variants modulate the smoking effect on IMT is unknown. We performed a genome-wide smoking-by-SNP interaction analysis in 1,010 individuals (15% white, 17% black, 66% Hispanic), who underwent carotid B-mode ultrasonography and genotyping with the Affymetrix 6.0 chip. Smoking was defined as 0, < 20, and ≥20 cigarette pack-years. An additive genetic effect was assumed for each SNP based on the minor allele number. A total of 722,379 SNPs with a minor allele frequency at least 5% were included in the analysis. Multiple linear regression analysis was conducted to test for smoking-by-SNP interaction on the segment-specific IMT while controlling for age, sex, and the top 3 principal components estimated to capture ancestry by EIGENSTRAT. The strongest interactions were found for 5 SNPs on 2q12, 6p21 and 13q14 for IMT in the internal carotid artery (ICA) (interaction p<5.0E-7) but not in the common carotid artery or bifurcation. Two SNPs on 13q14 near or within *RCBTB1* (including rs7331252) achieved a genome-wide significance (beta±se for smoking-SNP interaction: 0.04±0.01 mm; p=1.2E-8; the adjusted means of ICA IMT were similar among GG-carriers with 0.88, 0.88 and 0.87 mm, respectively, for 0, <20, and ≥20 cigarette pack-years, but substantially increased among AA carriers with 0.85, 0.92 and 1.00 mm for 0, <20, and ≥20 cigarette pack-years) in the total cohort and similar effects in whites (0.04±0.02 mm; p=0.06), blacks (0.04±0.01 mm; p=0.001) and Hispanics (0.03±0.01 mm; p=4.6E-6). SNP rs4851822 on 2q12, which is located near *NCK2*, had an interaction of p= 8.4E-8 (-0.04±0.01 mm) in the total cohort and similar effects in whites (-0.04±0.02 mm; p=0.007), blacks (-0.05±0.02 mm; p=0.004) and Hispanics (-0.03±0.01 mm; p=0.002). The other top 2 SNPs, located between *MDGA1* and *ZFAND3*, including rs6908598, had an interaction effect of p=7.4E-8 (0.05±0.01 mm) in the total cohort with similar effects in whites (0.05±0.02 mm; p=0.02) and Hispanics (0.05±0.01 mm; p=4.6E-6) but not in blacks (0.02±0.02 mm; p=0.37). The interaction effects remained very similar after further adjustment for hypertension, diabetes and hypercholesterolemia. Given the consistent findings across race-ethnic groups, further studies are warranted to perform in-depth investigations of functional genetic variants in these regions.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1636T

A Genome-wide Gene-by-environment Scan for Mobile Source Air Pollution in Atherosclerosis and Myocardial Infarction Identifies PIGR and HCN4. *C.K. Ward-Caviness[1,2], L. Neas[3], C. Haynes[2], C. Blach[2], E. Burns[2], K. LaRocque-Abramson[2], E. Dowdy[2], W. Cascio[4], R. Devlin[4], D. Diaz-Sanchez[4], W.E. Kraus[2,3], S.H. Shah[2,3], S.G. Gregory[2], M.L. Miranda[5], E.R. Hauser[2,3].* 1) Computational Biology, Duke University, Durham, NC; 2) Center for Human Genetics, Duke University Medical Center, Durham, NC; 3) Department of Medicine, Duke University Medical Center, Durham, NC; 4) Environmental Protection Agency, Chapel Hill, NC; 5) University of Michigan, Ann Arbor, MI.

Mobile source air pollution (MSAP) is a prevalent environmental exposure with known adverse cardiovascular effects. To investigate the role of genes and genetic variants in the pathological process triggered by MSAP exposure, we performed race stratified gene-environment interaction GWAS (GWAS) using distance from primary residence to the nearest roadway as a proxy for MSAP exposure. We performed GWAS for recent myocardial infarction (MI) and atherosclerosis, using an ordinal measure of atherosclerosis burden (CADindex), in a large cardiac catheterization cohort. To identify possible associations we first identified the 10 strongest gene-environment interaction associations. Among these 10 we removed SNPs not in genes (introns, exons or 3'-UTR) and examined these biologically relevant signals for consistency of direction and replication in separate African-American (AA) and European-American (EA) cohorts. The GWAS for MI among the EA (N = 1647) revealed three strong gene-environment interaction associations: RS1400207 (CNTN4, P = 1.3×10-5), RS2623997 (HCN4, P = 2.0×10-5), and RS925820 (CNTN4, P = 2.4×10-5). Of these RS2623997 was consistent in direction in the AA (P = 0.084, meta-analysis P = 1.6×10-4). As HCN4 is necessary for proper cardiac pacemaker function and associated with sick sinus syndrome 2, it is a strong candidate for a gene involved in MI. For the CADindex GWAS, three of the 10 strongest associations were in genes including RS291096, a synonymous mutation in PIGR in AA. An examination of other SNPs in PIGR revealed multiple SNPs in exons with nominally significant (P < 0.05) p-values for the gene-environment interaction term. Among these was RS291102 a missense mutation identified as probably pathogenic (A -> V, minor allele frequency (MAF) = 0.35, P = 0.0025). No PIGR SNPs replicated in the EA CADindex GWAS, but the direction of association was consistent for RS291096 (meta-analysis P = 0.0016) and RS291102 (meta-analysis P = 0.092). The MAF of RS291102 is 0.033 in the EA, making detection of a gene-environment interaction signal difficult in this cohort. These results show that genetic variants are associated with MI and atherosclerosis burden under exposure to MSAP. Future work will increase the sample size and add gene expression and metabolic profiling to further investigate joint effect of specific genes and MSAP exposure on MI and atherosclerosis.

1637T

Genome-wide association analysis of blood pressure incorporating gene-age interactions in the CHARGE, GBPgen, and ICBP Consortia. *J. Simino, CHARGE, GBPgen, and ICBP Investigators.* Division of Biostatistics, Washington University School of Medicine, St. Louis, MO.

BACKGROUND: Hypertension afflicts 26% of adults worldwide and accounts for 10% of global healthcare expenditures. Massive genome-wide association studies (GWAS) of single nucleotide polymorphism (SNP) main effects have successfully identified ≥50 blood pressure (BP) associated loci, however they collectively explain a small proportion of BP variance. We investigate whether additional loci can be discovered through an analysis of gene-age interactions using 9 studies (N=55,796) from the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium.

METHOD: Each CHARGE study stratified participants into 10-year age bins. GWAS main effects analyses were conducted within each cohort and age bin by regressing systolic (SBP), diastolic (DBP), mean arterial (MAP), and pulse pressure (PP) onto allele dosage, age, age-squared, body-mass-index, gender, and field center (28 analyses per trait). After applying genomic control, we conducted a fixed effects meta-regression of the SNP coefficient onto the median age of participants within each study and age bin. Significant findings using genomic-controlled chi-square tests were replicated in 43,493 participants of European descent and 8,682 Chinese, Indian, and Malay individuals from Singapore.

RESULTS: Three loci reached genome-wide significance (p-value<5×10-8) for previously unassociated BP traits via joint tests of the SNP main effect and SNP-age interaction (2 df). SNPs rs880315 (*CASZ1* intron), rs2004776 (*AGT* intron), and rs7599598 (*FER1L5* missense mutation) were associated with SBP, MAP, and DBP, respectively. Rs7233332 (*RAB31* intron) exhibited no main effect but significantly interacted with age to influence PP. Rs6782531 (*FHIT* intron) displayed a joint main effect and SNP-age interaction on SBP in 46,939 participants aged 40–70 years. Singapore participants aged 40–70 years replicated the main effects of rs2004776 on MAP (p-value=0.003) and rs7599598 on DBP (pvalue=0.005).

CONCLUSION: Age is a convenient surrogate for an array of complex biological and environmental phenomena that influence BP gene expression. Capitalizing on gene-age interactions using GWAS and next-generation sequencing data may enhance the identification of novel genes associated with common complex traits.

1638T

A genome wide association study of carotid intima-medial thickness, a preclinical marker of atherosclerosis. *S.H. Blanton[1,2], A.H. Beecham[1], L. Wang[1], C. Dong[2], D. Cabral[2], B.I. Hudson[2], M. Kiryiya[3], F. Jasmine[3], B. Pierce[3], R. Demmer[3], A. Habibul[3], M. Desvarieux[4], R. Sacco[1,2], T. Rundek[2].* 1) Dept of Human Genetics, Univ Miami, Miami, FL; 2) Dept of Neurology, Univ Miami, Miami, FL; 3) Dept of Health Studies, Univ Chicago, Chicago, IL; 4) Columbia Univ, New York, NY.

Carotid intima-medial thickness (CIMT) is a recognized risk factor for cardiovascular disease. The genetic underpinnings of CIMT are not fully understood. The aim of this study was to identify potential novel risk loci for CIMT by performing a genome-wide association study. CIMT for 1278 individuals from the population-based Northern Manhattan Study (NOMAS;1066) and the Oral Infections and Vascular Disease Epidemiology Study (INVEST;212) was measured by high-resolution B-mode ultrasound and expressed as the total IMT defined as mean of the maximum (IMTm). Other carotid segment-specific IMT phenotypes, bifurcation-BIF, common-CCA and internal carotid artery-CCA, were also examined. Genotyping was done in NOMAS using the Affymetrix SNP array 6.0 and in INVEST using Illumina 330k. With HapMap3 as the reference, IMPUTE2 was used to impute the NOMAS SNPs in INVEST which had not been genotyped, resulting in 598557 high quality SNPs (INFO ;≥ 0.4). SNPs were evaluated using linear regression first in Hispanics (795) and then in Blacks (263) and Whites (220) for replication. The model controlled for age and sex. Additionally, PCA1 and 2 from Eigenstrat were used to control for substructure in Hispanics and PCA1 was used in Blacks and Whites. We observed strong associations (p<1.0×10-5) in Hispanics that replicated (p<0.05) in either Blacks or Whites on chromosomes 3, 4, 14, 15 and 19. Five SNPs in the 4p15.31 region were associated with p<1.0×10-5 for BIFm and IMTm. One SNP, rs12507937, met genome-wide significance in Hispanics (p=4.48×10-8) and replicated in whites (p=0.016) for BIFm. Although no genes are located in this region, previous studies link this locus to type 2 diabetes and cholesterol levels. Interestingly, numerous SNPs identified on chromosome 14 were located in the T cell receptor alpha locus (TRA@). TRA are key genes in the activation of T-cells in the adaptive immune response, and play a role in regulating the inflammatory state of plaque. Lastly, 2 SNPs on chromosome 13 in the integrin beta-like protein (ITGBL1) gene were associated with BIFm in Hispanics alone. The integrin family of proteins is involved in cell adhesion and immune responses and critically involved in the cellular process of atherosclerotic plaque initiation and progression. In conclusion, we have identified several novel gene regions associated with CIMT which may further underpin the role of the adaptive and innate immune response in atherosclerosis.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1639T

Genome-wide association analyses of long-term blood pressure traits. *S.K. Ganesh[1,2], G.B. Ehret[1,3], A. Chakravarti[1] on behalf of the CHARGE-BP and ICBP consortia.* 1) McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University, Baltimore, MD 21205; 2) Department of Internal Medicine, Division of Cardiovascular Medicine, University of Michigan Medical Center, Ann Arbor, Michigan 48109; 3) Cardiology, Department of Specialties of Internal Medicine, Geneva University Hospital, Rue Gabrielle-Perret-Gentil 4, 1211 Geneva 14, Switzerland.

Blood pressure (BP) is an important cardiovascular trait, and arterial exposure to high BP is a major, modifiable risk factor for atherosclerotic diseases and cardiovascular mortality. Gene discovery is an important route to understand its mechanisms. BP has significant intra-individual variability and there are well-described sources of errors for noninvasive BP measurements. Most genetic studies utilize single visit traits but these do not correct for variability that occurs over days, months, and years. We conducted a genome-wide association study (GWAS) of four BP traits, systolic (SBP), diastolic (DBP), mean arterial (MAP), and pulse pressure (PP), for the purpose of improved gene discovery by increasing phenotypic accuracy through long-term averaging (LTA) of multiple measurements acquired over years. We also studied the statistical effects of such averaging. We identified 40 trait-variant associations across 20 independent loci associated with these traits (13 SBP, 10 DBP, 12 MAP and 5 PP loci with SNP association at P-value < 5×10-8) in individuals of European ancestry, of which four BP associations are novel. The novel loci include associations of SBP at 2p23 (rs1275988, near KCNK3), DBP at 2q11.2 (rs7599598, in FER1L5), and PP at chromosome 6p21 (rs10948071, near CRIP3) and 7p13 (rs2949837 near IGFBP3). Monte Carlo simulation studies to examine the effect of the averaging approach used here showed ~20% gain in statistical power for detection of associations. Thus, we demonstrate significant enrichment of our discovery findings by using LTA BP traits through improvement of the precision of the BP trait by averaging. We also highlight regions that might contain genes associated with phenotypic variation due to physiologic changes in blood pressure in which LTA does not improve signal. The expression level of genes identified with single-visit BP analysis were compared to those identified with LTA, and we found tissue-specific expression patterns with diurnal variation, suggesting a mechanism for the lack of enrichment seen in some regions with LTA. These discovery findings increase our understanding of the genetic determinants of BP and the implications of the averaging procedure on study design and power. This suggests that, as we enter an era of sequencing for complex diseases and traits, where power may be limited for the detection of rare variants, trait averaging is a useful strategy.

1640T

A genome-wide association study in a Lebanese population identifies three new loci for Coronary Artery Disease. *M. Ghassibe[1], J. Hager[2], JB. Cazier[3], D. Platt[4], Y. Kamatani[5], M. Farrall[6], D. Gauguier[3,7], P. Zalloua[1,8].* 1) Lebanese American University, School of Medicine, Beirut 1102 2801, Lebanon; 2) CEA-Genomics Institute, Centre National de Génotypage, 91057 Evry, France; 3) The Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Headington, Oxford OX3 7BN, UK; 4) Bioinformatics and Pattern Discovery, IBM T. J. Watson Research Centre, NY 10598, USA; 5) Centre d'Etude du Polymorphisme Humain (CEPH), 75010 Paris, France; 6) IntegraGenSA, Campus 1, Genavenir 8, 91000 Evry, France; 7) INSERM, UMRS872, Centre de Recherche des Cordeliers, 75006 Paris, France; 8) Harvard School of Public Health, Boston, MA 02215, USA.

The manifestation of coronary artery disease (CAD) follows a well-choreographed series of events that includes damage of arterial endothelial cells and deposition of lipids in the sub-endothelial layers. Genome-wide association studies (GWAS) of multiple populations with distinctive genetic and lifestyle backgrounds are a crucial step in understanding CAD pathophysiology. We carried out a genome-wide association study of arterial stenosis as measured by cardiac catheterization in 4 741 individuals of Lebanese ancestry. The locus of PHACTR1 showed association with coronary stenosis in a discovery experiment with genotype data from 2 002 individuals (OR= 1.37, P=1.57×10-5). Replication in 2 739 individuals yielded positive results (OR=1.31, P=8.85×10-6), leading to genome-wide significant association in a combined analysis (OR=1.34, P=8.02×10-10). GWAS data imputed with untyped single-nucleotide polymorphisms (SNPs) from the HapMap2 reference panel (for a total of 20 225 SNPs) replicated HNRNPA3P1-CXCL12 association with CAD and identified new significant associations of type 2 diabetes (T2D) genes CDKAL1, ST6GAL1, and PTPRD. In parallel, analysis of 10 candidate SNPs with CAD and/or MI in genotypic and allelic association models using logistic regression identified a protective effect of MTHFD1 against MI, and a positive association of CDKN2A/CDKN2B/ANRIL in the 9p21 locus with MI. Our study provides additional insights into mechanisms and genes underlying CAD/MI and show the potential for new discovery from genetic association studies in the Lebanese population that has an increased susceptibility to CAD.

1641T

An assessment of association of blood pressure with DNA variants in the Long Life Family Study and other studies. *A.T. Kraja[1], J.H. Lee[2], R. Straka[3], I.B. Borecki[1], C. Kammerer[4], M.A. Province[1].* 1) Div Statistical Genomics, Dep of Genetics, Washington U, St Louis, MO, USA; 2) Taub Institute, Sergievsky Center, Columbia U., New York, NY, USA; 3) Dep. of Experimental and Clinical Pharmacology, U. of Minnesota, MN, USA; 4) Dep. of Human Genetics, U. Of Pittsburgh Graduate School of Public Health, PA, USA.

The Long Life Family Study (LLFS, N >4,700, with a mean age of 70, gen. 2 with a mean of 90 and gen. 3 of 60, range 24–110 years; stroke prevalence of 3% and diabetes 6%), represents mainly Caucasian families identified for exceptional longevity in the United States and Denmark. Genome wide associations (2.5M SNPs) were performed for blood pressure (BP). The purpose of this work was to better understand different aspects of BP and its associated variants in relation with aging, health, and longevity. Clinical trials average-class constants were used to correct BP traits for anithypertensives' use, based on labels of medications participants presented. Thus up to 17 BP traits were studied. Similar procedures were used in the Family Heart Study for comparing results (N~5,000, with 2.5M `hybrid' SNPs, split into samples with familial history of high risk for CVD and one representing general population). After medication corrections 52% of LLFS subjects and 66% of FHS were classified with hypertension (MCHT). All traits were adjusted and analyzed with standard procedures. Results from samples of LLFS All; gen. 2; and 3; FHS All; with high risk for CVD; and with a general population CVD risk, yielded relatively weak association results. Of interest was rs407724 (MAF=0.4726) of *ROS1*. Only LLFS gen. 3 results were significant for additive effects of this SNP (p=5.9e-08) on MCHT versus ALL (p=1.4e-05), and gen. 2 (p=0.788), The C allele showed a protective effect against high BP. Referring to systolic BP, corrected for medication use (MCsbp), gen. 3 had 685 (C/C) with 132.8 (se=0.9), 1476 (C/T) with 135.0 (0.6), and 850 subjects (T/T) with a mean of 136.8 (0.8) mmHg, thus with almost 2 mmHg decrease for each copy of C allele. In contrast, in gen. 2, 356 subjects (C/C) were with 152.5 (1.5), 730 with 153.7 (1.0) and 435 subjects with a mean of MCsbp of 155.1 (1.3). After excluding married-ins from gen. 3, 524 subjects remained C/C with a mean MCsbp of 131.0 (1.0), 1103 as C/T with a mean of 134.8 (0.7) and 628 T/T with a mean of 136.0 (0.9) mmHg. This finding was not replicated with FHS samples. Yet, *ROS1*, a receptor tyrosine kinase is expressed relatively high in the kidney and intestines; It interacts among others with *PTPN* and *EGFR* genes; One intron-rs172409 is reported in association with heart rate; downstream 0.5M bps, rs654128 is reported to associate with telomere length (p=3.1e-06) and upstream rs281868 ~1M bps is associated with heart rate (p=3.9e-10).

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1642T

Novel Tissue-Specific Transcriptomic Signatures Revealed by Experimental Endotoxemia in Healthy Human Subjects. *Y. Liu[1], J. Feguson[2], I. Silverman[3], B. Gregory[3], M. Li[1], M. Reilly[2,4].* 1) Department of Biostatistics and Epidemiology; 2) Cardiovascular Institute; 3) Department of Biology; 4) Institute for Diabetes Obesity and Metabolism, University of Pennsylvania School of Medicine, Philadelphia, PA 19104.

Deregulated activation of innate immunity is a major pathogenic feature of multiple complex diseases including common cardiometabolic diseases such as diet-induced obesity, insulin resistance, type-2 diabetes and atherosclerotic cardiovascular disease. Indeed, inappropriate and sustained tissue-specific inflammatory responses to diet and lifestyle factors in genetically prone individuals may contribute to cardiometabolic disease initiation and clinical complications. However, little is known for the modulation of the transcriptome by inflammatory stress in disease relevant human tissues. Here we applied RNA-Seq during low-dose experimental endotoxemia (LPS) in healthy human subjects to explore unbiased tissue-specific transcriptomic responses that may directly relate to complex disease responses. We utilized adipose and blood from three healthy individuals recruited to the Genetics of Evoked Responses to Niacin and Endotoxemia study who underwent a standardized inpatient endotoxemia protocol. Specifically, we examined LPS modulated tissue-specific gene expression, alternative splicing (AS), and expression of lincRNAs. Our comprehensive analysis revealed a remarkable number of novel tissue-specific transcriptomic signatures during activation of innate immunity in vivo. In total, we identified 587 and 664 differentially expressed genes at 5% FDR level in adipose and blood, respectively, following LPS administration. These genes were further replicated by an independent microarray experiment. Genes up-regulated are highly enriched in immune responses and infection/cardiovascular diseases, whereas those down-regulated are mainly related to oxidative process and mitochondria. We also detected 186 and 47 genes showing differential AS in adipose and blood. By comparing with deep RNA-Seq data of technical replicate samples, we confirmed 75% and 87% of the splicing events in adipose and blood, respectively. Additionally, we detected 14 and 17 differentially expressed lincRNAs in adipose and blood, including linc-TP53, which was differentially expressed in both tissues. At present, we are experimentally validating our findings by qPRC experiments and replicating in obese vs lean individuals. In summary, we have identified multiple levels of tissue-specific transcriptome modulation during induced inflammatory stress, which provide insight into genomic regulation of complex diseases that cannot be detected by analysis of DNA variations alone.

1643T

Genome-wide association study for atrial fibrillation in the Japanese population. *K. Ozaki[1], Y. Onouchi[1], N. Kamatani[2], T. Tsunoda[3], M. Kubo[4], Y. Nakamura[5], T. Tanaka[1].* 1) Laboratory for Cardiovascular diseases, RIKEN, Ctr Genomic Med, Yokohama, Japan; 2) Laboratory for Statistical Analysis, RIKEN, Ctr Genomic Med, Yokohama, Japan; 3) Laboratory for Medical Informatics, RIKEN, Ctr Genomic Med, Yokohama, Japan; 4) Laboratory for Genotyping Development, RIKEN, Ctr Genomic Med, Yokohama, Japan; 5) Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, University of Tokyo, Tokyo, Japan.

Atrial fibrillation (AF) is the most common arrhythmia in developing countries including Japan associated with increased morbidity and mortality. AF is an abnormal heart rhythm characterized by rapid and irregular activation of the atria. Through the Genome wide association study (GWAS) from European descent, genetic variants that confer susceptibility to AF have been indicated to be present on several chromosomal loci so far. These studies, however, were conducted in individuals from only European decent. To identify additional susceptibility loci for AF and to increase the understanding of the complex genetic effect on AF, further GWAS of Asian and other ethnic populations are required. To identify novel AF susceptible loci, we conducted a GWAS for AF with Japanese population. The GWAS was genotyped by the Illumina Human610-Quad and Illumina Human Hap550v3 BeadChip platform. We applied stringent quality-control criteria and tested 843 cases and 3,350 controls for approximately 430,000 autosomal SNPs with minor allele frequency of > 0.01. We also assessed population stratification in our study population by comparison to HapMap samples using principal component analyses, and found that these subjects did not show any sign of population stratification. The inflation of test statistics, $\lambda_{\text{genomic control}}$ (λ_{gc}) was 1.03. In this GWAS, the locus corresponding to the SNPs close to *PITX2* on chromosome 4q25, identified previously for AF susceptibility in European desent, satisfied a genome wide significance threshold of $P < 1 \times 10^{-7}$ in this GWAS stage. We are currently conducting the validation analysis with additional 1,646 cases and 17,190 controls for high-ranking SNPs of this GWAS.

1644T

Cardiovascular genetics wooooooo. *D. Seo, Z. Liu, N. Sikka, P. Goldschmidt-Clermont, M.A. Pericak-Vance, G.W. Beecham.* University of Miami, Miami, Fl.

Coronary heart disease is the leading cause of death in the United States, with atherosclerosis of the coronary arteries (CAD) being one of the major underlying etiologies. Patients with atherosclerosis are more likely to suffer a myocardial infarction or sudden cardiac death. Though the role of genetics on CAD has been studied in European ancestry populations, its role in Hispanic populations is less clear. To further determine the genetic etiology of CAD, we are performing a genome-wide association study of CAD using a diverse set of cardiac catheterization patients. This study utilizes the Miami Cardiovascular Registry of the University of Miami, Miller School of Medicine. A diverse set of patients were ascertained through local cardiac catheterization labs; the dataset consists of over 65% Hispanic and 15% African American. All 2,000 patients were phenotyped using a research grade read of the coronary angiograms, describing the amount of atherosclerosis in each of the 16 major coronary artery branches. Genotype data were generated using the Affymetrix SNP array 6.0 platform for 900,000 SNPs; an additional dataset using the same samples but genotyped on Illumina's “exome chip” is currently being generated. Extensive quality controls tests were performed to ensure the integrity of the data, including sample filters (gender consistencies, efficiency thresholds, etc), SNP filters (Hardy-Weinberg equilibrium, genotyping efficiency, MAF, etc), and EIGENSTRAT methods to correct for population substructure. We are first testing the dataset for ethnic differences in the architecture of disease; that is: do different populations have different coronary branches affected at different levels? Additionally, we are testing for association between disease and genetics using regression-based methods, including covariates for smoking, cholesterol, age, sex, and vectors describing the population substructure of the samples. We are testing for association both between overall disease burden, and with disease in specific major branches; results will be presented. This study will yield important results, both about the phenotypic differences and similarities between populations, and about underlying genetic etiology of CAD.

1645T

Functional Relationship of the COL4A1/COL4A2 Locus on Chromosome 13q34 to Coronary Artery Disease (CAD). *A. Turner, P. Lau, S. Soubeyrand, O. Jarinova, R. McPherson.* University of Ottawa Heart Institute, Ottawa, Ontario, Canada.

The COL4A1 and COL4A2 genes on chromosome 13 have been identified as new loci associated with CAD (p<3×10-8) from the recently published CARDIoGRAM meta analysis (Nature Genetics, 2011). The goal of the current project is to elucidate how variants in the COL4A1 and COL4A2 genes associated with CAD functionally and mechanistically contribute to the CAD phenotype. Type IV collagen triple helices constitute the major structural component of basement membranes, consisting primarily of 2 COL4A1 chains arranged with 1 COL4A2 chain. COL4A1 and COL4A2 also have important functional roles in angiogenesis, and mutations are associated with diverse vascular abnormalities. A follow-up study to CARDIoGRAM with more cases and controls has identified an intronic SNP in COL4A2, rs9515203, as having the greatest CAD association. Bioinformatics analysis using data from the ENCODE project reveals high degrees of histone 3 lysine 27 acetylation in the region near this rs9515203 SNP, a mark often located near active regulatory elements. We therefore hypothesized that the intronic region near rs9515203 behaves as either an enhancer or a repressor. Furthermore, we hypothesized that the rs9515203 SNP, or a SNP linked with rs9515203, may disrupt the function of this regulatory element. ChIP-Seq data from the ENCODE project shows binding of transcription factors such as STAT3 and USF1 to this intronic region. 2 kb of this intronic sequence (containing rs9515203) was cloned into the pGL3-Promoter vector (Promega), in which insertion of functional enhancers leads to upregulation of luciferase expression in vitro. Luciferase assays in HT-1080 cells reveal this intronic sequence acts as an enhancer due to its insertion upregulating pGL3-Promoter activity over threefold relative to controls (p<0.0001, n=3 independent experiments). Currently, using deletion constructs we have narrowed down the functional sequence of this enhancer. Future work entails verifying what protein factors bind this enhancing region and SNP effects using techniques to investigate DNA-protein interactions (ie. DNA pull-down assays, electrophoretic mobility shift assay (EMSA)). Chromosome conformation capture will also be conducted to evaluate whether this enhancer acts locally or distally. These findings are important because misregulation of COL4A1 and COL4A2 could have important consequences relevant to CAD, including effects on basement membrane integrity and angiogenesis.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1646T

***De novo* Copy Number Variants in Non-syndromic Left Ventricular Outflow Tract Obstruction.** *L.A. Umana[1], N.A. Hanchard[1], G.A. Zender[2], S.M. Fitzgerald-Butt[2], G. Zapata[1], P. Hernandez[1], M.S. Azamian[1], D.J. Penny[3,4], J.W. Belmont[1,4], K.L. McBride[2].* 1) Department of Human and Molecular Genetics, Baylor College of Medicine, Houston, TX; 2) Center for Molecular and Human Genetics, Nationwide Children's Hospital, Department of Pediatrics, Ohio state University, Columbus, OH; 3) Department of Pediatric Cardiology, Texas Children's Hospital, Houston, TX; 4) Department of Pediatrics, Baylor College of Medicine, Houston, TX.

Background: Congenital cardiovascular malformations (CVMs) occur in 5-8/1000 live births. Defects leading to Left Ventricular Outflow Tract Obstruction (LVOTO), most commonly represented by Hypoplastic Left Heart Syndrome (HLHS), Aortic Valve Stenosis (AVS) and Coarctation of the Aorta (CoA), constitute 15–20% of medically significant CVMs, and contribute significantly to CVM-associated morbidity and mortality. Non-syndromic LVOTO is multifactorial in origin, but familial clustering of cases and an increased risk of LVOTO in first degree relatives suggest a significant genetic contribution. We evaluated the incidence of large (> 150 Kb) de novo chromosome copy number variants (CNVs) among a cohort of affected family trios.

Methods and results: Family members were genotyped with ~732,000 SNPs genome-wide using the Illumina OmniExpress SNP microarray chip. CNVpartition and PennCNV software were used to detect copy number variants larger than 150 kb in the affected probands that were not seen in either parent. Called CNVs were confirmed by visual inspection of the logR ratio and B-allele frequency tracks. A total of 16 large de novo autosomal events were observed among the 377 affected probands (4.2%). Observed events ranged in size from 152 Kb to 50.8 Mb and included 10 losses, 2 gains, 2 unbalanced translocations, 1 instance of copy neutral absence of heterozygosity of chromosome 16 and 1 instance of extensive AOH (459 Mb total) consistent with a consanguineous mating. An apparent recurrent deletion of chromosome 14q32.33 was observed in two patients, as were two atypical deletions of chromosome 22q11.2 and two cases with terminal deletions of 11q (Jacobsen syndrome). In 11 of the 16 cases (69%) the cardiac lesion was complex, involving a typical LVOTO lesion in combination with additional cardiac findings.**Conclusions:** These findings suggest that large chromosomal events may contribute to even apparently non-syndromic LVOTO CVMs. In such patients, complex cardiac lesions in particular, may warrant extended cytogenetic work-up. In addition, copy number analysis can identify CVM patients at high risk for non-cardiac complications even if they are not evident in the newborn period.

1647T

Genetic association of CVD-related loci with lipid traits in the Heart Strategies Concentrating on Risk Evaluation (Heart SCORE) study. *XB. WANG[1], S. Reis[2], M.M. Barmada[1], I. Halder[2], F.Y. Demirci[1], M.I. Kamboh[1].* 1) HUMAN GENETICS, University of Pittsburgh, PITTSBURGH, PA., US; 2) Cardiovascular Institute, University of Pittsburgh, Pittsburgh, PA, US.

Cardiovascular disease (CVD) is the leading cause of death worldwide. Plasma levels of total cholesterol (TC), high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C) and triglycerides (TG) are significant risk factor for CVD and about 50% of the variation in these four quantitative traits is under genetic control. A recent meta-analysis of 46 published lipid genome-wide association studies has identified 95 loci that explain ~25–30% of the genetic variance for each of the four traits, indicating that additional genetic factors remain to be identified. In this study, we used a novel Illumina CVD BeadChip that contains ~49,000 SNPs in approximately 2,100 candidate genes selected based on their potential role in CVD to test for association with four lipid traits in 464 African American (AA) and 770 White subjects as part of the Heart SCORE study. Forty-nine of the reported 95 genome-wide significant genes were represented on this chip. We observed significant associations (p<0.05) with SNPs in 39 of the 49 genes with strongest associations between HDL-C and CETP/rs7499892 (p=2.56E-10), LDL-C and APOE/rs7412 (p=6.59E-05), TC and ABCA1/rs10991405 (p=9.69E-05), TG and APOB/rs2678379 (p=5.21E-05). In addition to the replication of known loci, we also observed novel suggestive associations between HDL-C and MTNR1A (p=2.62-05), HDL-C and NR1I2 (p=1.56E-05), TG and ADD3 (p=1.67E-05), TC and OSBP2 (p=1.1E-05), HDL-C and GHRL (p=5.24E -06), HDL-C and F2RL1 (p=4.52-05). We are in the process of confirming these novel associations in an independent sample. In conclusion, we have replicated many of the known lipid loci and identified potentially new genes for the four lipid traits using a novel CVD chip.

1648T

Genetic analysis of vascular disease in the Diabetes Heart Study. *A.J. Cox[1, 2], M.C-Y. Ng[2], J. Xu[2], C.D. Langefeld[3], J.J. Carr[4], B.I. Freedman[5], D.W. Bowden[1, 2].* 1) Department of Biochemistry, Wake Forest School of Medicine, Winston Salem, NC; 2) Diabetes Research Center, Wake Forest School of Medicine, Winston Salem, NC; 3) Department of Biostatistical Sciences, Wake Forest School of Medicine, Winston Salem, NC; 4) Department of Radiological Sciences, Wake Forest School of Medicine, Winston Salem, NC; 5) Department of Internal Medicine - Nephrology, Wake Forest School of Medicine, Winston Salem, NC.

Understanding genetic risk factors for cardiovascular disease (CVD) in individuals with type 2 diabetes (T2D) may identify biological pathways contributing to the increased risk for macrovascular complications and CVD mortality in T2D. The aim of this study was to examine genetic association signals for vascular disease and associated risk factors in a sample enriched for T2D. A genome-wide association study (GWAS) for vascular disease was performed in 1180 European Americans from 474 families from the Diabetes Heart Study (DHS). Phenotypes of interest included history of CVD (prior event or intervention), coronary artery calcified plaque (CAC) and other known CVD risk factors. Genotyping was completed using the Affymetrix 5.0 Array. Poorly performing SNPs (CR<0.95; Hardy-Weinberg p-value<1×10^{-6}; MAF<0.01) were excluded from analyses. Association analyses were completed using variance components methods implemented in SOLAR v6.3.4. Among the DHS cohort, 84% were T2D affected with disease durations averaging 10.5±7.2 years. Regions on Chr6p13 (lead SNP: rs7341323; *KCNQ5/RIMS1*; p=1.9×10^{-5}) and Chr10p12 (rs10741097; *GPR158*; p=6.5×10^{-6}) were associated with history of CVD. Stratification by smoking status (n=694 current or prior smokers; n=483 non-smokers) demonstrated differential associations with history of CVD, revealing an additional region on Chr12q24 (rs10850155; *LHX5*; p=2.4×10^{-6}) for smokers; Chr6p13 was the most strongly associated region among non-smokers (rs7763575; *KCNQ5/RIMS1*; p=4.9×10^{-7}). Refinement of the CVD history definition to include CAC burden revealed additional associations with regions on Chr9p23 (rs10978059; *PTPRD*; p=5.7×10^{-6}) among smokers and on Chr13q31 (rs1331967; *SLITRK1*; p=1.2×10^{-7}) among non-smokers. For all other risk factors examined, the strongest phenotype association was seen for total Bilirubin concentrations (rs887829; *UGT1A8*; p=8.67×10^{-20}). SNPs from this region were nominally associated with all-cause mortality (p=0.0004–0.04), but not with history of CVD or CAC. This study identified genetic loci associated with vascular disease in a sample of European Americans enriched for T2D. Further fine mapping of these regions is required to more fully characterise and replicate these associations and to identify SNPs which are likely to have impacts on gene function/expression. The potential for a modifying effect of environmental factors, here smoking status, on genetic associations warrants further consideration.

1649T

SNPs associated with plasma triglyceride levels influence response during intensive cardiovascular risk reduction. *D.L. Ellsworth[1], A. Decewicz[1], M. Hicks[1], K.A. Mamula[1], A. Burke[2], M.J. Haberkorn[2], H.L. Patney[1], M.N. Vernalis[3].* 1) Integrative Cardiac Health Program, Windber Research Institute, Windber, PA; 2) Windber Medical Center, Windber, PA; 3) Integrative Cardiac Health Program, Walter Reed National Military Medical Center, Bethesda, MD.

Background: Triglycerides play a fundamental role in development and progression of atherosclerosis. Current guidelines advocate lifestyle change involving diet, physical activity, and weight control for management of hypertriglyceridemic patients as a first step. Recent genome-wide association studies (GWAS) identified single nucleotide polymorphisms (SNPs) associated with plasma triglyceride levels in the general population. We hypothesize that plasma triglycerides may be influenced by genetic composition in addition to lifestyle behaviors. Methods: We examined the influence of genetic variation on variability in triglyceride response in 178 participants who completed a prospective, non-randomized intervention designed to stabilize or reverse progression of cardiovascular disease (CVD) through dietary changes, exercise, and stress reduction. CVD risk factors were assessed at baseline, 12 weeks, and 52 weeks by standard methods. SNPs (n=19) associated with plasma triglycerides were genotyped by TaqMan® allelic discrimination assays. Results: Patients experienced significant improvement (P < 0.05) in CVD risk factors, including weight (-9%), blood pressure (-6%), total cholesterol (-7%), and triglycerides (-9%). Triglyceride response during the program differed significantly (P < 0.05) between genotypes for three SNPs (rs442177, rs3846662, and rs17145738) located close to the following genes: transcriptional activator AF4/FMR2 family member 1 (AFF1), 3-hydroxy-3-methylglutaryl-CoA reductase (HMGCR), which catalyzes the rate-limiting step in cholesterol synthesis, and MLX interacting protein-like (MLXIPL), which controls transcription of genes involved in glycolysis. Discussion: This is the first study to show that individual response to lifestyle modification for CVD risk reduction may be influenced by genetic composition. Genetic variation associated with CVD risk may provide a basis for personalized treatments to optimize cardiovascular health and requires further study.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1650T

Sex-specific association of cadherin with high density-lipoprotein cholesterol in the Long Life Study Family. *M.F. Feitosa[1], A.T. Kraja[1], J. Lee[2], K. Christensen[3], J. Wang[1], C. Kammerer[4], M.A. Province[1], I.B. Borecki[1].* 1) Division of Statistical Genomics, Department of Genetics, Washington University School of Medicine, MO; 2) Taub Institute, Columbia University, NY; 3) Institute of Public Health at the University of Southern Denmark, Denmark; 4) Department of Human Genetics, University of Pittsburgh Graduate School of Public Health, PA.

It is well known that high levels of high-density lipoprotein (HDL) cholesterol have protective effects on atherosclerosis and cardiovascular disease, with involvement in the prevention of low-density lipoprotein (LDL) oxidation. Studies have suggested sex-specific heritability of lipids. Also, it has been shown that elderly people, especially centenarians, possess high HDL cholesterol levels and low LDL cholesterol levels as compared to septuagenarians. To detect whether loci exhibit sex-specific associations influencing fasting levels of HDL cholesterol, we analyzed men and women separately in the Long Life Family Study (LLFS) which is designed to characterize exceptionally healthy elderly people. For subjects taking lowering-lipid medications, the HDL cholesterol levels were corrected by classes of lowering-lipid medication effects. Levels of HDL cholesterol were adjusted by field centers, age, and principal components using stepwise regression analysis and retaining terms significant at the 5% level. The residuals were calculated separately for men and women. We performed genome-wide association (GW) scans on HDL cholesterol using a mixed model approach, accounting for family structure with the kinship correlations. A total of 4,114 European-American subjects (480 families) was typed using ~2.3 million SNPs (Illumina Omni chip). We identified a strong association between *CDH12* (cadherin 12 type 2, 5p14) variants with HDL cholesterol ($p<5.0\times10^{-8}$) in women but not in men ($p>1.0\times10^{-1}$). Further, we tested if there was an interaction effect between sex and SNP by examining the statistical significance of the main effects and interaction term. Significant evidence was found for sex-specific association with HDL cholesterol (p=0.004, for sex*SNP term). There was no other association at GW significance level for men or women. The *CDH12* gene belongs to the cadherin superfamily that mediate calcium-dependent cell-cell adhesion and cadherins adhesion are regulated by the Wnt signaling pathway. The *CDH12* is expressed in the brain and may play a role in neuronal development. Modest associations of CDH12 variants with total cholesterol in women ($p>1.0\times10^{-5}$) and HDL cholesterol in sex-combined analysis ($p>1.0\times10^{-3}$) were previously found in a meta-GWA study. Our findings suggest that using more homogeneous subgroups in elderly people in LLFS data helped to identify the *CDH12* loci associated with HDL cholesterol in women.

1651T

A genome-wide association study identifies KNG1 as a genetic determinant of plasma factor XI level and activated partial thromboplastin time. *JM. Soria[1], M. Sabater-Lleal[2], A. Martinez-Perez[1], A. Buil[3], L. Folkersen[2], JC. Souto[4], M. Bruzelius[2], M. Borrell[4], J. Odeberg[2], S. Silveira[2], P. Eriksson[2], A. Almasy[5], A. Hamsten[2].* 1) Unit of Genomic of Complex Diseases, Hospital de Sant Pau, Barcelona, Barcelona, Spain; 2) Cardiovascular Genetics and Genomics Group, Atherosclerosis Research Unit, Department of Medicine, Karolinska Institutet, Karolinska University Hospital Solna, Stockholm; 3) Department of Genetic Medicine and Development, University of Geneva Medical School, Geneva, Switzerland; 4) Haemostasis and Thrombosis Unit, Department of Hematology, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain; 5) Department of Population Genetics. Southwest Foundation for Biomedical Research. San Antonio, TX, USA.

Objective: Elevated plasma levels of coagulation factor XI (FXI) are implicated in the pathogenesis of venous thromboembolism (VTE) and ischemic stroke, and polymorphisms in the F11 gene are associated both with risk of VTE and an elevated plasma FXI level. Methods and Results: Here, we report the first hypothesis-free genome-wide genetic analysis of plasma FXI levels. Two genome-wide significant loci were detected in the family-based Genetic Analysis of Idiopathic Thrombophilia 1 (GAIT-1) cohort: one located in the kininogen 1 gene (KNG1) (rs710446; P=7.98×10-10), and one located in the structural F11 gene (rs4241824; P =1.16×10-8). Both associations were replicated in a second population-based Swedish cohort. A significant effect on KNG1 mRNA expression was also seen for the two most robustly FXI-associated SNPs located in KNG1. Furthermore, both KNG1 SNPs were associated with activated partial thromboplastin time (aPTT), suggesting that FXI may be the main mechanistic pathway by which KNG1 and F11 influence aPTT and risk of thrombosis. Conclusions: These findings contribute to the emerging molecular basis of VTE, and more importantly, help understanding the biological regulation of a phenotype that has proved to have promising therapeutic properties in relation to thrombosis.

1652T

Gene Variants are Associated with Soft Lipid-rich Coronary Plaque in African Americans with Replication in European Americans. *L.R. Yanek, B.G. Kral, D. Vaidya, T.F. Moy, L.C. Becker, D.M. Becker.* Medicine, Johns Hopkins University, Baltimore, MD.

Introduction: Genome-wide association studies (GWAS) have identified significant single nucleotide polymorphisms (SNP) for coronary artery calcification (CAC) in European Americans (EA) but not in African Americans (AA). Further, no studies have examined soft lipid-rich plaque, the most important vascular substrate for acute coronary syndromes. New computed tomographic angiography (CTA) technology extends simple CAC measures to allow better discrimination of plaque characteristics and identifies soft lipid rich plaque as “low attenuation plaque” (LAP). We performed a GWAS on LAP volume in healthy AA and EA from families with early-onset acute coronary events. Methods: Traditional risk factors were assessed using standard methods. Coronary artery plaque was assessed using 256 multidetector dual-source CTA. LAP was defined as plaque with density <30 Hounsfield units and volume was quantified using an automated well-validated proprietary method. Genotyping was performed using the Illumina 1Mv1c human array. MACH was used to impute to 2.5 million SNPs in HapMap II. Race-specific mixed models adjusted for age, sex, and two principal components from EIGENSTRAT were run under an additive model to test each SNP-phenotype association. Results: Participants (350 EA, 198 AA) were 58% female, with mean age 51.5 ± 10.8 years. Two SNPs were significant at p<1E-5 in AA, and had nominal replication in EA, in the same direction (Table). Conclusions: This is the first GWAS to report any potentially important genetic signals for coronary plaque in AA, replicated with the same vector and nominal significance in EA, and also the first to report signals for soft lipid rich plaque, an important precursor for acute coronary syndromes.

SNP Chr	Gene	AA MAF (MA)	AA effect	AA p	EA MAF (MA)	EA effect	EA p
rs12204831 6		0.32 (T)	-1.22	3.4E-6	0.29 (T)	-0.486	0.022
rs10506016 12	STK38L	0.14 (G)	-1.36	5.9E-6	0.03 (G)	-1.37	0.008

1653T

Genetic variation for leukocyte telomere length predicts incident coronary heart disease The Framingham Heart Study. *S. Hwang[1], A. Aviv[2], D. Levy[1], M. Mangino[3], S. Hunt[4], A.L. Fitzpatrick[5], J.C. Bis[5], E.S. Smith[6].* 1) The National Heart, Lung, and Blood Institute, The Framingham Heart Study, Framingham, MA 01702; 2) The Center of Human Development and Aging, New Jersey Medical School, UMDNJ, Newark, NJ 07103; 3) Department of Twin Research and Genetic Epidemiology, King's College London, London, UK; 4) Cardiovascular Genetics Division, Department of Medicine, University of Utah, Salt Lake City, UT84101; 5) Departments of Epidemiology and Global Health, University of Washington, Seattle, WA 98195; 6) Genome Information Sciences at University of California San Diego, 3855 Health Sciences Drive, La Jolla, CA 92093.

Shorter leukocyte telomere length (LTL) correlates with higher cardiovascular disease risk. Whether genetic variation associated with LTL can be applied as a predictor of coronary heart disease (CHD) is unknown. We examined the association between LTL genetic risk score (GRS) derived from 6 SNPs with incident CHD in 4078 Framingham Heart Study participants who were free of CHD at baseline. Polymorphisms associated with LTL were identified through meta-analysis of genome-wide association in six cohorts with a total sample size of 9190 were used to construct a GRS for LTL. We applied proportional hazards regression to test for significant association of the GRS with CHD. Net reclassification index was calculated to characterize the incremental improvement of the GRS for the prediction of CHD above and beyond established CHD risk factors. A total of 324 participants (7.95%) developed CHD after an average 11.3 years of follow-up. Results of proportional hazards regression adjusted for baseline CHD risk factors showed an association of the GRS with CHD (hazards ration 1.1, 95% confidence interval 1.03–1.19, p=0.0068 for each unit of the GRS). Those with highest quartile of the GRS had a 45% increase in risk for CHD compared with those with the lowest quartile (95% CI 8% to 95%). Results of this community-based study suggest a potential role of a GRS for the prediction of CHD. These results warrant further investigation.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1654T

Genome-wide association analysis of plasma triglyceride in Korean population. *B. Gombojav[1], D. H. Lee[1], M. K. Lee[1], Y. M. Song[2], K. Lee[3], J. Sung[1].* 1) Complex Diseaseand Genome Epidemiology Branch, Department of Epidemiology and Instituteof Environmentand Health, School of Public Health, Seoul National; 2) Department of Family Medicine, Samsung Medical Center, and Center for Clinical Research, Samsung Biomedical Research Institute, Sung Kyun Kwan University School of Medicine, Seoul, Korea; 3) Department of Family Medicine, Busan Paik Hospital, Inje University College of Medicine, Korea.

Background: The metabolic syndrome and associated disorders such as insulin resistance and type 2 diabetes results from calorie excess, physical inactivity and underlying genetic susceptibility. Increased plasma triglycerides are key atherogenic lipid phenotype of the metabolic syndrome. Objective: To identify common genetic variants associated with plasma triglycerides, we did a genome-wide association study. Methods: A total of 537,159 markers were tested for their association with the plasma triglyceride in a GWAS. The GWAS sample consisted of 3,029 healthy individuals from 661 families of the Healthy Twin Study Korea. The most promising markers from the GWAS were subsequently tested in replication samples comprised of 8,834 individuals from the KARE study. Results: We replicate previously reported association of a synonymous SNP (rs651821) in APOA5 (P = 1.0 × 10-10) and identify four novel loci, which associated with plasma triglyceride level: 9q21.13 near GNA14 (rs6560604, P = 3.2 × 10-7); 7p21.3 near ICA1 (rs17144686, P = 3.3 × 10-7); 14q13.1 near AKAP6 (rs2300856, P < 0.0000053); 7p15.1 near CHN2 (rs3793269, P < 0.0000063). Among the four candidate loci, 9q21.13 has been reported to associate with plasma triglyceride and GNA14 controls the activation of phospholipase. Conclusion: We have identified genetic locus harboring common genetic variant that was associated with plasma triglycerides. Further studies are required to replicate these findings and to elucidate the biological mechanisms.

1655T

Genome-wide Association Study of Calcific Aortic Valve Disease. *S. Guauque-Olarte[1], M. Lamontagne[1], N. Gaudreault[1], P. Mathieu[1], P. Pibarot[1], Y. Bossé[1,2], D. Messika-Zeitoun[3,4].* 1) Centre de recherche Institut universitaire de cardiologie et de pneumologie de Québec, Laval University, Quebec, Canada; 2) Department of Molecular Medicine, Laval University, Quebec, Canada; 3) Cardiology Department, AP-HP, Bichat Hospital, Paris, France; 4) INSERM U698, University Paris 7, Paris, France.

Calcific aortic valve stenosis (AS) is a life-threatening disease that affects 2% of people older than 65 years. There is currently no medical treatment available. Only a few genetic studies were conducted to uncover the genetic architecture of AS. Some candidate genes have been identified but the causal genes remain to be identified. The objective of this study was to perform a genome-wide survey of susceptibility loci for AS. Cases from Quebec and France were collected independently to perform a genome-wide association study (GWAS) on AS among patients of European ancestry. The genotyping of both Quebec and French cohorts was carried out using the Illumina® HumanOmniExpress-12v1.0 BeadChip. After quality controls, the Quebec cohort consisted of 474 patients who underwent an aortic valve replacement, while the France cohort included 486 patients with at least mild AS confirmed using echocardiography. A control cohort of 3,294 individuals of European ancestry was taken from the Illumina® iControlDB. Two independent GWAS were performed. First using the cases from Quebec and 1,575 Illumina controls, and second using cases from France and 1,576 different Illumina controls. Imputation was performed with IMPUTE2 using SNPs in common between cases and controls. Nearly 6 million genotyped or imputed SNPs were analyzed in both GWAS. Single marker associations were tested using additive logistic regression models in PLINK v1.07. Eight principal components derived from EIGENSTRAT were used as covariates to adjust for population stratification. The genomic inflation factors were 1.0 for both GWAS. A meta-analysis was then performed to identify SNPs robustly associated with AS using the inverse variance meta-analysis implemented in PLINK. The Bonferroni correction was used to adjust for multiple testing. The best genetic association signals were observed on chromosomes 12 and 15. Replication of these susceptibility loci in additional populations is underway. Identifying genes associated with AS is important to find new molecular targets and therapeutic pathways for this frequent and life-threatening disease.

1656T

Genetic Variants Explaining Sodium Intake in a Population with Higher Sodium Intake Level: The Healthy Twin Study, Korea. *M. Kho[1], Y.-M. Song[2], K. Lee[3], JE. Lee[4], K. Kim[5], J. Sung[1].* 1) Department of Epidemiology, Institute of Health and Environment, School of Public Health, Seoul National University, Seoul, Korea; 2) Department of Family Medicine, Samsung Medical Center, Center for Clinical Research, Samsung Biomedical Research Institute, Sungkyunkwan University School of Medicine, Seoul, Korea; 3) Department of Family Medicine, Busan Paik Hospital, Inje University College of Medicine, Busan, Korea; 4) Department of Food and Nutrition, Sookmyung Women's University, Seoul, Republic of Korea; 5) Department of Statistics, Sookmyung Women's University, Seoul, Republic of Korea.

[BACKGROUND] Excess consumption of salt is an established risk factor of chronic diseases, notably hypertension as well as obesity and stomach cancer. In Asian countries including Korea, where rapid westernization of lifestyle is ongoing, traditional food has been reported to contain higher level of sodium, despite its potential beneficial effects. In our previous study, genetic factor was the most important determinant of sodium intake, with heritability of 0.32. Environmental factors including shared familial environments did not account significant proportion of salt intake. However, limited evidence has been accumulated on the specific variants explaining individual sodium intake. The aim of this study is to identify genetic variants associated with sodium intake through a genome-wide association study (GWAS). [METHOD] This study involved 2,209 healthy individuals from Healthy Twin Study cohort, Korea. We estimated 24 hour urinary sodium excretion from a Half-day urine (HU) sample. HU collection starts around 7 pm of the day before visit, after completely voiding when time record starts. Next morning, before health check, remaining urine was further voided and the time was recorded as final. We estimated 24 hour sodium intake using simple volume-time extrapolation with samples collected more than 8 hours. With samples less than 8 hours duration, we applied the Kawasaki Formula which is developed for estimating 24 hour sodium intake using spot urine. Total calorie was adjusted to correct sodium intake due to food volume. A total of 537,158 markers were tested for their association with the 24 hour sodium intake in GWAS. The association is conducted through Family-based Score Test for Association(FASTA) approach(genABEL). [RESULTS] Several candidates variants reached sufficient or probable genome-wide significance level; p16.2(p=2.35E-07), which harbors genes regulating voltage-gated Na+ channels(VGSCs) and Ca2+/calmodulin kinase in cardiac muscle; p15.33(p=8.78E-06) which harbors genes related to sodium reabsorption process in nephron; q11.21(p=8.30E-06) where genes related to potassium channel. [DISCUSSION & CONCLUSIONS] Although our findings await replications, we suggested a potential interesting genetic locus associated with individual preference to salt intake level. We expect the candidate genetic variants will provide references for future studies regarding sodium intake and its health effects.

1657T

Dietary intake of saturated fatty acids modulates the genetic effects on triglyceride concentrations in a Korean genome-wide study. *H. Lee[1], H. Jang[1], M. Go[2], J. Park[1], J. Choi[1], J. Park[1], Y. Ahn[2], B. Han[2], J. Lee[2], J. Song[1].* 1) Center for Biomedical Sciences, NIH Korea, Cheongwon-gun, Chungcheongbuk-do, South Korea; 2) Center for Genome Science, NIH Korea, Cheongwon-gun, Chungcheongbuk-do, South Korea.

Triglyceride is one of the key risk factor for the metabolic syndrome and the increased plasma triglyceride concentrations are influenced by dietary high intake of saturated fatty acid. To explore genetic loci interacting with dietary intake of saturated fatty acids participating in plasma triglyceride level, we analyzed the interaction between Korean genome-wide genotype data from the KARE project (the Affymetrix Genome-Wide Human SNP arrays 5.0) and dietary intake saturated fatty acids from Korean Genome Epidemiology Study (KoGES). We investigated 37 genome-wide significant signals interacting with dietary intake of saturated fatty acids explaining their potential involvement in triglyceride concentrations in Korean population (n = 7495). After the additional adjustment with smoking and alcohol drinking status, dietary intakes, and hyperlipidemic medication status, 1p13.2d (KCND3) and 6q25.1a (LRP11) loci appeared with the highest interaction effect signal (P for interaction < 3.47E-05) and 13q14.11d (EPSTI1), 13q21.1a (OLFM4), and 20p13e (RSPO4) loci ranked marginal interaction effect with dietary saturated fatty acid on plasma triglyceride. When the subjects were separated into two groups according to saturated fatty acid dietary intake level, high and low, KCND3 rs2165987 (1p13.2d) in the group with a high level of dietary saturated fatty acid showed the highest significant interaction with plasma triglyceride (P=0.0035). The triglyceride plasma concentration in the subjects with KCND3 rs2165987 (1p13.2d) genotypes TT (169.9 ±111.0 mg/dL) was higher than TC (152.7±83.7 mg/dL) which was higher than in CC carriers (150.5±85.8 mg/dL) in the subgroup exhibiting a higher dietary intake of saturated fatty acid. In the subgroup exhibiting a lower saturated fatty acid dietary intake, carriers of minor homozygous genotypes of LRP11 rs2342765 (6q25.1a), had lower plasma triglyceride concentrations than those who carry major homozygous genotypes (P= 0.0049. The effect of dietary intake of saturated fatty acid on plasma triglyceride might be varified by the susceptibility of KCND3 and LRP11 genotype.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1658T

A GWAS approach in search for modifiers of congenital heart disease in 22q11 microdeletion syndrome. *G.M. Repetto[1,2], C. Vial[1], S. McGhee[3], N.K. Henderson-MacLennan[4], M.L. Guzman[1], M. Palomares[5,6], G. Lay-Son[1,2], K. Espinoza[1], T. Guo[7], B. Morrow[7].* 1) Center for Human Genetics, Clin Alemana- Univ Desarrollo, Santiago, Chile; 2) Hospital Padre Hurtado, Santiago, Chile; 3) Stanford University, Stanford, CA; 4) UCLA, Los Angeles, CA; 5) Fundación Gantz, Santiago, Chile; 6) Hospital Dr Luis Calvo Mackenna, Santiago, Chile; 7) Albert Einstein College of Medicine, NY.

Chromosome 22q11 microdeletion syndrome (del22q11) has an estimated frequency of 1/4000 live births. Most patients share a common 3Mb deletion, but despite molecular similarity, the clinical phenotype shows marked variable expressivity. Approximately 50-70% of patients have congenital heart disease, predominantly involving the cardiac outflow tract. The cause of the incomplete penetrance of this manifestation of the syndrome is unknown. We performed a genome wide association study (GWAS) to search for genetic modifiers of the cardiac phenotype in Chilean patients with del22q11. DNA samples from 112 Chilean patients with del22q11 demonstrated by FISH with TUPLE1 probe (Abbott®), and know cardiac anatomy by ultrasound were analyzed with Affymetrix ® v. 6.0 SNP array. We compared cases with any cardiac or great vessel anomaly (n=59) to controls with normal cardiac or great vessel anatomy by echocardiogram (n=53). The most common cardiac defects in this series were ventricular septal defects and tetralogy of Fallot. After filtering and quality control, genotypes for 728,000 SNPs were analyzed with PLINK v1.07, using Fisher exact test for association. A Bonferroni corrected p value of 1 × 10-7 was considered as evidence of association. Five SNPs in a 100 kb region on chromosome 6q14 showed p values between 1 × 10-6 to 2 × 10-5. Although these results do not fulfill criteria for statistical significance, they point to potential regions of interest. A larger sample size in the GWAS, and sequencing of the region may help identify modifier genes for the cardiac phenotype in del22q11 syndrome. Funded by Fondecyt-Chile Grant #1100131.

1659T

Novel genes related to Hypertension: a gene-centric meta-analysis with over 127,000 individuals. *V. Tragante do O[1,4], S.K. Ganesh[2], W. Guo[3], Y. Guo[5], E.N. Smith[8], M.B. Lanktree[6], T. Johnson[7], B. Almoguera Castillo[5], J. Barnard[9], J. Baumert[10], A.G. Uitterlinden[11], P. van der Harst[12], Y.T. van der Schouw[13], N.J. Samani[14,15], P.B. Munroe[7], P.I.W. de Bakker[4,13,16], X. Zhu[3], D. Levy[17], B. Keating[5], F.W. Asselbergs[1,13], The IBC BP Consortium.* 1) Department of Cardiology, Division Heart and Lungs, University Medical Center Utrecht, Utrecht, The Netherlands; 2) Division of Cardiovascular Medicine, University of Michigan Health System, Ann Arbor, MI, USA; 3) Department of Epidemiology and Biostatistics, School of Medicine, Case Western Reserve University, Cleveland, OH, USA; 4) Complex Genetics Section, Department of Medical Genetics, University Medical Center Utrecht, Utrecht, the Netherlands; 5) Center for Applied Genomics, Abramson Research Center, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania 19104, USA; 6) Departments of Medicine and Biochemistry, Schulich School of Medicine and Dentistry, University of Western Ontario, London, Ontario, Canada; 7) Clinical Pharmacology and Barts and The London Genome Centre, William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, Charterhouse Square, London EC1M 6BQ, UK; 8) Department of Pediatrics and Rady's Children's Hospital, University of California at San Diego, School of Medicine, La Jolla, California 92093, USA; 9) Department of Quantitative Health Sciences, Lerner Research Institute, Cleveland Clinic, Cleveland, OH; 10) Institute of Epidemiology II, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany; 11) Depts. of Epidemiology and Internal Medicine, Erasmus Medical Center, Rotterdam, the Netherlands; 12) Department of Cardiology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands; 13) Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, the Netherlands; 14) Department of Cardiovascular Sciences, University of Leicester, Leicester, UK LE3 9QP; 15) Leicester NIHR Biomedical Research Unit in Cardiovascular Disease, Glenfield Hospital, Leicester, UK, LE3 9QP; 16) Division of Genetics, Brigham and Women's Hospital, Harvard Medical School, Boston, MA and Broad Institute of MIT and Harvard, Cambridge, MA, USA; 17) Center for Population Studies, National Heart, Lung, and Blood Institute, Framingham, MA, USA.

Blood pressure (BP) is a heritable determinant of risk for cardiovascular disease. To investigate genetic associations with systolic BP (SBP), diastolic BP (DBP), mean arterial pressure (MAP), and pulse pressure (PP), we genotyped ~50,000 single nucleotide polymorphisms (SNPs) that capture variation in ~2,000 candidate genes for cardiovascular phenotypes in 61,619 individuals of European ancestry from cohort studies in the US and Europe. We identified novel associations between gene HRH1 and SBP and between gene MDM4 and DBP and then replicated these associations in an additional set of 65,866 independent individuals of European ancestry. We also identified associations between a SNP from SOX6 gene and SBP, previously reported to be associated with MAP. We also confirmed nine previously known loci associated with SBP, DBP, MAP, or PP (ADRB1, ATP2B1, SH2B3/ATXN2, CSK, CYP17A1, FURIN, HFE, LSP1, MTHFR) at array-wide significance (P value < 2.4 × 10-6). Modest sex interactions were identified in the MDM4 locus, and findings from eQTL analysis showed associations of SNPs in the MDM4 region with MDM4 expression. Our findings extend our understanding of genes involved in BP regulation, some of which may eventually provide new targets for therapeutic intervention.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1660T

Additional variance of serum lipid levels explained by incorporating less significant genetic variants and allelic heterogeneity. *R. Deka[1], G. Zhang[2], R. Karns[1], G. Sun[1], S.R. Indugula[1], H. Cheng[1], D. Havas-Augustin[3], N. Novokmet[3], Z. Durakovic[3], S. Missoni[3], R. Chakraborty[4], P. Rudan[3].* 1) Dept Environmental Hlth, Univ Cincinnati Med Ctr, Cincinnati, OH; 2) Human Genetics Division, Cincinnati Children's Hospital Med Ctr, Cincinnati, OH; 3) Institute for Anthropological Research, Zagreb, Croatia; 4) Institute of Applied Genomics, Univ North Texas Health Science Ctr, Forth Worth, TX.

Serum concentrations of total cholesterol, low-density lipoprotein cholesterol, high-density lipoprotein cholesterol, and triglycerides are important heritable risk factors for cardiovascular disease. Large-scale genome-wide association studies (GWAS) have identified many genetic variants robustly associated with serum lipid levels; however, these variants explain only a small proportion of the overall phenotypic variance, leading to the issues of `missing' heritability. In this study we examined two possible sources of missing heritability: first, variants with smaller effects whose associations with the trait failed to reach genome-wide significance and second, allelic heterogeneity due to the effects of multiple variants at a single significant locus. For this purpose, we developed an analytical approach that accounts for local linkage disequilibrium patterns and dissects the allelic heterogeneity of significant loci from summary-level statistics from a meta-analysis of GWAS. We applied this method to the summary results reported by Teslovich et al, *Nature* (2010) and dissected heterogeneous allelic effects of lipids-associated loci identified at various significance levels. In a sample of 1,304 individuals from an island population of the Adriatic coast of Croatia, we further examined the extent of phenotypic variance of serum lipid levels explained by incorporating the effects of both the leading and additional variants of lipids-associated loci. Our results indicate that approximately half of the lipid loci that achieved stringent genome-wide significance (p-value $< 5\times10^{-8}$) showed significant evidence of allelic heterogeneity. In addition, including less significant loci (p-value $< 5\times10^{-5}$) and accounting for effects of allelic heterogeneity substantially improved the variance explained of serum lipid levels in our samples (>50% higher than the variance explained by using only leading SNPs with genome-wide significance). These observations suggest that allelic heterogeneity is common phenomenon in complex traits and an appreciable fraction of SNPs not meeting genome-wide significance might have small but genuine effects. Supported by NIH grants R01 DK069845, P30 ES006096, T32 ES010957.

1661T

Admixture mapping of coronary artery calcified plaque in African Americans with type 2 diabetes. *J. Divers[1], N.D. Palmer[2], L. lu[1], T.C. Register[3], J.J. Carr[4], P.J. Hicks[2], R.C. Hightower[4], S.C. Smith[2], J. Xu[2], A.J. Cox[2], K.A. Hruska[8], D.W. Bowden[2], C.E. Lewis[9], G. Heiss[10], M.A. Province[15], I.B. Borecki[15], K.F. Kerr[11], Y.D. Chen[12], W. Palmas[13], J.I. Rotter[12], C.L. Wassel[14], A. Bertoni[5], D. Herrington[7], L.E. Wagenknecht[5], C.D. Langefeld[1], B.I. Freedman[6].* 1) Department of Biostatistical Sciences, Wake Forest School of Medicine, Winston-Salem, NC; 2) Department of Biochemistry and Centers for Diabetes Research and Human Genomics,Wake Forest School of Medicine, Winston-Salem, NC; 3) Department of Pathology, Wake Forest School of Medicine, Winston-Salem, NC; 4) Division of Radiologic Sciences, Wake Forest School of Medicine, Winston-Salem, NC; 5) Department of Epidemiology and Prevention, Wake Forest School of Medicine, Winston-Salem, NC; 6) Department of Internal Medicine, Section of Nephrology, Wake Forest School of Medicine, Winston-Salem, NC; 7) Department of Internal Medicine, Section of Cardiology, Wake Forest School of Medicine, Winston-Salem, NC; 8) Pediatrics and Internal Medicine -Nephrology, Washington University School of Medicine, St. Louis, MO; 9) Division of Preventive Medicine, University of Alabama Birmingham, Birmingham, AL; 10) Department of Epidemiology, University of North Carolina, Chapel Hill, NC; 11) Department of Biostatistics, University of Washington, Seattle, WA; 12) Medical Genetics Institute, Cedars Sinai Medical Center, Los Angeles, CA; 13) Department of Medicine, Columbia University, New York, NY; 14) Department of Family and Preventive Medicine, University of California, San Diego, CA; 15) Division of Statistical Genomics, Department of Genetics, Washington University School of Medicine, St. Louis, MO.

Background: The presence and severity of coronary artery calcified plaque (CAC) differs markedly between individuals of African and European descent, suggesting that admixture mapping (AM) may be informative for identifying genetic variants associated with subclinical cardiovascular disease (CVD). Methods: AM of CAC was performed in 1,040 unrelated African Americans (AAs) with type 2 diabetes from the African American-Diabetes Heart Study (AA-DHS), Multi-Ethnic Study of Atherosclerosis (MESA), and Family Heart Study (FamHS) using the Illumina custom ancestry informative marker (AIM) panel. All cohorts obtained computed tomography scanning of the coronary arteries using identical protocols. For each AIM, the probability of inheriting 0, 1, and 2 copies of a European-derived allele was determined. Association testing was performed between each AIM using these probabilities and CAC, accounting for global ancestry, age, gender and study. Results: Markers on 1p32.3 in the GLIS1 gene (rs6663966, LOD= 3.7), 1q32.1 near the CHIT1 gene (rs7530895, LOD=3.1), 4q21.2 near PRKG2 (rs1212373, LOD=3.0) and 11q25 in the OPCML gene (rs6590705, LOD=3.4) had statistically significant LOD scores, while markers on 8q22.2 (rs6994682, LOD=2.7), 9p21.2 (rs439314, LOD=2.7), and 13p32.1 (rs7492028, LOD=2.8) manifested suggestive evidence of linkage. These regions were uniformly characterized by higher levels of European ancestry associating with higher level or odds of CAC. Findings were replicated in 1,350 AAs without diabetes and 2,497 EAs with diabetes from MESA and 1,181 EAs with diabetes from the DHS. Conclusions: Fine mapping these regions will likely identify novel genetic variants that contribute to CAC and clarify racial differences in susceptibility to subclinical CVD.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1662T

Genome-wide association studies of NMR-based lipoprotein subclasses in African American populations elucidate subfractions affected by SNP variation. *Y. Huang[1], WM. Chen[1], K.L. Keene[1], F. Chen[1], U. Nayak[1], X. Hou[1], I.J. Spruill[3], K.J. Hunt[3], J.K. Fernandes[3], K.H. Lok[2], J. Divers[4], J.C. Mychaleckyj[1], D.L. Kamen[3], W. Post[5], S.S. Rich[1], J.I. Rotter[6], G.S. Gilkeson[3], W.T. Garvey[2], M. Cushman[7], M.M. Sale[1].* 1) University of Virginia, Charlottesville, VA; 2) University of Alabama at Birmingham, Birmingham, AL; 3) Medical University of South Carolina, Charleston, SC; 4) Wake Forest University School of Medicine, Winston-Salem, NC; 5) Johns Hopkins University, Baltimore, MD; 6) Cedars Sinai, Los Angeles, CA; 7) University of Vermont, Burlington, VT.

We conducted a meta-analysis of GWAS scans of NMR measured lipoprotein subclasses in 5,438 African Americans (AA), including 1,491 family-based Gullah AA from the Sea Island Genetic NETwork (SIGNET), 2,370 unrelated AA from the Reasons for Geographic And Racial Differences in Stroke study (REGARDS) and 1,577 unrelated AA from the Multi-Ethnic Study of Atherosclerosis (MESA). We also attempted replication of significant associations discovered in AA in 4,550 non-AA MESA participants. Hapmap 3 imputed autosomal SNPs (r^2>0.3) plus genotyped autosomal SNPs (1,438,463 SNPs in total)were used for meta-analysis of 23 heritable inverse-normal-transformed quantitative lipoprotein traits. A regression-based QTL model for dosage data, adjusting for sex, age, diabetes status, lipid medication status, ancestry principal components and family structure, was applied for every SNP. We identified three putative novel loci in AA. A variant in *CD36* (7q11.2) was significantly associated with NMR lipoprotein traits small VLDL particle concentration (VSP, $P=6.91\times10^{-11}$) and VLDL & chylomicron particle concentration (VLDLCP, $P=1.10\times10^{-8}$), while its association with the total concentration of VLDL & chylomicron triglyceride (NVCTG) was marginal ($P=6.36\times10^{-3}$). Two variants in *DOCK6* (19p13.2) were significantly associated with HDL particle concentration (HDLP, $P=3.34\times10^{-9}$ and $P=3.57\times10^{-9}$) and a variant in *ATRN* (20p13) was associated with large HDL particle concentration (HLP, $P=1.53\times10^{-9}$), while none of them were significantly associated with conventionally-measured total HDL cholesterol (HDL, $P=5.29\times10^{-5}$, 4.37×10^{-4} and 3.86×10^{-4} respectively). No variants in these three genes were replicated in our non-AA samples with European, Asian and Hispanic ancestry. Associations were also found between well-known lipid loci variants and novel lipid subclass traits, such as rs247617 in CETP, which was associated positively with HLP ($P=1.16\times10^{-24}$), HDL mean particle size (HZ, $P=7.63\times10^{-18}$), large LDL particle concentration(LLP, $P=2.05\times10^{-8}$), LDL mean particle size (LZ, $P=4.72\times10^{-13}$), and negatively with small LDL particle concentration (LSP, $P=9.64\times10^{-13}$), besides its previously known association with total HDL concentration. Our study suggests that the additional information provided by NMR lipoprotein subclasses may be able to identify novel lipid loci in African American population, and may also help elucidate the specific lipoprotein subclasses that are impacted by known variants.

1663T

Redefining Fibromuscular Dysplasia of the Arteries as a TGF-β Pathway Disorder. *R. Morissette[1], S. Ganesh[2], B. Griswold[1], L. Sloper[1], N. McDonnell[1].* 1) National Institute on Aging, National Institutes of Health, Baltimore, MD; 2) Division of Cardiovascular Medicine, University of Michigan Health System, Ann Arbor, MI.

Fibromuscular dysplasia (FMD) is a disorder involving stenosis, aneurysms, and dissections of predominantly renal and carotid arteries. Other medium sized arteries of the abdomen, intracranial vessels, and coronary arteries may also be involved. The histopathology observed is aberrant extracellular matrix deposition in the arterial wall. The disorder preferentially becomes evident in women in their fourth through sixth decade. Autosomal dominant inheritance with reduced penetrance has been suggested. Although FMD-type histopathology can be seen in a number of genetic disorders, no underlying gene mutation has been identified in the vast majority of patients who have a diagnosis of primary FMD. Clinical examination of FMD subjects from NIA protocol 2003-086 (a screening protocol for hereditary disorders of connective tissue) revealed a high prevalence of joint laxity (as defined by a Beighton score > 5), scoliosis, low bone density, and other stigmata of hereditary connective tissue dysplasias such as pes planus, pectus deformities, and dolicocephaly. Family histories were notable for increased incidence of thoracic and abdominal aneurysms in first-degree male relatives. These clinical findings as well as the extracellular matrix pathology led to the hypothesis that derangement of the TGF-β pathway may play a role in the pathogenesis of FMD. To study the pathway, skin biopsy specimens were obtained from 18 FMD patients and fibroblast lines were established. Secretion of TGF-β1 and TGF-β2 from the FMD fibroblasts was found to be elevated compared to matched controls (p=0.0008 and p<0.0001, respectively) by ELISA. Interestingly, secreted TGF-β3 was not significantly different (p>0.05), suggesting biomarker specificity in FMD pathogenesis. FMD patients also had elevated circulating TGF-β2 in plasma relative to matched controls (p=0.005). Studies are underway to explore the role of TGF-β signaling and other candidate pathways in FMD, as well as formulation and investigation of alternative hypotheses to explain the observed elevation of circulating and secreted TGF-β1 and TGF-β2.

1664T

Integrated microRNA and mRNA profiling of the mouse ventricles during development of severe hypertrophic cardiomyopathy and heart failure. *R.D. Bagnall[1], T. Tsoutsman[1,2], R.E. Shephard[1,2], W. Ritchie[3], C. Semsarian[1,2,4].* 1) Agnes Ginges Center for Molecular Cardiology, Centenary Institute, Sydney, NSW, Australia; 2) Sydney Medical School, University of Sydney, NSW, Australia; 3) Department of Bioinformatics, Centenary Institute, Sydney, NSW, Australia; 4) Department of Cardiology, Royal Prince Alfred Hospital, Sydney, NSW, Australia.

Hypertrophic cardiomyopathy (HCM) is a primary disorder of the myocardium characterized by left ventricular hypertrophy, myocyte disarray and interstitial myocardial fibrosis. HCM is caused by autosomal dominant mutations primarily affecting genes encoding proteins of the sarcomere, however, it is less clear how these mutations alter intracellular signaling, leading to cardiac remodeling and hypertrophy. MicroRNAs (miRNAs) are short non-coding RNAs that regulate post-transcriptional gene expression during development and disease. We have used TaqMan Low Density Arrays to measure the expression levels of 335 murine miRNAs in ventricular tissue and determine the miRNA signatures of early- and end-stage HCM in a severe, transgenic mouse model of the disease. Seven miRNAs were dysregulated at an early stage of HCM development. Time-course analysis revealed that decreased expression of miR-1 commences at a pre-disease stage, with a consequent upregulation of target genes causal of cardiac hypertrophy and fibrosis, and thus represents an early disease change. At end-stage HCM, 21 miRNAs were dysregulated to form a cardiac stress signature resembling that of other forms of cardiac hypertrophy, suggesting common responses. Analysis of the mRNA transcriptome using Affymetrix GeneChip Arrays revealed that miRNAs potentially target 19.7% upregulated and 6.7% downregulated mRNAs at end-stage HCM, and regulate mRNAs associated with cardiac hypertrophy, calcium signaling, fibrosis and the Tgf-β signaling pathway. Collectively, these results highlight the regulatory roles of miRNAs in the development and progression of HCM, and shed light on critical miRNA regulated gene networks involved in disease pathogenesis. Furthermore, strategies to maintain miR-1 levels may represent a therapeutic opportunity in HCM.

1665T

Novel transcripts and pathways identified in blood one week following implant of continuous-flow left ventricular assist device (CF-LVAD). *J.L. Hall[1], W. Guan[2], A. Mitchell[1], R. Staggs[1], S. Grindle[1], N. Adhikari[1], S. Hozayen[1], P. Eckman[1].* 1) Dept Med, Univ Minnesota, Minneapolis, MN; 2) Dept Biostatistics, School of Public Health, Univ Minnesota, Minneapolis MN.

Introduction: Bleeding, thrombosis and infection are common adverse events following placement of continuous-flow left ventricular assist devices (CF-LVADs). We performed RNA sequencing to identify novel genes and pathways that may be involved. Methods: RNA was extracted from blood samples taken from 9 heart failure patients (8 males) prior to and 7 days after implant of a CF-LVAD. Libraries were sequenced on an Illumina Hiseq2000 and sequences mapped to the human ensembl GRCh37.67 genome assembly using TopHat. Annotations were assigned using the ensembl gene transfer format (GTF) reference. Cufflinks, Cuffmerge and Cuffdiff were used to generate FPKM values. All called gene FPKM values and upper, lower confidence interval bounds were analyzed with a Paired T-test and q-values generated to estimate the false discovery rate. Ingenuity Pathway Analysis was used to investigate pathways. Results: Five of the top ten differentially expressed genes were related to hematopoiesis, neuronal edema, or leukocyte trafficking (carbonic anhydrase I, 7.1 fold up, p< 0.001, q=0.035; alpha hemoglobin stabilizing protein, 6.1 fold up, p 0.002, q=0.043, ABCG2, 4.6 fold up, p< 0.0006 q = 0.02; FAM20A, 4.6 fold up, p 0.0003, q=0.02; CXCL9, 4.0 fold down, p < 0.0006, q=0.03). Of the major pathways represented two involved T cell signaling (CD28 signaling and CTLA4 signaling). Conclusions: To our knowledge this is the first characterization of changes in the human transcriptome in blood one week after LVAD insertion. These findings may represent new pathways for exploration to define mechanisms underlying the adverse events in this patient population.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1666T

Activation of TLR signaling in atherosclerosis and ischemic stroke. *C.-C. Huang[1], R. Sieberg[2], G. Feng[2], S. Wang[3], M. Yu[3].* 1) Dept of Preventive Medicine, Northwestern Univ, Chicago, IL; 2) Feinberg School of Medicine, Northwestern University, Chicago, IL; 3) Dept of Biostatistics & Medical Informatics, University of Wisconsin, Madison, WI.

Atherosclerosis is the major cause of cardiovascular diseases (CVD), the leading cause of morbidity and mortality in the United States and the developed world. Early efforts to understand the etiology of atherosclerosis from large population studies have established that atherosclerosis is primarily caused by dys-regulation of lipid metabolism, compounded by the influence of diverse modifiable risk factors such as smoking and hypertension. However, a wealth of recent investigations has revealed that inflammation also plays a central role in atherosclerosis. Compelling evidence shows that inflammation is a key mediator of atherosclerosis, rather than merely a bystander of this disease. Recently, the immune system has been shown to play a critical role in inflammatory response in atherosclerosis. However, the major genetic pathways and their components involved in regulating immune response in atherosclerosis remain to be elucidated. In this work, we perform an integrative analysis of four public available microarray gene expression data sets in human studies. These studies investigated transcriptomic changes in atherosclerosis, ischemic stroke, and acute inflammatory response in peripheral blood, as well as atherosclerotic lesion progression. Data sets are downloaded from NCBI Gene Expression Omnibus (GEO). Analysis on individual data set shows abundant differentially expressed (DE) genes comparing disease phenotype with control samples or advanced lesions with early lesions. Gene expression fold changes or t statistics were calculated for each study. Parallel comparisons of these statistics show striking correlation (Pearson correlation coefficient ρ: 0.56–0.76) across studies. Gene ontology analysis of DE genes identified several enriched immune and inflammatory pathways, including toll-like receptors (TLR), NF-$\kappa\beta$, p38 MAPK, B-cell, and T-cell signaling. More importantly, pathway mata-analysis identified toll-like receptor (TLR) signaling as the most significant pathway. Since TLR pathways play a pivotal role in initiating and shaping innate and adaptive immune responses, our result suggests that activation of TLR signaling promotes atherosclerosis and eventually lead to clinical events such as ischemic stroke.

1667T

Genetic insights into Nonalcoholic Fatty Liver Disease from the proteomic analysis of HDL particles. *K. Merath[1], M. Zickus[1], R. Komorowski[2], J. Wallace[3], M. Goldbladtt[3], S. Gawrieh[4], M. Olivier[1].* 1) Physiology, Medical College of Wisconsin, Wauwatosa, WI; 2) Pathology, Medical College of Wisconsin, Wauwatosa, WI; 3) Surgery, Medical College of Wisconsin, Wauwatosa, WI; 4) Medicine/Gastroenterology, Medical College of Wisconsin, Wauwatosa, WI.

Nonalcoholic Fatty Liver disease (NAFLD) defines a spectrum of disease manifestations that are associated with a liver fat content of greater than 5 percent. Obesity, insulin resistance, and dyslipidemia are predisposing factors. In a subset of NAFLD patients, the disease progresses to a state characterized by liver inflammation and fibrosis referred to as nonalcoholic steatohepatitis (NASH). Due to the variability in the development of steatosis and the subsequent variation in the progression of the state of liver disease, it is probable that genetic polymorphisms are a contributing factor. Three GWAS have been reported for NAFLD to date and a number of genetic modifiers have been previously reported. However, the likely complex genetic nature of NAFLD and the progression to NASH has not been fully elucidated. In this study, we are taking an alternative approach to identify potential genetic contributors. One of the characteristic changes in the lipid profile of obese individuals is a shift to smaller, denser HDL particles. Smaller and denser HDL particles are thought to lose their cardioprotective properties and thus are referred to as “dysfunctional” HDL. In previous studies, we have shown that this shift to smaller denser HDL particles correlates with quantitative changes in HDL-associated proteins. We have previously validated a technique of separating the larger HDL2 and the smaller HDL3 particles from human serum by size exclusion chromatography. Using nano-HPLC-ESI-tandem mass spectrometry, the proteomes of the particles can be quantitatively assessed. We characterized differences between the HDL particles of NAFLD and NASH patients and control individuals. All individuals were morbidly obese females of Northern European descent. We selected five representative individuals with NAFLD and age-matched control individuals for a comprehensive proteomic analysis using mass spectrometry. Overall, we identified 319 HDL-associated proteins in our analysis. Across all pairs, we identified a total of 12–17 significantly different protein levels, with 4–11 proteins increased in NAFLD patients. The proteins identified included proteins involved in immune signaling, and members of the complement and the coagulation system. The ultimate goal is to use this proteomic data to identify specific proteins and assess the corresponding gene for variations within our cohort that might contribute to the development of NAFLD and/or to the progression to NASH.

1668T

Do variations in genes encoding the argonaute proteins confer risk or protection from cardiovascular disease in Europeans? *I. Predazzi[1], W. Bush[1], S. Williams[1,2].* 1) CHGR, Vanderbilt University, Nashville TN, TN; 2) Dartmouth Medical School, Hanover, NH.

Micro RNAs (miRNAs) are small non-coding RNAs that bind argonaute proteins and regulate post-transcriptional gene expression of specific targets. Many miRNAs have been implicated in the pathogenesis of cardiovascular disease (CVD); however, mechanisms through which miRNAs affect CVD are still unclear. In particular, little is known about how genetic variation in miRNAs and associated genes may affect their function and ultimately disease. We hypothesized that genes encoding proteins that interact with all miRNAs (e.g., Drosha, Dicer-1 and the argonaute proteins) should be conserved across human populations; whereas miRNA sequences that have specific, more limited functions should be less conserved. We assessed the patterns of selection of genes that encode generic miRNA interacting proteins, Drosha, Dicer-1 and the argonaute proteins. We also tested whether regions that encode for miRNA and miRNA targets associated to CVD are subject to population specific patterns of selection, depending on the environment. We found strong signals of selective sweeps in three out of four genes encoding for argonaute proteins (EIF2C1, EIF2C2 and EIF2C3), but only in European samples. Through whole-genome eQTL analysis, we identified SNPs in the genetic regions that regulate expression of these genes and found that these SNPs are most polymorphic in European populations and exhibit large Fst compared to other continental ancestry groups (Fst ~ 0.400). Notably, two of these variants (rs7540413 and rs4026409) have been reported to associate with CAD in a previous genome wide association study. Our data indicate that miRNAs could have population specific expression patterns. Furthermore, SNPs affecting miRNA function could confer risk to disease, and this effect could be different depending on population. Further studies will be necessary to assess the role of these and other genetic variation in miRNA regulation and disease.

1669T

Gene expression profiling in peripheral blood cells in patients with myocardial infarction at young age. *T. Zeller[1], C. Müller[1], A. Schillert[2], M.O. Scheinhardt[2], S. Szymczak[2], F. Ojeda[1], C.R. Sinning[1], R.B. Schnabel[1], S. Wilde[1], P.S. Wild[3,4], K.J. Lackner[5], T. Munzel[4], A. Ziegler[2], S. Blankenberg[1].* 1) Clinic for General and Interventional Cardiology, University Heart Center Hamburg, Hamburg, Germany; 2) Institut für medizinische Biometrie und Statistik, Universitätsklinikum Schleswig-Holstein, Campus Lübeck, Maria-Goeppert-Str. 1, 23562 Lübeck, Germany; 3) Center for Thrombosis and Hemostasis, University Medical Center Mainz, Johannes Gutenberg University Mainz, Germany; 4) Department of Medicine 2, University Medical Center Mainz, Johannes Gutenberg University Mainz, Germany; 5) Institute for Clinical Chemistry and Laboratory Medicine, University Medical Center Mainz, Langenbeckstraβe 1, 55131 Mainz, Germany.

Myocardial infarction (MI) is one of the main causes of mortality. It usually is a disease of the middle-aged and elderly and MI in young adults is a rare phenomenon and explanations about the molecular mechanisms of the genetic predisposition of MI at young age remain elusive. Transcriptome analysis by large-scale gene expression profiling is a promising tool to explore disease-related genes and biological pathways. The aim of this study was to identify differentially expressed genes in young MI individuals, to functionally characterize these genes by pathway analysis and to investigate the relation to the underlying pathophysiology. Using Affymetrix GeneChip Human Exon ST1.0 gene expression data we performed linear regression analysis to identify differentially expressed genes between 235 young MI individuals and 419 population-based controls. After adjustment for classical cardiovascular risk factors age, gender, hypertension, smoking, BMI, diabetes, LDL/HDL ratio and principal components we identified 217 genes being differentially expressed in the young MI group. Most significantly deregulated was *GPR15*, an orphan G-protein-coupled receptor ($p = 5.96*10^{-17}$, up-regulation, fold change 1.78). Pathway analysis using the Ingenuity Pathways Analysis tool (IPA) identified several canonical pathways including the pathways of “PPARα/RXRα Activation”, “Role of NFAT in Cardiac Hypertrophy”, and the “P2Y Purigenic Receptor Signaling” Pathway. These analyses revealed key signaling processes (G-protein signaling, including *P2RY2*, *DUSP1*, *GPR107*, *GPR133*, *PRKAG2*, *PIK3CD*, *ADCY2*, *ADCY7*, *GNAL*) and lipid metabolism (including *ABCA1*, *ABCG1*, *FDFT1*, *GALNT2*, *LIPG*) to be affected in MI patients at young age. Expression quantitative trait analysis to investigate the genetic background of differentially expressed genes revealed SNPs surrounding genes involved in lipid metabolism to be associated with expression of those genes, with the strongest association found for the *FDFT1* gene and SNPs rs2740592 ($p = .19*10^{-13}$), rs4839 ($p = 2.96*10^{-13}$), rs1692821 ($p= 6.39*10^{-13}$). This comprehensive transcriptome analysis of circulating peripheral blood cells in patients affected by MI at young age revealed several profoundly changed genes involved in key pathways. Our results provide novel insights in the genetic basis and expand our current understanding of the pathophysiology.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1670T

A genetic study of familial intracranial aneurysm in French-Canadian and Inuits through exome sequencing. *S. Zhou[1], L. Xiong[1,2], C. Bourassa[1], M. Bojanowski[3], N. Dupré[4], M. Dubé[5], A. Dionne-Laporte[1], D. Spiegelman[1], E. Henrion[1], O. Diallo[1], P. Dion[1], G. Rouleau[1,6].* 1) Centre of Excellence in Neuroscience of Universite de Montreal(CENUM), CHUM Research Center and the Department of Medicine, Montréal; 2) Centre de recherche Fernand-Seguin, Hopital Louis H. Lafontaine, Montreal; 3) Department of Surgery, Universite de Montreal, Montreal; 4) Department of Neurology, Universite Laval, Quebec; 5) Pharmacogenomics Centre, Montreal Heart Institute, Universite de Montreal, Montreal; 6) Ste-Justine Hospital Research Center, Universite de Montreal, Montreal.

Background: Intracranial Aneurysm (IA) is a localized dilation or ballooning of the cerebral artery; a lesion may lead to a consequence of subarachnoid hemorrhage. The cause of primary IA is mainly unknown except for its genetic origin and co-exist factors such as hypertension and smoking. Moreover, genetic predisposition is even higher in familial IA. **Materials & Methods:** We performed whole exome sequencing in 19 affected individuals from three French-Canadian (FC) and one Quebec Inuit families. Targeted capture was done using the SureSelect Human All Exon 50Mb Kit and was sequenced on Illumina Hiseq 2000. Bioinformatic analytical pipline includes FastQC, BWA, GATK and Annovar were used for QC, alignment, variant calling and annotation, respectively. An in-house algorithm was applied for the family-wide variant segregation analyses and based on a selection of 583 in-house controls. **Results & Conclusions:** The initial goal of this study is to look for rare or unique exonic nonsynonymous variants or indels co-segregating within IA individuals of each family. 30 such variants from the exome data of 3 FC families were selected and were validated in all available family members, from which 7 variants showed a >70% penetrance in a separate family. A missense variant located in gene *NOG* was particularly interesting. The variant P84H was presented in all 5 affected individuals from one family but in none of its unaffected members, and was also absent from more than 700 other in-house exomes and 1000 Genome Database. Six different prediction programs indicated that the variant had an extremely deleterious effect to this single exon gene. However, *NOG* mutations have been implicated to cause severe defects such as multiple synostosis and proximal symphalangism which doesn't include IA as part of their clinical descriptions. Therefore, the contribution of this variant to IA in this family is unknown and need further research. At last, the study of combining genome-wide linkage analysis and exome data of the large Inuit IA pedigree is currently undergoing, as well as the additional sequencing of thirty-two IA affected Inuits. From which we hope they will add more information to the discovery of candidate region of Inuit IA.

1671T

Rare variant *APOC3* R19X is associated with cardio-protective profiles in a diverse population-based survey as part of the Population Architecture using Genomics and Epidemiology (PAGE) study. *D.C. Crawford[1,2], L. Dumitrescu[1], R. Goodloe[1], K. Brown-Gentry[1], C. Sutcliffe[1], R. Wiseman[1], P. Baker[1], H.H. Dilks[1,2], J. Boston[1], B. McClellan, Jr[1], P. Mayo[1], M. Allen[1], N. Schnetz-Boutaud[1], J.L. Haines[1,2], T.I. Pollin[3].* 1) Center for Human Genetics Research, Vanderbilt University, Nashville, TN; 2) Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, TN; 3) Division of Endocrinology, Diabetes and Nutrition, Department of Medicine, University of Maryland School of Medicine, Baltimore, Maryland.

High levels of low-density lipoprotein cholesterol (LDL-C) and triglycerides (TG) and low levels of high-density lipoprotein cholesterol (HDL-C) have long been recognized as risk factors for coronary artery disease (CAD), a leading cause of morbidity and mortality. Approximately 100 common genetic variants have been associated with lipid traits through genome-wide association studies, collectively accounting for <15% of the observed heritability. Multiple large effect rare variants in aggregate may account for much of the "missing heritability" as well as provide better predictors of cardio-risk and cardio-protective profiles in a personalized medicine setting. Indeed, a founder mutation was recently discovered and described as conferring favorable lipid profiles and reduced subclinical disease in a Pennsylvania Amish population, and preliminary data have suggested that this null mutation *APOC3* R19X (rs76353203) is rare in the general population. To better describe the frequency and associated lipid profile in a general population setting, we as part of the Population Architecture using Genomics and Epidemiology (PAGE) study and the Epidemiologic Architecture for Genes Linked to Environment (EAGLE) genotyped rs76353203 in the National Health and Nutrition Examination Surveys (NHANES III, 1999–2002, and 2007–2008). Out of 19,613 participants, we identified 30 participants heterozygous and one homozygous for the 19X allele, for an overall allele frequency of 0.08% in this US-representative survey. Among the three major groups, the 19X allele was observed at a higher frequency among Mexican Americans (0.23%) compared with non-Hispanic whites (0.20%) and non-Hispanic blacks (0.02%). Among fasting adults, the 19X allele was associated with lower ln(TG) (n=5,990; β= -0.55; p = 3.6×10-4) and higher HDL-C (n=8,891; β = 15.65; p = 2.1×10-4) and trended towards lower LDL-C (n=6,502; β= -4.85; p = 0.68) after adjustment for age, sex and race/ethnicity. On average, participants with the 19X allele had approximately half the TG levels (geometric means 65 vs. 141 mg/dL), >20% higher HDL-C levels (geometric means 68 vs. 52 mg/dL), and, although not statistically significant, had lower LDL-C levels compared with RR participants (geometric means 110 vs. 121 mg/dL). Overall, these data demonstrate that *APOC3* R19X exists in the general U.S. population in multiple racial/ethnic groups and that 19X is associated with cardio-protective lipid profiles.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1672T

Resequencing of the Cholesteryl Ester Transfer Protein gene (CETP) in U.S. Whites and African Blacks with extreme HDL-C levels. *D. Pirim[1], F.Y. Demirci[1], X. Wang[1], J.E. Hokanson[2], R.F. Hamman[2], C.H. Bunker[3], C.M. Kammerer[1], M.M. Barmada[1], M.I. Kamboh[1].* 1) Human Genetics, GSPH, Univ Pittsburgh, Pittsburgh, PA; 2) Epidemiology, Colorado School of Public Health, Univ Colorado Denver, Aurora, CO; 3) Epidemiology, GSPH, Univ Pittsburgh, Pittsburgh, PA.

Coronary heart disease (CHD) is a global health problem and the main cause of morbidity and mortality in developed countries. In order to decrease the global burden of CHD across the world, the risk factors for CHD have been investigated by conducting both epidemiological and genetic studies. Genetic findings suggest that *CETP* genetic variation has an impact on the inter-individual variation in CHD risk by affecting the plasma lipid levels. Although several common genetic variants were reported to be associated with *CETP* deficiency and CHD risk, undetermined variations in this gene still stir up considerable research interest. To investigate the role of both common and rare variation in *CETP* in regulating plasma HDL-C levels, we resequenced the entire gene (22kb) and 1kb flanking regions in individuals with HDL-C levels of lower 5th percentile (48 Whites, 47 Blacks) and upper 5th percentile (47 Whites, 48 Blacks), selected from a non-Hispanic White (NHW) U.S. population and an African Black population. We identified a total of 280 variants, of which 45 were present only in NHWs and 157 were present only in Blacks. Out of 280 variants, 136 were common with a minor allele frequency (MAF) ≥0.05. Fourteen variants were located in coding regions, of which 9 were predicted to cause non-synonymous changes. Thirty-six common variants showed significantly different allele frequencies between high and low HDL-C subgroups in at least one ethnic group. The variants that showed most significant MAF distribution between the two subgroups were rs17231506 (p=1.00×10-6) and rs17231520 (p=2.36×10-6) in Whites and Blacks, respectively. These two variants were already reported to be strongly associated with HDL-C levels in different ethnic groups; hence, our data confirmed previous associations. We observed 74 rare or relatively uncommon variants (MAF<0.05) unique to either high or low HDL-C subgroups in at least one ethnic group, suggesting the potential importance of *CETP* rare variation in determination of the HDL-C levels. Our complete sample sets of NHWs (n=623) and Blacks (n=788) were comprehensively genotyped for the rare and common variants identified by resequencing; the results are being analyzed for associations with plasma lipid levels using single-site, haplotype, and burden tests. The results of this comprehensive analysis will help us to better understand the role of both common and rare variants in *CETP* in regulation of plasma HDL-C levels and other lipids.

1673T

Comprehensive evaluation of the effects of *APOE* genetic variation on plasma lipoprotein traits in U.S. Whites and African Blacks. *Z.H. Radwan[1], F.Y. Demirci[1], X. Wang[1], F. Waqar[1], J.E. Hokanson[2], R.F. Hamman[2], C.H. Bunker[3], C.M. Kammerer[1], M.M. Barmada[1], M.I. Kamboh[1].* 1) Human Genetics, GSPH, Univ Pittsburgh, Pittsburgh, PA; 2) Epidemiology, Colorado School of Public Health, Univ Colorado Denver, Aurora, CO; 3) Epidemiology, GSPH, Univ Pittsburgh, Pittsburgh, PA.

Cardiovascular diseases (CVD) continue to be a leading cause of mortality and morbidity worldwide. CVD risk factors are well established and dyslipidemia with high levels of low-density lipoprotein cholesterol (LDL-C) and low levels of high-density lipoprotein cholesterol (HDL-C) is a significant risk factor. Thus, unraveling the genetic determinants of the interindividual variation in plasma lipoprotein levels that modulate the CVD risk has public health importance. Apolipoprotein E (ApoE) has a crucial role in lipoprotein metabolism and its genetic variation was found to have effects on lipoprotein levels and CVD risk. The objective of this study was to identify the common and rare genetic variants of *APOE* through comprehensive resequencing of all exons, introns and 1 kb flanking regions of the gene in 95 individuals with extreme HDL-C levels (upper and lower 5th percentiles), selected from two distinct well-defined epidemiological samples [U.S non-Hispanic Whites (NHWs) and African Blacks], and then evaluate the tagSNPs and all relevant rare variants in entire sample sets (623 NHWs and 788 Blacks) to examine their effects on lipid profiles. Analysis of sequencing data revealed a total of 42 variants (40 substitutions and 2 indels), of which 11 were present in both ethnic groups. In NHWs, 20 variants were observed, of which 4 were located in exons, 2 were indels, and 56.0% had ≥5% minor allele frequency (MAF). In Blacks, out of 32 variants observed, 3 were located in exons, 2 were indels, and 30.0% had ≥5% MAF. In NHWs, 28% of the individuals in the high HDL-C group (13/47) had at least one rare or relatively uncommon variant (MAF<5%) versus 13% in the low HDL-C group (6/48). Similarly in Blacks, 52% of the individuals in the high HDL-C group (25/48) had at least one rare or relatively uncommon variant versus 36% in the low HDL-C group (17/47). Following comprehensive genotyping of the identified variants in entire sample sets, single-site association analysis of variants with MAF≥1% revealed 8 variants in NHWs (MAFs ranging between 0.021 and 0.477) and 9 variants in Blacks (MAFs ranging between 0.017 and 0.366) significantly associated with one or more lipoprotein traits. Haplotype analyses and rare variant burden tests have also been performed. Our findings indicate that APOE genetic variation has a significant role in regulating plasma lipoprotein levels, further supporting the importance of ApoE in lipoprotein metabolism.

1674T

Filamin-C: a novel candidate for familial restrictive cardiomyopathy identified by whole-exome sequencing. *M. Tariq, E. Miller, S. Ware.* Division of Molecular Cardiovascular Biology and The Heart Institute, Cincinnati Children's Hospital Medical Center, Cincinnati, OH.

Restrictive cardiomyopathy (RCM) is a rare and distinct form of cardiomyopathy characterized by normal ventricular chamber dimensions, normal myocardial wall thickness, and preserved systolic function. The abnormal myocardium, however, demonstrates impaired relaxation. RCM accounts for <5% of all cardiomyopathies in the US and Europe. To date, dominant mutations causing RCM (~10% of all cases) have been reported in a small number of sarcomeric or cytoskeletal genes including *DES, ACTC1, TNNI3, TNNT2*, and *MYH7*. In this study, we investigated a large family with multiple affected individuals diagnosed with autosomal dominant isolated RCM. All affected individuals presented in childhood without evidence of myopathy or other extra-cardiac findings. Clinical genetic testing for sarcomeric genes was negative in the proband. To identify the cause of disease in the family, whole-exome sequencing of the proband was undertaken with >30x coverage resulting in 26,065 coding variants. After filtering by comparison to 1000 Genomes and dbSNP135 datasets, 2,913 non-synonymous changes were left. Targeted filtering of these changes for ~300 genes previously implicated in cardiomyopathy or related-phenotypes, heart development, or cardiac function identified 4 heterozygous changes: one each in *MYPN, LDB3, PKP2, FLNC*. After screening 4 affected and 2 unaffected individuals of the family, only *FLNC* (c.C6893T; p.Pro2298Leu) segregated appropriately with the disease-phenotype. Filamin C is an actin crosslinking protein found in skeletal and cardiac muscle and localizes to the z-disk, sarcolemma, and sarcoglycan-related complex. A small number of mutations in the FLNC N-terminal actin binding and rod or dimerization domains have been identified in distal myopathy and myofibrillar myopathy, respectively. The *FLNC* p.Pro2298Leu mutation falls in the immunoglobulin-like repeat of the protein affecting a highly conserved residue, and is predicted to be damaging by 5 different bioinformatic programs. In analogy to the domain-specific phenotypes caused by mutations in related family members such as *FLNA* and *FLNB*, we predict that the *FLNC* p.Pro2298Leu mutation causes isolated RCM without skeletal myopathy. Functional analyses are ongoing. In conclusion, we identified a new candidate for autosomal dominant RCM using whole-exome sequencing. This study extends the current clinical and genetic spectrum of filaminopathies by associating a novel *FLNC* mutation to a distinct phenotype.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1675T

Mutations in SMAD3 in a British cohort of thoracic aortic aneurysm and dissection (TAAD) patients. *G. Arno[1], J.A. Aragon-Martin[1], O. Song[1], N.P. Kamali[1], A. Saggar[2], M. Jahangiri[3], A.H. Child[1].* 1) Cardiac and Vascular Sciences, St George's University of London, London, United Kingdom; 2) Clinical Genetics Unit, St George's University of London, Cranmer Terrace, London SW17 0RE, UK; 3) Department of Cardiothoracic Surgery, St. George's Healthcare NHS Trust, London, UK.

Background: Thoracic aortic aneurysm and dissection (TAAD) is a feature of several genetic conditions, such as Marfan (MFS) and Loeys-Dietz (LDS) syndromes. However, there is a growing body of work suggesting that non-syndromic TAAD is also of genetic origin. Recently studies have implicated mutations in *SMAD3* (MIM*603109) as a cause of aneurysm-osteoarthritis syndrome (AOS) and 2% of non-syndromic familial TAAD. AOS is characterised by thoracic aortic aneurysm with a risk of early dissection, aneurysm and tortuosity throughout the arterial tree, joint anomalies with osteoarthritis, osteochondritis dissecans (OCD) and meniscal anomalies. It was the aim of this study to determine the frequency of *SMAD3* mutation in a consecutive series of TAAD patients attending a British MFS cardiac genetics clinic. Diagnosis was based on pedigree, physical examination, echocardiogram, CT and MRI studies and operative reports. **Method:** A total of 87 UK patients (71M:16F, mean age 51±13.1) with known TAAD who did not fulfill the revised Ghent criteria for MFS, and with no demonstrable mutations in *FBN1*, *TGFBR2* or *ACTA2* were recruited to this study. These patients were screened for mutations in all exons of *SMAD3* including intron/exon boundaries by bi-directional sequencing. **Results:** Five novel mutations were identified in 6/87 (6.9%) probands. These mutations comprised: in exon 1, a missense mutation - p.Ile67Ser (c.200T>G) (found in two unrelated probands); in exon 2, a missense mutation - p.Arg90Cys (c.268C>T); in intron 2 - c.401-6G>A - predicted to create a cryptic splice site; in exon 6, a nonsense mutation - p.Tyr238* (c.714C>A); and also in in exon 6, a duplication leading to a frameshift and premature termination codon - p.Arg287Thrfs*24 (c.858dupA). None of the mutations were found in 208 control chromosomes or are reported in any SNP databases. The cryptic splice site mutation found in intron 2 was identified in an affected sibling of the proband, familial screening is ongoing. **Conclusion:** This study supports data from previous studies that link mutations in *SMAD3* with TAAD. Furthermore, our findings suggest it may be a more important cause of aneurysmal disease in our British population than equivalent similar US or European studies.

1676T

Distribution of sarcomere vs. non-sarcomere gene mutations in over 2400 hypertrophic and dilated cardiomyopathy patients. *S. Baxter[1], A. Daly[1], D. Macaya[1], N. Smaoui[1], G. Richard[1], W. Chung[2].* 1) GeneDx, Gaithersburg, MD; 2) Department of Pediatrics, Columbia University Medical Center, New York, New York, USA.

Hypertrophic and dilated cardiomyopathy (HCM/DCM) are primarily disorders of the sarcomere. Rarely, mutations in genes encoding non-sarcomere proteins have been associated with both forms of cardiomyopathy, including non-syndromic (Z-disk genes) and syndromic forms (protein storage and mitochondrial disorders). We assessed 1838 HCM and 630 DCM cases that were sent for clinical diagnostic testing over the course of two years to study the mutation distribution and contribution of genes encoding non-sarcomere proteins to HCM and DCM. The HCM cases were sequenced for 18 genes including 9 sarcomere-encoding genes (ACTC1, MYBPC3, MYH7, MYL2, MYL3, TNNC1, TNNI3, TNNT2, and TPM1) and 9 non-sarcomere-encoding genes (CAV3, GLA, LAMP2, MTTG, MTTI, MTTK, MTTQ, PRKAG2, and TTR). Of the 1838 HCM cases, 38% (n=706) had a positive result and 5% (n=85) had one or more variant of unknown clinical significance (VUS), 4% in sarcomere genes and 1% in non-sarcomere genes. A disease-causing mutation in a sarcomere gene was identified in 677 cases compared to 34 cases with mutations in non-sarcomere-encoding genes, including CAV3 (n=1), GLA (n=9), LAMP2 (n=4), mitochondrial genes (n=5), PRKAG2 (n=9), and TTR (n=6). Five percent (n=84) of all cases had more than one mutation or VUS. The DCM cases were sequenced for 27 genes including 6 sarcomere-encoding genes (ACTC1, MYBPC3, MYH7, TNNI3, TNNT2, and TPM1) and 21 non-sarcomere-encoding genes (DES, LAMP2, LDB3 (ZASP), LMNA, MTND1, MTND5, MTND6, MTTD, MTTH, MTTI, MTTK, MTTL1, MTTL2, MTTM, MTTQ, MTTS1, MTTS2, PLN, SGCD, TAZ, and TTR). Of the 630 DCM cases, 26% (n=165) were positive and 10%(n=65) had VUSs evenly split between sarcomere and non-sarcomere genes. Compared to HCM, DCM has a higher frequency of non-sarcomere mutations among positive cases, 37% (n=61). The majority of pathogenic mutations in non-sarcomere genes were identified in LMNA (n=34) and mitochondrial genes (n=10). Like HCM, 5% (n=33) of the DCM cases also had more than one mutation or VUS. Our data supports the notion that HCM and DCM are predominantly diseases of the sarcomere; however, DCM has a significantly higher prevalence of mutations in non-sarcomere genes ($p<0.001$). Although storage disorders and mitochondrial disorders are relatively rare causes of HCM and DCM, they are important to include in molecular testing panels for prevention, early diagnosis, and management of cardiac and extra-cardiac symptoms.

1677T

Burden of rare sarcomere gene variants in the Framingham and Jackson Heart Study cohorts. *A. G. Bick[1,2], J. Flannick[2,3], K. Ito[1], S. Cheng[4], R. S. Vasan[5,6], M. G. Parfenov[1], D. S. Herman[1], S. R. DePalma[1], N. Gupta[2], S. B. Gabriel[2], B. H. Funke[7], H. L. Rehm[7], E. J. Benjamin[5,6,8], J. Aragam[4], H. A. Taylor[9,10], E. R. Fox[9], C. Newton-Cheh[3,5,11], S. Kathiresan[3,5,11], C. J. O'Donnell[5,12], J. G. Wilson[13], D. M. Altshuler[2,3], J. N. Hirschhorn[1,2,14], J. G. Seidman[1,2], C. Seidman[1,2,4].* 1) Department of Genetics, Harvard Medical School, Boston, MA; 2) Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, MA; 3) Center for Human Genetic Research, Massachusetts General Hospital, Boston, MA; 4) Division of Cardiovascular Medicine, Brigham and Women's Hospital, Boston, MA; 5) National Heart, Lung, and Blood Institute's Framingham Heart Study, Framingham, MA; 6) Departments of Cardiology and Preventive Medicine, Boston University School of Medicine, Boston, MA; 7) Laboratory for Molecular Medicine, Partners Center for Personalized Genetic Medicine, Cambridge, MA; Department of Pathology, Harvard Medical School, Boston, MA; 8) Epidemiology Department, Boston University School of Public Health, Boston, MA; 9) Department of Medicine, University of Mississippi Medical Center, Jackson, MS; 10) Jackson State University, Jackson, Mississippi, USA; Tougaloo College, Tougaloo, MS; 11) Cardiology Division, Massachusetts General Hospital, Boston, MA; 12) Division of Intramural Research, National Heart, Lung, and Blood Institute, Bethesda, MD; 13) Department of Physiology and Biophysics, University of Mississippi Medical Center, Jackson, MS; 14) Divisions of Genetics and Endocrinology and Program in Genomics, Children's Hospital, Boston, MA.

Rare non-synonymous mutations in sarcomere genes cause hypertrophic and dilated cardiomyopathies, dominant disorders that are characterized by cardiac remodeling and dysfunction. To evaluate whether allelic variants in these genes are associated with cardiac morphology and function in the general population, eight sarcomere genes (*ACTC1, MYBPC3, MYH7, MYL2, MYL3, TNNI3, TNNT2, TPM1*) were sequenced in 1,637 individuals from the Framingham Heart Study (FHS) Offspring cohort and 1,963 individuals from the Jackson Heart Study (JHS) cohort; echocardiography was used to assess cardiac structure and function. Amongst the sequenced individuals, 11.2% had one or more rare (MAF<1%) non-synonymous coding variants. The prevalence of rare likely pathogenic sarcomere variants was 0.6%, at least two times higher than previously anticipated. Individuals with rare sarcomere variants had significantly increased mean left ventricular wall thickness (LVWT) of 3% ($p=.01$) and 2% ($p=.05$) in the FHS and JHS cohorts, respectively. Sarcomere gene variants previously identified as pathogenic that were detected here conferred an 8.6% greater LVWT in the FHS cohort ($p=.05$, n=14) but no change in the JHS cohort (n=8). Only 4 out of the 22 individuals carrying likely pathogenic variants had clinical manifestations of HCM implying that pathogenic variants identified in affected individuals may have reduced penetrance when ascertained in the general population. Additionally individuals with the combination of a common variant rs3729799 (*MYBPC3*, p.R1002W) and a second rare *MYBPC3* variant had a 15% decreased LVWT in comparison to individuals with rs3729799 and no rare variant ($p<10^{-5}$) implying that multiple variants may have more effect on cardiac morphology than each alone. Rare sarcomere variants were associated with increase left ventricular dimensions over time and an increased risk for adverse cardiovascular events over a follow-up period of 30 years (hazard ratio: 2.3; 95% confidence interval: [1.2, 4.1]) in the FHS cohort. Our sequence analysis of eight sarcomere protein genes, suggests that analysis of rare variants may potentially contribute to cardiovascular risk assessment in the general population.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1678T

Whole-exome sequencing for identification of novel genes and variants underlying the risk of Lonq-QT syndrome. *M. Chaix[1], P. Goyette[1], A. Alikashani[1], F. Latour[1], L. Robb[1], B. Mondesert[1], L. Rivard[1,2], S. Foisy[1], M. Ladouceur[1], R.M. Hamilton[3], M. Talajic[1,2], J.D. Rioux[1,2].* 1) Montreal Heart Institute, Montreal, Canada; 2) Universite de Montreal, faculty of medicine; 3) The Hospital for Sick Children, University of Toronto.

Long QT syndrome (LQTS) is a cardiac disorder characterized by elongation of the cardiac repolarisation characterized on the ECG by an elongation of the QT interval. This syndrome is a cause of sudden cardiac death mostly in young individuals without any medical history. To date, 13 genes have been associated with LQTS risk (LQT1-13), with the majority of cases showing mutations in the LQT1 (KCNQ1), LQT2 (KCNH2) and LQT3 (SCN5A) genes. The different LQTS genes have been associated with differing clinical expression of the disease, and all implicate ion channel complexes (for sodium, potassium and calcium) or accessory proteins interacting with them. High genetic heterogeneity has also been observed in this disease, with several hundred mutations identified in associated loci, many of which being private/familial. Genetic testing of individuals with elevated disease risk (based on the Schwartz's clinical diagnostic score), through sequencing of the most common genes in the clinic identifies causal genetic variants in up to 60–70% of cases, suggesting that additional genes may be involved in this disease. In the current study, we have recruited 24 patients with elevated disease risk for LQTS, but reported negative for mutations in the known LQTS loci in clinical testing, and have used a Next generation sequencing (NGS) whole-exome DNA sequencing approach. We have developed a stepwise analytic approach that focuses on (1) the known LQTS genes, (2) genes in loci identified in genome-wide association studies of QT-interval and sudden cardiac death, (3) known ion channels and accessory proteins, and (4) loci with expression patterns consistent with a role in LQTS; we will also weight any locus showing multiple variants across our different patients. Identification of novel genes implicated with LQTS through this approach will allow a better understanding of the pathophysiologcal mechanism underlying this condition and should improve clinical diagnosis of patients and families that are currently reported as negative for the known LQTS loci.

1679T

A rare genetic mutation (FBLN-4 mutation) found in a patient presenting with diffuse severe aortopathy. *K. Coleman[1], P. Fernhoff[1], S. Shankar[1], C. Hebson[2], M. Clabby[2], B. Kogon[3], B. Loeys[4].* 1) Children's Healthcare of Atlanta, Department of Human Genetics, Emory University, Atlanta, GA; 2) Sibley Heart Center, Children's Healthcare of Atlanta, Atlanta, Ga; 3) Department of Cardiothoracic Surgery, Emory University, Atlanta, Ga; 4) Antwerp University Hospital/University of Antwerp, Antwerp, Belgium.

A 6-month-old female presented with acute onset difficulty breathing progressing to respiratory failure. Chest films showed a large mediastinum; chest MRI eventually detailed a severely dilated aorta from aortic root to aortic hiatus. The patient was diagnosed with RSV bronchiolitis as the etiology of the acute respiratory distress and compression of the main left bronchus by the dilated aorta as the etiology for failure to extubate. She thus required a tracheostomy prior to weaning from mechanical ventilation. The aorta was eventually surgically treated with an “elephant trunk” type ascending aorta repair and plication with reimplantation of the all of the strap vessels. Months after surgical repair, the patient was successfully decannulated and no longer requires any airway support. The eventual plan is a second and hopefully final surgery on her descending aorta to reduce its caliber. Otherwise, during the clinical progression, an ongoing attempt at defining an underlying genetic etiology was undertaken. Specific testing for Loeys-Dietz syndrome (TGFBR2 and TGFBR1) and arterial tortuosity syndrome (SLC2A10) was negative. DNA was sent to Dr. Bart Loeys in Brussels, Belgium for sequencing of the Fibulin-4 (FBLN4) gene that is a member of the fibulin family, a group of extracellular matrix proteins prominently expressed in medial layers of large veins and arteries. The result was a homozygous FBLN4 mutation (p.E126K) in this patient. Both parents who are consanguineous were heterozygous for the same mutation. Mutations in FBLN4 have previously been described in four patients including one patient from Australia with the identical mutation as this patient. This case report thus describes a severe clinical presentation of an extremely rare genetic mutation, which has to date been successfully surgically managed.

1680T

Rare Variants in Notch Pathway Genes Among Individuals with Congenital Heart Defects. *K. McBride[1,2], G. Zender[1], S. Fitzgerald-Butt[1], J. Stemming[1], A. Hooper[1].* 1) Molecular & Human Genetics, Nationwide Children Hospital, Columbus, OH; 2) Dept Pediatrics, College of Medicine, Ohio State University, Columbus, OH.

The Notch pathway is an important signaling system involved in development of the myocardium, valves, and outflow tract of the heart, regulating epithelial to mesenchymal transition (EMT). Mutations in *NOTCH1* were first identified in two families demonstrating autosomal dominant inheritance for aortic valve disease. We subsequently found *NOTCH1* variants among individuals with other apparently sporadic congenital heart defects of the left ventricular outflow tract (LVOT), including aortic valve stenosis, bicuspid aortic valve, coarctation of the aorta, hypoplastic left heart syndrome. These variants caused dysfunction of Notch signaling, leading to downstream effects on EMT. We hypothesized that variation in other genes of the Notch pathway might also cause LVOT defects. Multiple genes of the Notch pathway were screened in over 450 individuals with an LVOT defect, coding for ligand *DLL4*, NOTCH1 intracellular domain cleavage *PSEN1*, transcriptional repressors *HEY1*, *HEY2*, *HEYL*, and effectors of EMT *SNAI1, SNAI2*. Variants causing amino acid substitutions were found in *DLL4* (p.R248Q, p.L622G); *PSEN1* (p.R35Q, p.R358Q, p.Q459K); *HEY1* (p.K249T p.F275L); *HEY2* (p.S12N p.S172P p.P179L p.T228A); *HEYL* (p.R141C p.P173S p.W177C p.F179L); but not in *SNAI1* or *SNAI2*. These variants were not identified in 300 control chromosomes and were absent in the 1000 Genomes database or present at a MAF of <0.001. We used a cell based assay to study the effects of select Hey gene variants on repression of GATA4/6. Hey wild type and mutant expression vectors, GATA4/6 expression vector and ANF promoter luciferase reporter were transfected into HeLa or 293 cells. A statistically significant loss of repression was demonstrated for *HEYL* p.W177C and p.F179L ($p<0.001$ for difference in ANF luciferase activity). In conclusion, we identified a number of rare variants in genes important in the Notch signaling pathway among individuals with an LVOT defect, several of which have demonstrated loss of function.

1681T

TTN Mutations from Exome Sequences of 17 DCM Families: Clinical Implications. *A. Morales[1], N. Norton[1], D. Li[1], E. Rampersaud[2], E. Martin[2], S. Zuchner[2], S. Guo[2], M. Gonzalez[2], R. Hershberger[1].* 1) Cardiovascular Division, Department of Medicine, University of Miami Miller School of Medicine, Miami, FL; 2) Hussman Institute for Human Genomics, University of Miami Miller School of Medicine, Miami, Florida.

BACKGROUND: Truncating variants in TTN were recently implicated in up to 25% of DCM and 3% of controls. METHODS: To evaluate the role of TTN in a broader genetic background, we analyzed exome data from 17 families (48 cases) in which the proband was found not to carry mutations in 16 known DCM genes. A filtering strategy for segregating, rare (<0.5% controls), and conserved (Phastcons >0.4 or GERP >2) coding or splice site variants in genes with known myocardial expression was followed. TTN variants meeting filtering criteria were confirmed and evaluated in family members by Sanger sequencing. Exome data was also queried for DCM variants meeting all criteria except for segregation. RESULTS: Variants in an average of 28 candidate genes per family were identified (range 1–80). TTN was among the candidates in nine families (7 with truncating, 2 with missense variants). Truncating variants may be more relevant, as the average number of TTN truncating variants in the Exome Sequencing Project (ESP) was 0.014 per individual while the average number of missense variants per individual in our DCM cases and non-DCM ESP exomes was 23. Among the 7 families with truncating variants, 22 of 35 TTN carriers were affected. Median age of onset was 42, ejection fraction 33.5% and left ventricular size 62 mm. In addition to TTN truncations, 2 families had potentially relevant variants that met filtering criteria in SYNE2 and SOS1, respectively. Non-segregating variants in ACTN2, ANKRD1, CRYAB, CSRP3, DMD, DES, LMNA, MYBPC3, MYH6, MYH7, RBM20, TCAP, TMPO, TPM1, and TTN were identified in 6 families. Non-segregating TTN truncating variants were identified in 2 additional families. DISCUSSION: Our data confirm that TTN truncating variants are frequent in DCM. While genetic testing for DCM should encompass as many known genes as possible and clinical exome sequencing is now considered for patients with negative known gene panels, clinicians face increasingly complex issues of testing strategy and variant interpretation. Whole panel testing should be considered in more than one family member, especially with widely disparate ages of onset, clinical presentation or meiotic distance. Missense TTN variants are expected; their relevance remains unknown. Rare, non-segregating variants in TTN and other relevant genes occur. Comprehensive analysis of much greater numbers of DCM exome sequences, ideally in well phenotyped families, will be needed to fully understand DCM genetics.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1682T

A novel Ser33045Ala variant of TTN gene in a family with left ventricular non-compaction and syncope. Is there a linked etiopathogenesis? . *A. Psychogios, A. Marsidi.* Dept Pediatrics, Division of Medical Genetics and Genomic Medicine, East Tennessee Univ/James H Quillen Col Med, Johnson City, TN.

Left ventricular non-compaction (LVNC) is a rare cardiomyopathy defined by ventricular trabeculations resulting in arrhythmia, heart failure, cardioembolic events, and sudden death. Several LVNC genes are reported but the spectrum remains unclear. We report a 28-year old Caucasian female with history of arrhythmia and heart failure secondary to LVNC. Her growth and developmental history are normal. Family history is significant for maternal cardiomyopathy and hypertension, paternal history of syncope and sudden death, and a brother who died of infantile heart disease. **On examination,** she is well developed, thin, without dysmorphic features. **Clinical molecular testing:** Initial sequencing analysis for LVNC at John Welsh Cardiovascular Diagnostic Laboratory, Houston, TX including Taffazin (TAZ, G4.5) gene, Z-band alternatively spliced PDZ motif-containing protein gene (ZASP), and Lamin A/C (LMNA) was normal. Further DNA sequencing of the coding regions and splice sites of 46 cardiomyopathy genes (Pan Cardiomyopathy Panel) at the Laboratory for Molecular Medicine, Center for Personalized Medicine, Harvard Medical School, Cambridge, MA identified two novel variants. A heterozygous c.99133T>G (p. Ser33045Ala) in exon 309 of the TTN gene and a heterozygous c.1159-13C>T in intron 7 of the JUP gene. The Ser33045Ala variant (TTN) has not been reported in the literature nor previously identified. Computational analyses (biochemical amino acid properties, conservation, AlignGVGD, PolyPhen2, and SIFT) did not provide strong support for or against an impact to the protein. This variant has been seen in 0.03% (2/6666) of European American chromosomes in a broad population by the NHLBI Exome Sequencing Project (http://evs.gs.washington.edu/EVS). In addition, over 30 benign TTN gene polymorphisms were identified in our patient. **Conclusion:** A novel Ser33045Ala TTN gene sequence variant is reported in a 28-year-old Caucasian female with LVNC, heart failure, arrhythmia, and family history of syncope and sudden death. The TTN gene role in the pathogenesis of LVNC and/or syncope in this family is unclear. Frequent follow-up and medical therapy for arrhythmias and possible increased risk for syncope were recommended independently of her carrier mutation status. Further studies are needed to clarify the association of TTN gene in the etiopathogenesis of LVNC or syncope.

1683T

Molecular autopsy for sudden death using whole genome sequencing. *M.J. Puckelwartz[1], L. Dellefave-Castillo[1], D. Wolfgeher[1], V. Nelakuditi[2], M.T. Campbell[3], J.R. Golbus[1], E.M. McNally[1,2].* 1) Medicine, University of Chicago, Chicago, IL; 2) Human Genetics, University of Chicago, Chicago, IL; 3) Integrated DNA Technologies, Inc., Coralville, IA.

Sudden, unexpected death in the young is rare and is often attributed to cardiac arrhythmia events. Cardiac arrhythmia may leave few pathological clues, and pathogenic DNA sequence variation in arrhythmia genes may be helpful in identifying the cause of death and providing useful information to family members. Genetic testing for sudden death has focused on genes encoding cardiac ion channels with very limited sensitivity. We applied whole genome sequencing to a young individual with sudden death to examine genes beyond those previously associated with heritable cardiac arrhythmia. A rare variant in RYR1 encoding the ryanodine receptor 1 was identified. RYR1 mutations are associated with malignant hyperthermia, heat stroke, seizures and death. Because the decedent had pre-mortem symptoms attributable to heat intolerance, the RYR1 mutation was concluded to be a primary risk allele. Several variants of high population frequency were also identified that may have had a contributory role because of their predicted pathogenic nature. We conclude that the combination of genetic variants and environmental conditions contributed to sudden death in this individual and that autopsy-derived whole genome analysis should be considered in cases of young unexplained deaths.

1684T

High incidence of *GLA* variants in a non-selected heart disease patient population suggests that the Fabry trait is a common cardiovascular genetic risk factor. *R. Schiffmann[1], S. Forni[1], C. Swift[1], X. Wu[2], D.J. Lockhart[2], M. Chee[3], T. Kitaoka[3], E. Chudin[3], S. Pond[3], N.H. McNeill[1], K. Sims[4], E.R. Benjamin[2], L. Sweetman[1].* 1) Inst Metabolic Disease, Baylor Res Inst, Dallas, TX; 2) Amicus Therapeutics, Cranbury, NJ; 3) Prognosys Biosciences, Inc., La Jolla, CA; 4) Neurogenetics DNA Diagnostic Laboratory, Massachusetts General Hospital, Boston, MA.

We have initiated a large screening research protocol for the Fabry disease biochemical and genetic trait in a non-selected population of patients, age 19–85 years, with any form of cardiac disease with either coronary artery disease, conduction or rhythm abnormalities, cardiomyopathy or valvular defects. We hypothesized that α-galactosidase A deficiency is a genetic risk factor for non-specific cardiovascular complications that occur in the general population. The verification of our hypothesis will confirm the Fabry trait as a potentially modifiable cardiovascular risk factor, allowing for better management and prevention of heart disease. Screening is performed by measuring urinary globotriaosylceramide (Gb_3) in random samples of whole urine using tandem mass spectrometry, measuring α-galactosidase A activity in dried blood spots, and looking for *GLA* gene mutations by parallel sequencing of the whole gene in pooled genomic DNA samples. Individual samples from selected pools are further analyzed by conventional Sanger sequencing. Thus far we studied 1110 unselected consecutive adult patients, age 19–85 years, with a variety of cardiac ailments including coronary artery disease, conduction and rhythm defects, hypertrophic cardiomyopathy and valvular disease. Exonic mutations were confirmed in four female cardiac patients. Intronic variants were confirmed in nine other cardiac patients, eight males and one female, with above normal levels of urinary Gb_3. In four of these nine patients, α-galactosidase A activity ranged from 30–50% of the mean seen in normal controls. We conclude that the general heart disease population contains up to 1% of individuals with exonic *GLA* variants and intronic *GLA* variants associated with elevated urinary Gb_3. Variations in the *GLA* gene may be a relatively common genetic risk factor (the Fabry trait) for all forms of heart disease.

1685T

Exome sequencing identifies homozygous truncation mutations in *MYOM2* in arrhythmogenic ventricular dysplasia/cardiomyopathy. *J. Wang, C. Silverstein, H. Lee, E. Mokhonova, A. Seki, S.P. Taylor, M. Fishbein, M. Spencer, S.F. Nelson.* UCLA, Los Angeles, CA.

Background: Arrhythmogenic right ventricular dysplasia/cardiomyopathy (ARVD/C) is a genetic disorder of the myocardium characterized by progressive fibrofatty replacement that predisposes to ventricular tachycardia and sudden death. Although it classically affects the right ventricle, left and biventricular forms are increasingly recognized. A mutation in one of the 8 known genes, primarily encoding desmosomal proteins, is found in fewer than half of patients. We identified a family with onset of dilated cardiomyopathy in the 2nd or 3rd decade in four out of four siblings with progression to cardiac transplantation so far in three. Gross and histopathologic examination of the three explanted hearts demonstrated the characteristic fibrofatty replacement of ARVC/D involving predominantly the left ventricle. The family reported a high degree of consanguinity.

Methods: To identify the molecular basis of these findings in the family, exome sequencing was carried out in the index patient at the UCLA Clinical Genomics Center. Exome variants were filtered for rare heterozygous variants and homozygous variants to identify new mutations. Mean oversampling of bases was 163X and at least 95% of the ~33.4 Mb of protein coding sequence defined by RefSeq was read to at least 10X depth.

Results: A single plausible mutational target, a homozygous truncation mutation in *MYOM2*, was identified genome-wide. MYOM2 is a sarcomere structural protein localized to the M-line, and these individuals would thus be predicted to make no functional MYOM2. The mutation was confirmed by Sanger sequencing, and the index patient's siblings' and father's DNA were collected to confirm the mutation. Quantitation of the regions of homozygosity was consistent with a closer than 1st cousin relationship in the subject's parents.

Conclusions: Our findings implicate *MYOM2* as a new gene for ARVD/C and provide new evidence of a critical role for MYOM2 in cardiac sarcomere structure and function. The significance of an MYOM2 truncation mutation in the heterozygous state is to be established.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1686T

Replication of Associations from Large Scale Re-sequencing Studies. *D. Waterworth[1], L. Li[2], R. Scott[3], C. Gillson[3], J. Aponte[2], L. Warren[2], M. Nelson[2], M. Ehm[2], N. Wareham[3], S. Chissoe[2].* 1) Gen Div, GlaxoSmithKline, King of Prussia, PA; 2) Genetics, GlaxoSmithKline, Research Triangle Park, NC; 3) MRC Metabolic Unit, Institute of Metabolic Science, Addenbrookes Hospital, Cambridge, UK.

202 drug-target genes were sequenced in 14,002 individuals as part of a drug-repositioning and target validation experiment. In focusing on coding variation it was hoped that we would be able to identify gene-specific functional tool variants that could be utilized in predicting drug efficacy, alternative indications and safety. An abundance of rare, putatively functional, variation was found, providing opportunities for translational research; however large sample sizes will be needed for genotype-phenotype association studies. Some of the most significant associations in the cardiovascular and metabolic disease areas were with low frequency variants and intermediate traits, however it was not known how robust these findings were due to their relatively low power. Therefore we selected the 43 most promising variants from 27 genes (two thirds of the variants with MAF less than 0.05) for the purpose of replicating initial association results. These variants were typed in up to 6 different cohorts (CoLaus, Lolipop, EPIC-Norfolk, Norfolk-T2D, ELY and Fenland) by either Sequenom or Kaspar technologies, up to a total of 39,290 subjects and analyzed for a broad range of cardiovascular and metabolic phenotypes. Eight of the variants could not be designed or failed to generate viable genotypes and many did not replicate with their original trait. However, a low frequency (MAF 0.015) coding variant in GLP1R (A316T) was associated with glucose (p=0.0004), HOMA-B (p=0.0009) and waist (p=0.001), recapitulating the therapeutic effect of the GLP1 agonists, an effect not seen with GWAS SNPs. A common coding variant in MMP9 (Q279R, MAF 0.35), also not covered in GWAS platforms, was consistently associated with HDL-cholesterol (p=2.5e-5) and to a lesser extent triglyceride (p=0.006). A rare novel coding variant T258I also in MMP9 was found to be associated with higher interleukin levels (p=1.07e-6), supporting a finding in the initial study (p=3.4e-6), though not the same interleukin. While following up results identified from sequencing experiments will require large, well phenotyped cohorts, genotyping technologies capable of providing robust data for rare variants as well as analysis methods tailored to rarer variants, the insights obtained from full characterization of drug target genes and their associations with cardiovascular, metabolic and inflammatory phenotypes provide support for existing and future drug development efforts.

1687T

Rare Copy Number Variants in Congenital Heart Disease Affecting Gene Structure. *K. Dumas[1], Y. Bouhlal[1], M. McKeon[1], JTC. Shieh[1,2].* 1) Division of Medical Genetics, Department of Pediatrics, University of California San Francisco, San Francisco, CA, USA; 2) Institute for Human Genetics, University of California, San Francisco, San Francisco, CA, USA.

Developments in genome-scale technologies have greatly facilitated the discovery of copy number variants (CNVs) in the human genome. Some rare variants may constitute predisposing alleles to diseases of early age onset and may be associated with decreased reproductive fitness. To assess for the presence of such variants in individuals with congenital anomalies, we performed CNV analysis on eleven individuals affected with congenital heart disease, their phenotypically normal parents, and HapMap sample NA10851 using Affymetrix 6.0 genotyping arrays. Replicates of NA10851 revealed 100% concordance of CNV calls above an average of 85 markers. CNVs in the CHD trio samples were stringently called to decrease the likelihood of false positives. Rare CNVs were filtered from common variants using known databases then compared to Decipher. We found five rare CNVs including variants on 13q, Xq, and 10p. The latter two variants were not present in parents, while the rare 13q CNV was inherited. Of particular interest, we noted that some of the CNVs alter gene structure by interrupting exons. We cloned the precise CNV boundary and verified the presence of a novel transcript from 13q that crossed the CNV boundary using RT-PCR. Since this predicted early translation termination, we transfected GFP-tagged truncated or full-length cDNA constructs and found that the truncated protein was also stable in cell culture. These data demonstrate CNVs can create altered transcripts, and some truncated genes may be expressed. Studying these products may help in determining the phenotypic consequences of rare CNVs in congenital anomalies.

1688T

Rare variants in the *APOA5* promoter are associated with a paradoxical HDL-C decrease in response to fenofibric acid therapy. *A. Brautbar[1,2], M. Barbalic[3], F. Chen[3], J. Belmont[4], S. Virani[1], S. Scherer[4], R. Hegele[5], C. Ballantyne[1].* 1) Department of Medicine, Baylor Col Med, Houston, TX; 2) Department of Genetics, Marshfield clinic, Marshfield, WI; 3) Human Genetics Center, University of Texas Health Science Center at Houston, Houston, TX; 4) Department of Human Genetics, Baylor College of Medicine, Houston, Texas; 5) Robarts Research Institute, London ON Canada.

Background: The role of rare genetic variants in determining response to lipid modifying therapy is not fully understood. Individuals with mixed dyslipidemia are characterized by elevated triglycerides (TG), depressed high density lipoprotein cholesterol (HDL-C), and high risk for coronary events. Fenofibric acid (FA) is commonly used to treat mixed dyslipidemia by increasing HDL-C and apolipoprotein (apo) A1 (APOA1), decreasing TG, and possibly reducing coronary events. We sought to examine the effect of rare genetic variants in the *APOA5* gene region on plasma HDL-C, APOA1, and TG response to fenofibric acid (FA) and its combination with statins in individuals with mixed dyslipidemia. Methods: The *APOA5* gene region was sequenced in individuals (N=1612) with mixed dyslipidemia participating in a randomized, double-blind, active controlled, clinical trial examining the response to FA and its combination with statins. Student's t-test and rare variant burden tests were used to examine plasma HDL-C, APOA1 and TG response to therapy. Results: Instead of the expected increase in HDL-C and APOA1 following FA therapy, rare variants in the *APOA5* promoter region were associated with a paradoxical decrease in HDL-C and APOA1 levels in response to FA therapy, while rare missense variants were associated with better TG response to the statin-FA combination. Conclusion: Rare variants in the *APOA5* promoter region were associated with paradoxical decreases in HDL-C and APOA1 in response to FA therapy in the mixed dyslipidemia population, while missense variants were associated with improved TG response. Further study is indicated to examine the effect of these rare variants on coronary outcomes in individuals with mixed dyslipidemia receiving FA and its combination with statins.

1689T

***COL4A1* and *COL4A2* mutations cause genetically modifiable hemorrhagic stroke.** *M. Jeanne[1], J. Jorgensen[1], C. Labelle-Dumais[1], Y. Weng[1], W.B. Kauffman[1], M. de Leau[1], S.M. Greenberg[2], J. Rosand[2,3,4], J. Favor[5], D.B. Gould[1].* 1) Departments of Ophthalmology and Anatomy, Institute for Human Genetics, University of California San Francisco, San Francisco, CA; 2) Department of Neurology, Massachusetts General Hospital Boston, MA; 3) Center for Human Genetic Research, Massachusetts General Hospital Boston, MA; 4) Division of Neurocritical Care and Emergency Neurology, Massachusetts General Hospital Boston, MA; 5) Institute of Human Genetics, Helmholtz Zentrum Munchen, Neuherberg, Germany.

Patients with mutations in the collagen type IV alpha 1 gene (*COL4A1*) suffer from a broad spectrum of highly penetrant cerebrovascular diseases (CVDs), including subclinical cerebral microbleeds, leukoaraiosis, infantile porencephaly and fatal intracerebral hemorrhage (ICH). This phenotypic heterogeneity suggests that environmental and other genetic factors may influence pathology expressivity. COL4A1 associates with COL4A2 in heterotrimers before being secreted in the extracellular matrix where they form a network critical for basement membrane stability and function. We recently screened a cohort of patients with non-familial, spontaneous ICH. We identified novel *COL4A1* mutations and importantly, we discovered the first *COL4A2* mutations causing hemorrhagic stroke. Notably, our findings show that *COL4A1* and *COL4A2* mutations could contribute to 5–10% of non-traumatic, sporadic ICHs. We are now using cell biological and genetic approaches to decipher the pathogenic mechanisms underlying the disease. Using a cell-based assay, we found that mutations impair COL4A1 and COL4A2 secretion. We showed that mutant COL4A1 or COL4A2 proteins accumulate within cells where they titrate normal COL4A1 and COL4A2 and can ultimately trigger endoplasmic reticulum stress and the Unfolded Protein Response. Using cerebral magnetic resonance imaging and histological analysis, we determined that the hemorrhage localization and severity were dependent on the position and nature of the mutation in the *Col4a1* or *Col4a2* gene, and also on their genetic context. These findings suggest that allelic heterogeneity and the presence of genetic modifiers contribute to the CVD heterogeneity in human patients. Our ongoing experiments to better understand the cell biology of *COL4A1* and *COL4A2* mutations and the mechanisms of genetic modification could lead to targeted therapeutics to reduce the risk of CVD in patients with *COL4A1* or *COL4A2* mutations.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1690T

Association of 87 traits related to coronary heart disease and rare sequence variants in the ClinSeq™ Study. *H. Sung[1], B. Suktitipat[1], K. Lewis[2], D. Ng[2], S. Gonsalves[2], J.K. Teer[2], N.F. Hansen[3], J.C. Mullikin[3,4], L.G. Biesecker[2], A.F. Wilson[1], NISC Comparative Sequencing Program.* 1) Genometrics Section, Inherited Disease Research Branch, NHGRI, NIH, Baltimore, MD; 2) Genetic Disease Research Branch, NHGRI, NIH, Bethesda, MD; 3) Comparative Genomics Unit, Genome Technology Branch, NHGRI, NIH, Bethesda, MD; 4) National Institutes of Health Intramural Sequencing Center (NISC), NHGRI, Bethesda, MD.

ClinSeq™ is a large scale medical sequencing study designed to investigate associations of rare sequence variants with traits related to coronary heart disease (CHD). The study currently includes 976 non-smoking patients, ages 45 to 65, with normal to severe coronary artery calcification with both traits and whole exome data. More than 200 CHD-related traits were measured at the NIH Clinical Research Center in Bethesda, MD; 87 quantitative traits were selected for the first pass analyses. Whole exome sequencing with two different capture regions was performed at the NIH Intramural Sequencing Center; 387 and 325 individuals were sequenced with the Agilent SureSelect 38 mb and 50 mb capture regions, respectively. Sequence variants (SVs) in common over both capture regions with calling rates > 98% and minor allele frequency (MAF) > 1% were used to check for misspecified population stratification by multidimensional scaling analysis - 635 Caucasians were included in these analyses. The two groups of capture regions were merged, yielding 645,364 SVs. Of these SVs, 68% and 31% had MAFs < 0.01 and < 0.001, respectively. The SVs with MAF < 0.01 were collapsed into a single derived variant for each genomic region defined by hotspot blocks. Collapsed variants were coded as the proportion of the minor allele within each region; common variants were coded as the number of minor alleles (scaled from 0 to 1). Tests of association of each SV with each trait were performed on age- and sex-adjusted traits on the untransformed, log-transformed and Box and Cox transformed data with simple linear regression. Associations between 23 traits and at least one SV were significant over both untransformed and transformed traits after Bonferroni correction (p value < 7.78e-08). Traits and the number of significant SVs (indicated in parentheses) included: coronary calcium score (208), echo right atrium pressure (117), corrected QT interval (38), direct bilirubin level (19), total bilirubin level (18), echo left atrial dimension (16), echo AV peak velocity (11), glucose (10), QT interval (8), echo E/A ratio (8), head circumference (7), platelet count (6) and progesterone (3). Tiled regression implemented in TRAP will be performed to identify the set of independently significant SVs that affect each trait. The genes for the associations will be described at the meeting.

1691T

Exome Sequencing Identifies a Novel Splice Site Mutation in *LMNA* and Multigenic Heterozygosity of Potential Modifiers in a Multi-generation Family with Sick Sinus Syndrome, Dilated Cardiomyopathy and Sudden Cardiac Death. *M.V. Zaragoza, E. Jensen, F. Oh, C.K. Tran, V. Hoang, S. Hakim.* UC Irvine Cardiogenomics Program, Dept of Pediatrics, Division of Genetics & Metabolism, and Dept of Biological Chemistry, University of California, Irvine.

Sick Sinus Syndrome (SSS) is a cardiac arrhythmia caused by abnormal pacing of electrical activity via the sinus node; the defect causes an irregular heart beat leading to an increased risk of cardiac dysfunction (cardiomyopathy) and sudden death. The main goals of this study are to understand the primary genetic mechanism that causes SSS and to identify the secondary factors that may result in clinical variability within a large, multi-generation family. The proband is a 61 year-old male with SSS and a family history of individuals (>10) with atrial flutter, ventricular tachycardia, cardiomyopathy, heart failure and/or sudden death. After our initial Sanger sequencing studies for mutations were negative, we conducted exome sequencing to identify the primary mutation and potential modifiers. We used NimbleGen SeqCap enrichment and Illumina HiSeq paired-end sequencing. This produced 9 Gb of sequencing data of 97 million reads at an average length of 93 bases. Bioinformatics analysis using DNAnexus showed 94% of the reads mapped the reference (hg19) and 98% of the target exons (191,866 of 194,954) were completely sequenced. The average mean read coverage of each exon was 129X (range: 0 to 7108X), and 92% of the exons had mean coverage above 30X. We identified 128,563 unique variants with 108,795 (85%) located in 16,319 genes of 19,056 target genes. To identify candidate mutations, we focused on 2,077 variants located in 237 genes of 283 known arrhythmia, cardiomyopathy or ion channels genes. We filtered the candidates to 41 variants in 33 genes using zygosity, protein impact, database searches and clinical association. We found that half of these variants were real; only 20 variants were validated by Sanger sequencing. We then selected nine confirmed variants with minor allele frequencies (<1%) for extensive family studies of affected and unaffected individuals. These results identified *LMNA* c.357-2A>G, a novel heterozygous splice site mutation as the primary mutation and rare or novel heterozygous variants in *HCN4*, *MYBPC3*, *PKP4*, *TMPO*, *TTN*, *DMPK* and *KCNJ10* as potential modifiers. In conclusion, we used exome sequencing to identify a novel *LMNA* mutation and demonstrated the importance of Sanger validation and extensive family studies to evaluate candidate variants. We also discovered multiple variants in known arrhythmia, cardiomyopathy or ion channels genes that may serve as secondary factors in disease expression.

1692T

Whole-exome sequencing and quantitative analysis of genetic variants associated with blood pressure: the NHLBI Exome Sequencing Project. *S.L. Pulit[1,2], E.S.P. Blood Pressure Project Team[2], The NHLBI Exome Sequencing Project.* 1) Center for Human Genetics Research, Massachusetts General Hospital, Boston, MA, USA; 2) Medical and Population Genetics Program, The Broad Institute, Cambridge, MA, USA.

Background: Hypertension is a highly heritable trait that increases the risk of congestive heart failure, stroke, and other cardiovascular diseases. Rare mutations in salt-handling genes underlie Mendelian hypo- or hypertension syndromes and, for three of these genes, blood pressure (BP) variation at the population level [1]. Genome-wide association studies (GWAS) have implicated 45 common SNPs associated with BP, but these SNPs explain only 1% of the heritability of BP. Low-frequency and rare variants, thus far untested by GWAS, may contribute to the genetic architecture of BP. **Methods:** We exome-sequenced 6,823 European- and African-ancestry samples selected for heart and lung traits. Of these samples, 815 were selected from the top and bottom 1% of the BP distribution measured in 100,000 samples. Another 4,155 samples selected for other traits had available BP measures. We applied sample- and variant-level quality control (QC) including principal component analysis and filtering by Ti/Tv ratio. **Results:** We performed quantitative trait analyses of baseline systolic and diastolic BP (SBP and DBP), adjusted for age, sex, BMI, cohort, and ancestry, on all samples with these measures (N=4,769 after QC). BP measures were imputed in individuals on antihypertensive treatment. Variants with frequency >0.5% were tested individually using linear regression. No exonic variants achieved exome-wide significance ($p<10^{-6}$). The top exonic variant was a missense SNP in *RGS22* (chr8:101011612, MAF=0.008, $p=1.08\times10^{-6}$) associated with DBP. We used permutation-based burden tests to test the aggregate effects of rare variants. No genes achieved significance after multiple-test correction for 18,000 genes ($p<3\times10^{-6}$). Further, we did not replicate 3 genes previously shown to contain rare variants associated with BP (p>0.05). **Conclusions:** Analysis of baseline SBP and DBP using whole-exome data from 4,769 samples failed to identify single mutations or genes influencing BP variation. Our current results could reflect limited power due to sample size [2] or phenotypic variability. To improve power to identify a potential role of rare mutations in BP variation, analysis of long term average BP and imputation in >10,000 individuals with GWAS data are underway. **Citations:** [1] Ji, W., et al. Nat Genetics. 40 (592-9). May 2008. [2] Kryukov, G. et al. PNAS. 106 (3871-6). Feb 2009.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1693T

Identification of Novel Genes and Their Mutations Associated with Hypertrophic Cardiomyopathy. *X. Liu, Q. Huang, X. Dai, W. Xie.* Institute of Life Sciences, Nanjing, Jiangsu, China.

Hypertrophic cardiomyopathy (HCM) is a common form of genetic heart muscle disease that occurs in 1 out of 500 people and is also the leading cause of sudden cardiac death in young athletes. HCM is inherited as an autosomal dominant trait and is attributed to mutations in genes that encode for the sarcomere proteins. About 50–60% of patients with a high index of clinical suspicion for HCM will have a mutation identified in at least 1 of 9 sarcomeric genes. So far mutations of over 30 genes are can cause HCM. There are around 20% of HCM without known genetic causes. To identifiy novel HCM associated genes and their mutations, we performed exon sequencing of 20 Chinese HCM samples without mutations in two of the most common HCM genes (MYH7 and MYBPC3). Three of 20 samples did not have any mutation in known HCM associated genes. To further analyze these three samples, we performed a series of filtration analysis to exclude all the common polymorphism and non-coding variants, we retained 30–240 variants per sample. Combined with functional relevance analysis, we identified three novel HCM associated genes for functional study and population analysis.

1694T

Linkage Analysis and Exome Sequencing to Identify Causal Variants for Familial Combined Hyperlipidemia. *A. Mak, C.R Pullinger, M.J. Malloy, P.L.F. Tang, R. Deo, P.Y. Kwok.* Cardiovascular Research Institute, University of California San Francisco, San Francisco, CA.

Familial combined hyperlipidemia (FCH), the most prevalent hyperlipidemia, is a complex metabolic disorder characterized by variable occurrence of high low-density lipoprotein (LDL) and high triglycerides (TG) - a condition that is commonly associated with coronary artery disease (CAD). Over the years, numerous studies have contributed to the elucidation of the genetic basis of FCH, which is still largely unknown. The need to search for additional genetic variants remains due to the highly variable and complex genetic profile specific to each affected family.

Exome sequencing provides a fast and comprehensive study of genetic variants that are associated with FCH. The large number of genetic variants typically detected in a human genome can, however, be overwhelming and makes it difficult to differentiate disease-causing variants from common variants that are not disease-associated. Linkage analysis greatly reduces the search space for disease-causing variants by pinpointing the most-likely chromosomal regions that are associated with disease phenotypes in the family. Using Affymetrix Axiom Custom array, we genotyped 42 individuals from a 60-member family with history of FCH and low levels of high-density lipoprotein (HDL) cholesterol. We identified various chromosomal regions associated with low HDL, high LDL, high TG and obesity in this family. To identify the exact causal variants that were associated with these traits, we performed exome sequencing on five representative family members using the Illumina HiSeq2000 platform with the TruSeq exome enrichment protocol.

The exome data obtained reached a high mean target coverage of 66X with 10X coverage over 90% of the targeted regions. A union set of 35739 exonic / splicing single nucleotide variants and indels were found in the five family members. Among these variants, 1263 non-synonymous, frameshift or stop-gain/loss variants were identified in one or more linkage peaks associated with various lipid phenotypes in the family. This includes variants found in 22 genes related to lipid metabolism and 256 novel variants that possibly represent rare variants responsible for the FCH phenotype. Verification of potential causal variants is underway. The findings from this study will contribute to the understanding of the genetic basis of FCH.

1695T

Exome Sequencing Reveals a Truncation Mutation in the Novex-3 Isoform of TTN that Disrupts Cardiomyocyte Ultrastructure and Segregates with Familial Dilated Cardiomyopathy. *J.L. Theis[1], V.V. Michels[2], S. Middha[3], S. Baheti[3], P.C. Abell Aleff[4], T.M. Olson[1].* 1) Cardiovascular Genetics Laboratory, Mayo Clinic, Rochester, MN; 2) Department of Medical Genetics, Mayo Clinic, Rochester, MN; 3) Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, MN; 4) Department of Biochemistry and Molecular Biology, Mayo Clinic, Rochester, MN.

Dilated cardiomyopathy (DCM) is a heritable, genetically heterogeneous disorder characterized by progressive heart failure. The molecular and cellular pathobiology of DCM remains unknown in a majority of cases. Massively parallel sequencing and comparative variant filtering in well-characterized families enable discovery of novel, unsuspected genetic underpinnings of DCM.

Echocardiography in 15 members of a three-generation family revealed 6 with DCM, which segregated as an autosomal dominant trait. Whole-exome sequencing was performed in 3 affected 2nd–3rd degree relatives using the Agilent Sure Select Capture Kit (50 Mb) and the Illumina HiSeq 2000. An iterative bioinformatics process was used to filter >40,000 annotated SNVs (single nucleotide variants) and INDELs (insertion-deletion variants) for novel, non-synonymous, shared variants, culminating in identification of 7 candidates. A sole heterozygous mutation in TTN, encoding the giant protein titin, segregated with DCM. The mutation deleted a conserved base pair within an alternatively-spliced exon unique to the novex-3 isoform of titin. The N-terminal 3454 amino acids of the novex-3 isoform are 97% identical to the N2A and N2B isoforms, but the C-terminus is made up of a single distinct exon encoding 2150 amino acids. The resultant frameshift mutation truncated 758 amino acids of the 5604-residue full-length protein, causing loss of an obscurin binding domain as well as two phosphorylation sites. Electron microscopy of cardiac biopsy tissue from an affected family member revealed marked myofibrillar dissolution in multiple cardiomyocytes with complete absence of the classic sarcomeric structure.

Comparative analysis of exome sequences from non-nuclear affected family members is a powerful strategy to discover pathogenic mutations. While truncating mutations in *TTN* are relatively common in DCM, none have been reported to disrupt novex-3, a short isoform thought to play a minor role in myocardial physiology. Moreover, the impact of *TTN* mutations on cardiomyocyte ultrastructure in human DCM has not been described. Our findings implicate a central role for the unique novex-3 isoform in sarcomere assembly and stability. Genetic testing platforms for DCM should include the novex-3 encoding exon of *TTN*.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1696T

Challenges in interpreting secondary variants from massively parallel sequencing, perspectives from the ClinSeq[™ study. *D. Ng[1], J.J. Johnston[1], K.L. Lewis[1], S.G. Gonsalves[1], L.N. Singh[1], L.C. Peller[1], J.K. Teer[1,2], J.C. Mullikin[2,3], L.G. Biesecker[1,2].* 1) Genetic Disease Research Branch, NHGRI, NIH, Bethesda, MD; 2) NIH Intramural Sequencing Center, NHGRI, NIH, Bethesda, MD; 3) Genome Technology Branch, NHGRI, NIH, Bethesda, MD.

Massively parallel sequencing (MPS) has been used successfully to identify the genetic cause of rare disease (primary findings). MPS also generates incidental disease susceptibility variants (secondary findings). Interpreting secondary findings in asymptomatic individuals is a challenge as there is no consensus on the analysis and return of incidental genetic results. As part of our ongoing endeavor to study secondary variants, we selected 22 arrhythmia-associated and 41 cardiomyopathy-associated genes for analysis and present our findings. Methods: 572 whole exome sequences were annotated for variants in 63 genes with an algorithm that filtered results based on genotype quality, allele frequency, mutation type, and information in locus-specific databases. Variants were assigned pathologic scores ranging from 0–5 (0=poor genotype quality, 1=not pathogenic, 2=likely not pathogenic, 3=uncertain, 4=likely pathogenic, 5=pathogenic). Results: Seven arrhythmia-associated variants were designated likely pathogenic and three variants were pathogenic (KCNE1 p.Thr10Met, p.Arg98Trp; KCNH2 p.Arg3-12Cys). Two individuals had prolonged QTc interval (male with SCN4B p.Ser206Leu and female with SCN3B p.Leu10Pro). Three participants had relatives who died suddenly and one had a relative who died in infancy of unknown cause. There were 20 likely pathogenic and 3 definitely pathogenic (MYBPC3 IVS+1G>A; MYH7 p.Arg787Cys; PLN p.Leu39X) cardiomyopathy-associated variants. Among these 23 participants, 16 had abnormal ECG (sinus bradycardia n=13, premature ventricular contractions n=2, supraventricular contractions n=1). Echocardiogram showed one participant with a borderline thickened septum. Family history showed relatives with cardiomyopathy (n=1), congestive heart failure (n=5) and cardiac arrhythmia (n=11). Conclusion: We identified 6 asymptomatic individuals with pathogenic variants associated with Long QT syndrome (LQTS n=3), hypertrophic cardiomyopathy (HCM n=2) and dilated cardiomyopathy (DCM n=1). The incidence of these cardiac-associated variants in the ClinSeqTM cohort exceeds the estimated population frequency of LQTS (1/2500), HCM (1/500) and DCM (1/7500). The elevated incidence may be due to self-selection, false positives, or underestimation of the incidence of these disease phenotypes. Further family and functional studies are underway to address the pathogenicity, penetrance and medical management of secondary variants identified through MPS.

1697T

An intronic mutation affecting pre-mRNA splicing in the COL3A1 gene as novel mechanisms causing vascular Ehlres-Danlos syndrome. *A. Watanabe[1,2], B.T. Naing[1], M. Sasaki[2], K. Akutsu[3], S. Kosaihira[4], A. Gemma[4], T. Shimada[1,2].* 1) Dept Biochem & Molec Biol, Nippon Med Sch, Tokyo, Japan; 2) Div Clinical Genetics, Nippon Med Sch Hosp, Tokyo, JAPAN;; 3) Div Intensive and Cardiac Care Unit, Nippon Med Sch Hosp, Tokyo, Japan; 4) Div Pulmonary med, Infectious dis, and Oncology, Dept Internal Med, Nippon Med Sch, Tokyo,.

Vascular type of Ehlers-Danlos syndrome (vEDS), also known as EDS type IV (MIM#130050), is a life-threatening autosomal dominant inherited disorder connective tissue mainly caused by mutations in type III collagen, *COL3A1* gene. vEDS is usually caused by mutations in single amino acid substitutions for glycine in the GLY-X-Y repeat of the triple helical region or invariant splice sites. Recently, premature translation termination mutations are reported affecting nonsense-mediated mRNA decay of mutated allele. Here, we described a 35 y/o male with clinical features and family history of vEDS and characterized the functional consequences of an intronic *COL3A1* mutation. The coding regions and flanking intronic bases of the *COL3A1* gene were screened by high resolution melting curve method using genomic DNA. The patient is a heterozygote for a variant with IVS15 -10 A>G in genomic DNA. By RNA/cDNA sequencing analysis from fibroblast, this variant generated a new acceptor site within intron 15, causing the insertion of a 9-bp intronic sequence between exon 15 and 16 (c.996_997insTTTTTTCAG), and resulting in-frame three amino acids insertion. Consistent with this finding, minigene splicing assays demonstrated that this variant is sufficient to create a cryptic splice site and cause the observed 9-bp insertion. This is the first report of creation of a splice site by an intronic single nucleotide change in vEDS. Intronic mutations may be responsible for vEDS in some families with otherwise negative mutation screening of *COL3A1*.

1698T

Impaired basement membrane function causes dilated cardiomyopathy in Marfan syndrome. *J.R. Cook[1], L. Carta[1], L. Benard[1], E.R. Chemaly[1], E. Chiu[1], T. Hampton[2], N. Clayton[3], C. Nelson[3], B. Wentworth[3], R.J. Hajjar[1], F. Ramirez[1].* 1) Pharmacology & Systems Therapeutics, Mount Sinai School of Medicine, New York, NY; 2) Neuroscience Discovery Core, Mouse Specifics, Inc., Framingham, MA; 3) Genzyme Corporation, Framingham, MA.

Contrary to the general view that cardiomyopathy in Marfan syndrome (MFS) is a secondary complication of arterial and/or valvular disease, some clinical findings have suggested intrinsic myocardial insufficiency. Our characterization of mice with early onset of progressively severe MFS ($Fbn1^{mgR/mgR}$ mice) supports this hypothesis. First, histological and physiological analyses demonstrated that $Fbn1^{mgR/mgR}$ mice display dilated cardiomyopathy (DCM) prior to the emergence of aortic aneurysm and valvular regurgitation. Second, conditional Fbn1 inactivation in cardiomyocytes, but not in the ascending aorta or cardiac valves, was associated with DCM. Even though fibrillin-1 regulates TGFβ bioavailability, additional experiments excluded involvement of promiscuous TGFβ activity in DCM progression by showing unremarkable Smad2/3 signaling in fibrillin-1 deficient hearts and unabated myocardial dysfunction in $Fbn1^{mgR/mgR}$ mice systemically treated with the pan-TGFβ neutralizing antibody 1D11. Furthermore, normal cardiac size and function were noted in $Fbn1^{mgR/mgR}$ mice that were treated with the angiotensin type I (AT1) receptor antagonist losartan or those lacking the AT1 receptor. Collectively, the data imply that basement membrane-associated fibrillin-1 assemblies protect the mammalian heart against DCM by providing structural support for cardiomyocyte contraction. More generally, our study suggests that monitoring myocardial insufficiency should become standard care in MFS and that losartan treatment is an effective therapy to restore cardiac function in afflicted patients.

1699T

Oxidative stress as a modifier of cardiovascular disease in a mouse model of Williams Beuren Syndrome. *B. Kozel[1], L. Ye[1], R. Knutsen[2].* 1) Gen & Genomic Med, Washington Univ Sch Med, St. Louis, MO; 2) Cell Biology and Physiology, Washington Univ Sch Med, St Louis, MO.

Introduction: Elastin haploinsufficiency causes the cardiovascular complications associated with Williams-Beuren syndrome (WBS) and isolated supravalvular aortic stenosis. However, severity of vascular disease ranges broadly in this population. Using the Eln+/− mouse, we sought to identify the source of this variability. Previous investigation showed that crossing the C57Bl/6 Eln+/− to 129×1/SvJ increased hypertension and decreased vascular compliance. Conversely, the DBA/2J-crossed Eln+/−s were protected from the pathologic changes associated with elastin insufficiency. Methods/Results: Quantitative trait locus (QTL) analysis performed on F2 intercrosses revealed a largely additive effect of the 129×1 background on the severity of Eln+/− vascular disease. Highly significant peaks for blood pressure (BP) or vascular stiffness (VS) were identified on chromosomes 1 (BP, LOD 9.7), 9 (VS, LOD 8.7), and 5 (BP and VS, LOD 4.5). In the C57 Eln+/−×DBA F2s, the modifier effect was epistatic. Peaks on chromosomes 14 (VS, LOD 4.7) and 5 (BP, LOD 7.6) were identified only in the Eln+/− and a peak on chromosome 2 (VS, LOD 5.6) mainly affected Eln+/+s. Bioinformatic analysis of the QTL peaks revealed multiple genes affecting oxidative stress, including both the NADPH oxidase (NOX) and nitric oxidase synthase (NOS) pathways. Preliminary Eln+/−; Ncf1+/− (a member of the NOX2 complex) and Eln+/−; Nos1+/− double mutant studies show opposite modifying effects on elastin mediated vascular disease, with Ncf1 haploinsufficiency being protective and Nos1 deficiency causing more severe disease. Conclusions: QTL analysis successfully identified genes that modulate the severity of vascular disease in elastin insufficient mice. Follow-up studies with elastin modifier double mutants suggest that the balance and type of reactive oxygen species generated markedly affect the vascular phenotype of these animals. These initial results are promising and may suggest rational therapies for the treatment of vascular disease in individuals with WBS.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1700T

The National Registry of Genetically Triggered Thoracic Aortic Aneurysms (GenTAC): registry progress and research successes. *C. Maslen, GenTAC Consortium.* Cardiovascular Med, Oregon Hlth & Sci Univ, Portland, OR.

Funded by the National Institutes of Health, the GenTAC Registry is a multicenter, longitudinal, observational cohort study of patients with conditions involving genetically triggered thoracic aortic aneurysm and/or dissection (TAAD). GenTAC was established to provide a biospecimen inventory and bioinformatics infrastructure that will enable research to advance the clinical management of such patients. Primary diagnoses include Marfan syndrome, bicuspid valve with aneurysm and/or a family history of aneurysm, idiopathic TAAD in patients < 50 years of age, Turner syndrome, familial TAAD, other congenital heart disease with TAAD, Ehlers-Danlos syndrome, and Loeys-Dietz syndrome. To date GenTAC has recruited 2826 subjects, with a final goal of 4000 registrants. Initial analyses of GenTAC data/specimens have included studies of genetic causes for aortic syndromes via gene sequencing, SNP, CNV and genome-wide association studies, potential usefulness of TGF-β blood levels as prognostic or therapeutic marker in Marfan subjects, surgical approaches/outcomes for ascending aortic conditions and gender differences in TAAD. Other studies in progress include cross-sectional and longitudinal data regarding phenotype-genotype correlations of disease risk factors, features, treatment, and outcomes, and analysis of imaging methods/integration of imaging findings with clinical and genetic data. GenTAC phenotyping and imaging cores have been established to facilitate these inquiries. The GenTAC Registry is a resource for anyone in the scientific community interested in advancing our understanding of genetically mediated TAAs and their causes, diagnosis, and optimal treatment. Investigators interested in utilizing GenTAC for ancillary studies should contact the registry at gentac-registry@rti.org or apply at http://gentac.rti.org.

1701T

Association of CRP gene variation in Asian Indian Takayasu's arteritis patients. *K. Shah[1], D. Danda[2], R. Shah[2], S. Prasanna[3], P. Chandna[4], R. Chopra[4], S. Danda[1].* 1) Clinical Genetics Unit, Christian Medical College, Vellore, India; 2) Department of Clinical Immunology and Rheumatology,Christian Medical College, Vellore, India; 3) Department of Biostatistics, Christian Medical College, Vellore, India; 4) AceProbe Technologies (India) Pvt. Ltd., B-8, Namdhari Chambers, Karol Bagh, NewDelhi, India.

BACKGROUND: Takayasu's arteritis (TA) is a rare, systemic inflammatory disease, typically involving the aorta and its main branches where cell-mediated autoimmunity has been strongly implicated in its pathogenesis. CRP gene polymorphisms are associated with serum C-reactive protein (CRP) concentrations and may play a role in pathogenesis of Takayasu's arteritis. The association of CRP gene polymorphism with Takayasu's arteritis is being reported for the first time. METHODS: SNP genotyping was done with Sequenom iPlexTM Gold Assay and MALDI-TOF platform for 54 TA patients and 100 controls (Indian population). Four single nucleotide polymorphisms (SNPs) from the CRP gene (rs1205, rs1417938, rs3093058 and rs1800947) were evaluated. Single nucleotide polymorphisms in MCP-1, IL-23R, MMP-2, MMP-9, IL-10, IL-18, IL-12RB2, CD-24 and ICAM-1 were also investigated. Statistical analysis was done using the PLINK software. RESULTS: Detected frequencies of heterozygous genotype rs1205 were 53% (53/100) in control group and 22.3% (11/54) in TA group (P = 0.0001, odds ratio: 0.2, 95% CI: 0.09, .45). rs1417938, rs1800947 and rs3093059 genotyping did not show significant association. Rest of the studied polymormphisms also did not show significant association with takayasu's arteritis. CONCLUSIONS: This is a pilot study from India, first time looking at association of various gene polymorphisms in takayasu's arteritis. A significant association between heterozygous genotype rs1205 in the CRP gene and TA was revealed which decreased the risk of TA. No significant association of rs1417938, rs3093058, and rs1800947 polymorphisms in the CRP gene was found in our study population. The current results suggest that CRP polymorphisms may play a major role in pathogenesis of TA patients. Our results for the first time show an association between CRP gene polymorphisms and TA.

1702T

Cardiomyopathy as an emerging phenotype in Hereditary Inclusion Body Myopathy (HIBM). *S. Siebel[1], S. McGehee[2], A. Brofferio[2], C. Ciccone[1], M. Huizing[1], L. Medne[3], R. Finkel[3], J. McKew[4], W.A. Gahl[1], N. Carrillo-Carrasco[4].* 1) National Human Genome Research Institute, NIH, Bethesda, MD; 2) National Heart, Lung and Blood Institute, NIH, Bethesda, MD; 3) Children's Hospital of Philadelphia, Philadelphia, PA; 4) Therapeutics for Rare and Neglected Diseases (TRND), National Center for Advancing Translational Sciences (NCATS), NIH, Bethesda, MD.

Background: HIBM is an autosomal recessive neuromuscular disorder presenting during early adulthood with slowly progressive muscle weakness and atrophy. It is caused by GNE mutations in either the epimerase or kinase domains of this enzyme of sialic acid biosynthesis. More than 60, mainly missense, mutations have been described in different ethnic groups. The complete phenotypic characterization of HIBM is ongoing. **Case report**: We describe a 37-year-old patient with HIBM and cardiomyopathy, evaluated under the Natural History Study of HIBM (CT.gov: NCT01417533). She is a compound heterozygote for missense mutations in exon 4 and exon 11 (p.V216A/p.A631V), affecting both the epimerase and kinase domains. Prior to her evaluation she denied any cardiac symptoms, but screening cardiac evaluation detected an arrhythmia with Holter monitoring; an echocardiogram showed global left ventricular hypokinesis with a reduced ejection fraction of 40%. A follow-up echocardiogram, while the patient was managed with a low-dose AT1 subtype angiotensin II receptor antagonist, showed an ejection fraction of 46%, global left ventricular hypokinesis and paradoxical septal motion. Cardiac stress MRI showed no evidence of ischemic disease. Her 29 year-old sister diagnosed with HIBM had a normal echocardiogram. **Conclusion:** We describe cardiomyopathy in a patient with HIBM and no other apparent predisposition to heart disease. The association of cardiomyopathy and HIBM remains to be established. The only other case report of dilated cardiomyopathy in HIBM was recently described in two siblings who were compound heterozygous for mutations in the kinase domain (p.F528C/A631V) (Chai, 2011). One of the mutations, p. A631V, was present in our current patient. In a recent study, cardiac myocytes derived from murine Gne-/- KO embryonic stem cells showed early degradation and a rapid decrease in their beating capacity, suggesting GNE may be an important factor in the development of cardiac tissue (Krentsis, 2011). Our case further supports the association between cardiomyopathy and HIBM, raises the possibility of a genotype-phenotype correlation and highlights the evolving and variable phenotype of HIBM. These data are encouraging to continue the systematic clinical evaluation and follow up of this patient group and the study of the role of GNE in cardiac development.

1703T

Role of the Pro12Ala, G972R, G1057D polymorphism of the PPAR-γ 2, IRS1 and IRS2 genes in type 2 diabetes and Coronary Artery Disease. *S. Vats, K. Matharoo, V. sambyal, A.J.S Bhanwer.* Human Genetics, Guru Nanak Dev University, Amritsar, India.

Diabetes mellitus (DM) is a metabolic disorder of multiple aetiology characterized by chronic hyperglycemia with disturbances ofcarbohydrate,fat and lipid metabolism resulting from defects in insulin secretion, insulin action or both. Diabetic subjects have 2 to 4 fold increased risk of the most common and life threatening cardiovascular complication; coronary artery disease (CAD) as compared to non diabetic subjects. As insulin resistance characterizes T2D, obesity and cardiovascular diseases, genes involved in the insulin resistance pathway can be the important candidate genes for CAD also. Insulin receptor substrates (IRS) are the important mediators of insulin signalling and have an important role in insulin resistance. Insulin stimulated tyrosine phosphorylation of insulin receptor substrate 1(IRS1) and insulin receptor substrate 2 (IRS2) is promoted by Peroxisome proliferator-activated receptor-gamma (PPARγ), hence all of them plays a major role in insulin resistance. In the present study, association of PPARγ Pro12Ala, G972R of IRS1 and G1057D of IRS2 polymorphisms in T2D and CAD patients of Punjabi population (North India) was investigated in 444 subjects divided into four groups: 122 healthy controls, 103 T2D+CAD, 117 T2D only and 102 CAD only patients. Binary logistic regression was applied to test the association of risk factors with T2D, T2D+CAD and CAD only after adjusting for age, sex and BMI. The minor allele (G) frequency for Pro12ala polymorphism was observed to be 11.5%, 12%, 6.9% and 4.6% in controls, T2D+CAD, CAD only and T2D only group respectively. In case of G972R polymorphism the minor allele (A) frequency for G972R polymorphism was observed to be 4.9%, 3.4%, 5.9 % and 4.3% in controls, T2D+CAD, CAD only and T2D only group respectively. 33.2%, 34.0%, 34.3% and 32.9% was found out to be the frequency of D allele in controls, T2D+CAD, CAD only and T2D group respectively. In case of Pro12Ala statistically significant difference ($p<0.05$) was observed in the allele frequencies of T2D patients with controls (Odds ratio : 3.014 (1.38–6.96) The allelic and genotypic frequencies of G972R and G1057D were found to be nonsignificant in all the groups($p>0.05$). For Pro12Ala, CC homozygotes appeared more likely to develop T2D than GG (Ala) carriers. Thus G allele may be protective for T2D patients of Punjabi population in the absence of CAD.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1704T

Genetic causes of cardiomyopathy in pediatric transplant recipients. *L. Zahavich, A. Al-Maawali, A. Dipchand, C. Manlhiot, S. Bowdin.* The Hospital for Sick Children, Toronto, Canada.

Background: Cardiomyopathies are a heterogeneous group of disorders. There are many different etiologies including myocarditis, familial isolated cardiomyopathy, neuromuscular disorders, inborn errors of metabolism and malformation syndromes. In children, cardiomyopathy is associated with a high degree of mortality and morbidity. Death within 2 years or requirement for heart transplant is seen in about 40% of children. Pre transplant risk factors and different etiologies have a major impact on prognosis. With advances in genetic investigations and availability of testing, the number of patients with genetic diagnoses will continue to expand. However, underlying genetic diagnoses may impact patient eligibility for heart transplantation. Objectives: To describe the genetic aetiology of cardiomyopathy in pediatric heart transplant recipients. Methods: We retrospectively reviewed the clinical phenotypes and results of cardiomyopathy gene panel testing in pediatric heart transplant recipients with cardiomyopathy, during the period of 2001–2012. Results: A total of 92 children have received heart transplants at the Hospital for Sick Children. Of these, 34 (36.9%) were excluded due death prior to 2001, transfer to another centre, anthracycline- related cardiomyopathy or other non-genetic etiology. Of the 58 remaining patients, 4 (6.9%) were diagnosed with hypertrophic cardiomyopathy, 42 (72.4%) with dilated cardiomyopathy, 10 (17.2%) with restrictive cardiomyopathy and 2 (3.4%) with other. A total of 20 (34.5%) patients had not undergone genetic testing to date and the results of testing for 3 (5.1%) patients were pending. Of the 35 who underwent genetic testing, 2 (5.7%) were diagnosed with a metabolic disease, 5 (14.3%) had a neuromuscular disorder and 8 (22.9%) patients were found to have a mutation in one of the 8 sarcomeric genes. As well, 8 (22.8%) of patients had variants of unknown significance and 12 (34.3%) had negative cardiomyopathy panels. Conclusions: Preliminary results suggest that about one third of our cohort of pediatric heart transplant patients have “idiopathic” cardiomyopathy. However, with the rapid increase in knowledge regarding genetic causes of cardiomyopathy, this number is likely to decrease. At present, all patients with confirmed genetic etiology of their cardiomyopathy are living, however, studies are ongoing to assess the clinical course and outcomes in transplant recipients with genetically caused cardiomyopathies.

1705T

Is your mother or father's family history more important when assessing your risk of having high blood pressure: Studies in humans and mice. *C.L. Chiu, C.T. Morgan, S.J. Lupton, S. Lujic, J.M. Lind.* University of Western Sydney, School of Medicine, Australia.

Background High blood pressure (HBP) is caused by a combination of environmental and genetic factors. A family history of disease is known to increase an individual's risk of having HBP. The contribution of paternal versus maternal effect is poorly understood. We aimed to determine the contribution of maternal versus paternal family history of HBP towards the risk for having HBP in humans. Animal models were also used to study parent of origin influences of angiogenic molecules implicated in HBP. **Methods** Self-reported data from 220,711 individuals recruited from the 45 and Up Study, Australia, were included in the study. Family history was stratified according to whether your father, mother, both, or neither parents had HBP. Relative risks (RR) and 99% confidence intervals were estimated using generalised linear models, adjusting for demographic and lifestyle factors. Total RNA was extracted from heart tissue of two homozygous mouse strains, A/J (AJ) and 129×1/SvJ (129) and the F1 of their reciprocal cross (n=8 males and n=8 females per group). Expression of *Vegfa* and *Hif1a* were measured using qPCR. GLM was used to compare expression differences within both the parental and F1 hybrid groups. **Results** Paternal and maternal family history of HBP was associated with an elevated risk of having HBP, when compared to no family history (RR, 99% CI: paternal 1.83, 1.79–1.87; maternal 1.79, 1.75–1.83; both 2.42, 2.36–2.47). Paternal family history of HBP was associated with a significantly increased risk of having HBP for males, compared to maternal family history (1.20, 1.14–1.27). This association was not observed in females. Within mice, *Vegfa* expression was 1.45 fold higher in male AJ compared to male 129 parentals. *Vegfa* experssion was 1.72 fold higher in F1 males with an AJ father, compared with F1 males with a 129 father; and 1.32 fold higher in females with an AJ father compared with females with a 129 father. *Hif1a* expression was 1.35 fold higher in male AJ compared to male 129 parentals, and 1.33 fold higher in male F1 with an AJ father compared to male F1 with a 129 father. **Conclusions** Males with a paternal family history had 20% increased risk of having HBP compared to males with a maternal family history. In mice, genes involved in angiogenesis that were elevated in male parents were shown to be elevated in the F1 offspring of these males. The sex of the patient and parent should be considered when assessing an individual's risk of having HBP.

1706T

Analysis of Variations from Exome Sequence Data with Subclinical Cardiovascular Disease in the Diabetes Heart Study. *J. Adams[1,2,3], A. Cox[2,3,4], B. Freedman[5], J. Carr[6], D. Bowden[2,3,4].* 1) Molecular Genetics and Genomics, Wake Forest School of Medicine, Winston Salem, NC; 2) Center for Genomics and Personalized Medicine Research, Wake Forest School of Medicine, Winston-Salem, NC; 3) Center for Diabetes Research, Wake Forest School of Medicine, Winston-Salem, NC; 4) Department of Biochemistry, Wake Forest School of Medicine, Winston-Salem, NC; 5) Department of Internal Medicine - Nephrology, Wake Forest School of Medicine, Winston-Salem, NC; 6) Department of Radiologic Sciences, Wake Forest School of Medicine, Winston-Salem, NC.

An estimated 8% of the US population is affected by diabetes with 95% of these cases being type 2 diabetes (T2D). Diabetes is a major risk factor for cardiovascular disease (CVD). Up to 65% of T2D individuals die due to vascular disease events. Identification of heritable risk factors for CVD and other complications of T2D are important to help understand individual disease risk. A recent GWAS in the CHARGE consortium identified loci associated with coronary artery calcification (CAC) and myocardial infarction. A meta-analysis of GWAS also from CHARGE identified several loci associated with carotid intima media thickness (IMT) and plaque. CAC and IMT are clinically relevant measures of subclinical CVD. In this study SNPs from exome sequencing were selected in the top hit genes from the CHARGE GWAS and meta-analysis using the NHLBI ESP-GO database. The aim of this study was to examine the association of rare coding SNPs with measures of subclinical CVD in a sample enriched for T2D. 28 SNPs from 14 genes were determined in 1208 European Americans from 474 families in the Diabetes Heart Study. Measures of subclinical CVD included CAC, carotid artery and abdominal aortic calcified plaque (CarCP and AaCP respectively), determined by CT scanning, and carotid IMT, determined by B-mode ultrasound. Additional risk factors including blood lipid levels were also analyzed. Average minor allele frequency for the genotyped SNPs was 0.030 (0.00040–0.16). Three SNPs were significantly associated with CVD risk factors: rs3135506 (Ser19Trp) in *APOA5* was associated with increased triglycerides (TG) (p=5×10-5), decreased concentrations of LDL (p= 0.00070), and HDL (p=0.0054). Also in the *APOA5*, rs651821 (5' UTR), was associated with increased TG concentration (p=0.00080) where as rs13832449 (splice donor) in *APOC3* was associated with decreased TG concentrations (p=0.0015). In addition, several other SNPs including rs45456595 (*CDKN2A*, Gly63Arg), rs5128 (*APOC3*, 3' utr), and rs 72650673 (*SH2B3*, Gly400Lys) were nominally associated (at p<0.010) with history of CVD, subclinical CVD, or CVD risk factors. While we were not able to replicate the associations with CAC and IMT, we did find that the genes identified in the CHARGE consortium GWAS and meta-analysis contain additional variants associated with CVD risk factors in individuals with T2D.

1707T

Common familial effects on ischemic stroke and myocardial infarction: A Swedish national cohort study. *K. Kasiman[1,2], C. Lundholm[1], S. Sandin[1], N. Malki[1], P. Sparén[1], E. Ingelsson[1].* 1) Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden; 2) Saw Swee Hock School of Public Health, National University of Singapore, Singapore.

Introduction: Recent genome-wide association analyses suggest some overlap of genetic determinants of ischemic stroke (IS) and myocardial infarction (MI). This study aimed to assess shared familial risk between IS and MI in a large population-wide cohort study. Methods: Study participants free of IS and MI were identified through linkage with their affected siblings extracted from the Swedish Hospital Discharge and Cause of Death Registers between 1987 and 2007, forming an exposed sib-pair. The Swedish Multi-Generation Register was used to match each exposed sib-pair, by birth year of both siblings and calendar period, to up to five unexposed sib-pair. Stratified Cox regression analyses were used to assess familial risk of incident MI and IS in those exposed to having a sibling with IS and MI, respectively, compared to unexposed. Results: A total of 31,659 exposed and 143,728 unexposed to siblings with IS, and a total of 62,766 exposed and 265,974 unexposed to siblings with MI were included in the analyses. The overall risk of incident MI when exposed to having a sibling with IS was significantly increased (RR, 1.44; 95% CI, 1.34–1.55, p<0.001) to a similar extent to risk of incident IS when exposed to having a sibling with MI (RR, 1.41; 95% CI, 1.32–1.50, p<0.001). The familial risks were higher among full siblings in both groups (RR for MI, 1.46; 95% CI, 1.35–1.58, p<0.001; and RR for IS, 1.40; 95% CI, 1.30–1.40, p<0.001) compared to half siblings (RR for MI, 1.29; 95% CI, 1.05–1.59, p<0.001; and RR for IS, 1.38; 95% CI, 1.16–1.65, p<0.001). Conclusion: Our large, population-wide study indicates that there is a considerable overlap of familial risk between ischemic stroke and myocardial infarction. The sources of this overlap should be evaluated using other study designs.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1708T

Jervell and Lange-Nielson Syndrome: Homozygous missense mutation in KCNQ1 in a Turkish family. *S. Temel[1], O. Bostan[2], H. Cangul[3], E. Cil[2].* 1) Department of Medical Genetics, Faculty of Medicine, University of Uludag, Bursa, Turkey; 2) Department of Pediatric Cardiology, Faculty of Medicine, University of Uludag, Bursa, Turkey; 3) 3University of Birmingham, Department of Medical and Molecular Genetics, School of Clinical and Experimental Medicine.

Long QT syndrome is one of the most common cardiac ion channel disease which its morbidity and mortality rate can be lessened with an early diagnosis and proper treatment. It is a cardiac repolarization abnormality that is characterized by prolonged OT interval and propensity for ventricular tachycardia (VT) of the torsades de pointes type are the charecteristics of the disease. This syndrome represent high risk of presyncope, syncope, cardiac arrest and sudden death. Jervell and Lange-Nielson syndrome (JLNS) is one of the inherited form of long QT syndromes. It is inherited recessively and characterized by profound sensorineural deafness and prolongation of the OT interval, thus representing abnormal ventricular repolarizasiton. JLNS has been shown to occur due to homozygous and copmpound heterozygous mutations in KCNQ1 or KCNE1. There was one clinical report on JLNS in Turkey; however, it was not confirmed by a molecular study. We identified a homozygous mutation in KCNQ1 in a 3.5-yr-old female child with JLNS, who visited the hospital due to recurrent syncope and seizures and had congenital sensorineural deafness. Her electrocardiogram revealed a markedly prolonged QT interval. The sequence analysis of the proband revealed the presence of homozygous missence mutation (c.728G>A, p.R243H). Heterozygous mutation in KCNQ1 was identified on the maternal, paternal and sister side. Even if with a high dose β-blocker therapy the patient has twice VT attacks, because of this reason the implanted cardiac defibrillator (ICD) was planned and implanted. We suggest early genetic diagnosis for proper management of the disease and genetic counseling.

1709T

Genetic Risk Profiles For Incident Coronary Heart Disease Using a 28 SNP Marker Panel In Four Prospective Cohorts. *E. Tikkanen[1,2], A.S. Havulinna[2], A. Palotie[1,3,4,5], V. Salomaa[2], S. Ripatti[1,2,3].* 1) Institute for Molecular Medicine Finland FIMM, Helsinki, Finland; 2) National Institute for Health and Welfare, Helsinki, Finland; 3) Wellcome Trust Sanger Institute, Hinxton, UK; 4) Department of Medical Genetics, Haartman Institute, University of Helsinki and Helsinki University Central Hospital, Helsinki, Finland; 5) Broad Institute, Cambridge, USA.

Genome-wide association studies have identified several single-nucleotide polymorphisms (SNPs) associated with coronary heart disease (CHD). The clinical and practical utility of these findings in prediction of CHD has been under debate. We tested 28 genetic variants in four Finnish prospective cohorts (N=24,124). During the median follow-up time of 11 years (IQR 8.75–14.25 years), we observed 1552 (6%) cardiovascular disease events (1093 CHD events). We compared the associations of previously reported 13-SNP genetic risk score and our novel 28-SNP score using Cox proportional hazards model and tested how well the scores predict future CHD events over traditional risk factors and family history of CHD. We found that the genetic risk scores were strongly associated with incident cardiovascular events after adjusting for traditional risk factors. The 28-SNP risk score was more strongly associated with CVD events than the 13-SNP score, particularly in individuals with high genetic risk load, thus showing evidence for non-linear relationship to risk. Subjects in the highest 10% of GRS had 2.1-fold (95% CI 1.68–2.56) increased risk for coronary heart disease in the 28-SNP score and 1.6 (95% CI 1.26–1.91) in the 13-SNP score, when compared with the subjects in the middle 20% of the GRS distribution. The 28-SNP score improved discrimination of CHD cases above traditional risk factors and 13-SNP score, when a non-linear form of genetic effects was allowed in the risk model (AUC difference = 0.4%, P=0.009). We conclude that the genetic risk score based on 28 CHD SNPs could help to identify individuals at high genetic risk for the first CHD event, but the accumulation of genetic risk alleles may not increase the CHD risk linearly. The more accurate classification of subjects using the GRS, particularly for high risk individuals, may have potential for better targeting intervention strategies, but the clinical utility of the genetic risk score needs to be further tested.

1710T

Vascular Stiffness in a Healthy High Risk African American Population is Modified by the Extent of European Admixture. *D. Vaidya, R.A. Mathias, L.R. Yanek, L.C. Becker, D.M. Becker.* Medicine, Johns Hopkins University, Baltimore, MD.

Background: Compared to European Americans (EA), African-Americans (AA) have stiffer peripheral vessels, reflected in reduced carotid distensibility coefficient (DC). To determine whether this racial difference may be genetically determined, we examined the extent to which the variance in carotid distensibility in AA could be explained by EA admixture either at a global or local at genomic level. **Methods**: We examined data from 344 AA, 62% women, aged 25–76 years, enrolled in a large study (GeneSTAR) of apparently healthy people with a family history of early-onset coronary artery disease. DC was assessed using 2D ultrasound, calculated as 2*(fractional change in diameter from diastole to systole)/(systolic -diastolic blood pressure). By its calculation DC is inherently corrected for blood pressure levels. EA admixture was determined using a panel of 50,000 ancestry informative markers (deCODE Genetics), and local ancestry was calculated on Illumina Human 1M genomewide SNP panel using LAMP. Associations of log-transformed DC were tested using mixed model regressions adjusted for age, sex, sex*age interaction and within-family correlations. LAMP models were adjusted for population stratification PCAs derived from the Illumina 1M SNPs (EIGENSTRAT). **Results**: The median [interquartile range] of the DC was 0.0017 [0.0012–0.0024] $mmHg^{-1}$. Every 10% incremental level of EA admixture was associated with 5% higher DC (95% CI: 1% to 9%, p= 0.005), reflecting more distensibility, and less stiffness. In genomewide local ancestry analysis adjusted for sex, age, sex*age interaction, population stratification PCAs and within-family correlations, of 2756 genome segments in local ancestry LD, the highest association for local ancestry was found in Chromosome 8, positions 8.3M to 10M (Build 37.3), p=0.0012. On adjusting for local ancestry in this region, population stratification PCA1 representing global Caucasian ancestry was no longer significantly associated with DC (p=0.93). **Conclusions**: The racial difference in arterial distensibility between AA and EA is likely to have a basis in genetic admixture. We have identified a candidate region on chromosome 8 that may be responsible for this global admixture association.

1711T

A Dutch founder mutation in the cardiac regulatory light chain (MYL2). *I.P.C. Krapels[1], G.R.F. Claes[1], M.B. Hoos[1], Y.E.G. Barrois[1], J-W. Sels[2], A.T.J.M. Helderman-van den Enden[1], P.G.A. Volders[3], H.J.M. Smeets[1], A. Van den Wijngaard[1].* 1) Clinical Genetics, Maastricht University Medical Center, Maastricht, Netherlands; 2) Cardiology, Catharina Hospital, Eindhoven, The Netherlands; 3) Cardiology, Maastricht University Medical Center, Maastricht, The Netherlands.

Hypertrophic cardiomyopathy (HCM) is a common disease defined as left ventricular hypertrophy in the absence of other cardiac or systemic disease explaining the hypertrophy with an estimated prevalence is 0.2% in the general population. HCM is most often inherited; genetic causes of HCM are merely found in the sarcomeric genes. In approximately 60% of patients an autosomal dominant mutation is identified. Most of the reported mutations are distinct and unique for each family. In addition, clinical heterogeneity and incomplete penetrance are observed in families with a mutation. Sometimes, sudden cardiac death can be the first manifestation of the disease. We report on a novel Dutch founder mutation in MYL2, a gene that is rarely mutated in HCM (less than 1% of identified HCM mutations). We identified a p.Glu22Lys (c.64G>A) mutation in 10 different HCM families of which 3 large pedigrees. We show that the MYL2 p.Glu22Lys mutation is actually a founder mutation which probably originated 450 years ago in the region around the city of Eindhoven in the South-East of the Netherlands. The availability of clinical cardiologic data of multiple MYL2 mutation carriers in these families gives a better understanding of the clinical phenotype and risk stratification for this mutation. Our data show that MYL2 mutation carriers usually have a benign disease course with low penetrance, no sudden cardiac deaths, moderate left ventricular hypertrophy and late disease onset in the 5th or 6th decade of life. However, when an additional risk factor is present like hypertension or another mutation, the disease penetrance is higher with an earlier onset.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1712T

Genome-wide DNA methylation profiling in myocardial infarction. *A. Russo[1,2], G. Fiorito[1], S. Guarrera[1], C. Di Gaetano[1,2], F. Ricceri[1,2], F. Rosa[1], A. Allione[1], F. Voglino[1], L. Iacoviello[3], M.C. Giurdanella[4], R. Tumino[4], S. Grioni[5], V. Krogh[5], A. Mattiello[6], S. Panico[6], P. Vineis[1,7], C. Sacerdote[1,8], G. Matullo[1,2].* 1) Human Genetics Foundation, Turin, Turin, Italy; 2) Genetics, Biology and Biochemistry Department, University of Torino, Torino, Italy; 3) Fondazione di Ricerca e Cura "Giovanni Paolo II", Catholic University, Campobasso, Italy; 4) Cancer Registry and Histopathology Unit, "Civile-M.P. Arezzo" Hospital, ASP7, Ragusa, Italy; 5) Department of Preventive and Predictive Medicine, Epidemiology and Prevention Unit, Fondazione IRCSS Istituto Nazionale dei Tumori, Milano, Italy; 6) Department of Clinical and Experimental Medicine, Federico II University, Napoli, Italy; 7) Epidemiology and Public Health, Imperial College London, UK; 8) Cancer Epidemiology, CPO-Piemonte, Torino, Italy.

Methylation of cytosine bases in DNA CpG islands is an important epigenetic regulation mechanism in the organ development and differentiation, aging and several diseases, ranging from neurological and autoimmune disorders to cancer. To investigate the role of differential methylation on the onset of myocardial infarction (MI), we examined the methylation levels of more than 450K CpG sites in 129 cases and 129 matched controls belonging to the Italian section of the European Perspective Investigation into Cancer and Nutrition (EPIC) cohort. EPIC healthy volunteers were recruited between 1994–98 and followed up for cancers or other diseases; blood and other biological samples were stored in liquid nitrogen until use. For the CpG methylation level assessment we used the HumanMethylation450 BeadChip from Illumina Inc. (San Diego, CA, USA). Data were quality checked, normalized, and analyzed according to standard procedures using the R statistical environment (Bioconductor's open source Methylumi package). To account for sex specific methylation and risk profiles, logistic regression analyses were conducted separately for males (95 matched pairs) and females (34 matched pairs). All analyses were corrected by matching variables (age, season and center of recruitment) and cardiovascular risk factors when significantly different between cases and controls (smoke, BMI, waist/hip ratio). In males, the most differentially methylated CpG site between cases and controls was found on chromosome 1 (p=8.12*10-6) in a gene encoding for a zinc finger protein with transcriptional repression function. The following CpG significant signal (p=2.11*10-5) resides in the gene-body region of a gene (Chr. 8) involved in the lipid metabolism supporting its possible involvement in the onset of MI. In females, no CpG signal reached the statistical significance of 10-5, possibly due also to the limited sample size. The present results suggest to further investigate the blood methylation profiles in a larger number of subjects in independent cohorts.

1713T

Common cardiovascular disease risk factors are associated with mitochondrial DNA levels. *S. Mitchell[1], K. Brown-Gentry[1], M. Allen[1], L. Hunt[2], P. Mayo[1], N. Schnetz-Boutaud[1], D.C. Crawford[1], D.G. Murdock[1].* 1) Center for Human Genetics Research, Vanderbilt University, Nashville, TN; 2) Department of Molecular Physiology & Biophysics, Vanderbilt University, Nashville, TN.

Mitochondria are the primary energy producers of the cell and are essential for cellular metabolism. They are unique organelles in that they maintain their own DNA separate from the nuclear genome. Maintenance and expression of mitochondrial DNA (mtDNA) are essential for mitochondrial fitness; decreased mtDNA levels have been linked to mitochondrial dysfunction. Various classes of disease have been associated with mitochondrial dysfunction including neurodegenerative disorders, cancer, diabetes, and cardiovascular disease (CVD). Cardiovascular phenotypes, such as cardiomyopathy and cardiac dysrhythmias, are commonly observed in diseases resulting from mtDNA point mutations or deletions. The different roles mitochondria play in normal cardiac function, and exactly how mitochondrial dysfunction contributes to CVD, have not been fully elucidated. In the present study, we determined mtDNA levels in more than 6,000 samples from three populations (non-Hispanic Whites (NHW), non-Hispanic Blacks (NHB), and Mexican Americans (MA)) from the National Health and Nutrition Examination Surveys (NHANES) 1999–2002 and examined the relationship between mtDNA levels and common CVD risk factors. Using quantitative real-time PCR, mtDNA levels (mitochondrial genome equivalents, or MGEs) were determined as a ratio of expression of the mitochondrial ND5 and nuclear RNASEP genes. Linear regression was performed to examine the relationship between MGEs and known CVD risk factors.Preliminary analysis revealed the mean MGEs were similar across populations: 249, 247, and 252 in NHW, NHB, and MA, respectively. Low density lipoprotein was significantly associated with mtDNA levels in MA, while high density lipoprotein was significantly associated in NHW. Additionally, GWAS-identified variants associated with lipid profiles were also significantly associated with mtDNA levels in all three populations. Further conditional analysis is required to facilitate elucidation of the link between common CVD risk factors, such as lipid profile and genetic variants, and mtDNA levels and how this relationship contributes to the pathological mechanisms of CVD. Results of this study may identify factors influencing mtDNA maintenance and reveal novel phenotype associations with mtDNA levels, both of which are likely critical to uncovering the role mitochondrial dysfunction plays in the etiology of CVD.

1714T

Assessing change in quality of life in patients participating in Additional KIF6 Risk Offers Better Adherence to Statins (AKROBATS) trial. *HR. Superko[1], SL. Charland[2], BC. Agatep[2], V. Herrera[2], M. Ryvkin[2], BJ. Schrader[2], J. Shabbeer[3], JJ. Devlin[1], EJ. Stanek[2].* 1) Celera Corporation, Alameda, CA.,; 2) Medco Research Institute, LLC/an Express Scripts Company, Bethesda, MD; 3) Celera Corporation, currently with Ariosa Diagnostics Inc., San Jose, CA.

Background: Few studies have prospectively evaluated the psychosocial and behavioral impact of genetic testing, and there remains some concern about potential negative effects. We evaluated the effect of providing kinesin-like protein 6 (KIF6) genotype results directly to the patient on health-related quality of life (QOL). Methods: Patients (N=647) new to statin therapy enrolled in the AKROBATS trial completed a validated QOL questionnaire (SF-12v2) at baseline and were then provided an interpretation of their KIF6 results. The SF-12v2 was completed again after 6 months of follow-up; it yields 3 summary scores: Physical Component Summary (PCS) which incorporates physical functioning and pain, Mental Component Summary (MCS), and a utility score (SF-6D). Categorical and continuous variables were analyzed using chi-squared and t-tests, respectively (significance at p<0.05). Multivariate analysis was used to evaluate the relationship between KIF6 carrier status and QOL scores. Results: Carriers (N=423) and non-carriers (N=224) did not differ significantly by age (60±12 vs 61±11 years), sex (48% vs 42% males), Deyo-Charlson score (0.59±1.10 vs 0.66±1.18), or BMI (30.5±6.7 vs 30.6±7.0 kg/m2). Non-carriers were more likely to be white (86% vs 88%, p<0.001), while carriers were prescribed simvastatin more often (46% vs 38%, p=0.046). At baseline, KIF6 carrier and non-carriers did not differ by any QOL score: PCS (46.5±10.9 vs 45.8±11.0), MCS (54.0±8.4 vs 54.5±8.6), and SF-6D (0.79±0.13 vs 0.79±0.14), respectively. After adjusting for age, sex, race, income, BMI, statin adherence, concomitant diseases, medications, education, and baseline QOL scores there were no differences between carriers and non-carriers after 6 months: PCS (47.5; 95% CI, 46.8–48.2 vs 47.11; 95% CI, 46.2–48.1); MCS (53.9; 95% CI, 53.2–54.6 vs 54.5; 95% CI, 53.5–55.4); and, SF-6D (0.80; 95% CI, 0.79–0.81 vs 0.80; 95% CI 0.79–0.82). Conclusion: Patients' quality of life remained stable over 6 months and was unaffected by knowledge of their KIF6 carrier status. As genetic testing advances in clinical practice, it is important to fully understand the physical and mental effect of informing patients of their genetic makeup and potential drug response or disease risks.

1715T

Heritability of John Henryism, and Correlation Between John Henryism and Hypertension in the Jackson Heart Study. *S.G. Buxbaum[1], P. Goel[1], W. White[2], M. Gregoski[3], S.H. Dunn[4].* 1) Jackson Heart Study, Jackson State Univ, Jackson, MS; 2) Tougaloo College; 3) Medical University of South Carolina; 4) University of Pittsburgh.

Purpose: The John Henryism Scale of Active Coping is measured in the Jackson Heart Study, a cohort of African Americans in the Jackson, MS area. The purpose of this study is to assess the heritability of John Henryism (JH) and to determine whether or not it is correlated with hypertension. The null hypothesis for this study was twofold: 1) John Henryism is not associated with hypertension; 2) John Henryism is not genetic. The initial data set comprised 4,567 participants who answered questions in the third JHS annual follow-up questionnaire and 64.5% are women. **Method**: After recoding the JH survey responses, a score was created by summing the values. Spearman correlations between JH and both hypertension and age were determined. Heritability of the JH score within the family subset, excluding singletons (N= 1040), was determined using S.A.G.E. and SOLAR, with adjustment for hypertension and age. **Results**: 3,987 participants answered all 12 questions in the John Henryism questionnaire. Among these, the correlation between JH and hypertension was not significant (p-value = 0.39). However, JH *is* inversely correlated with age (r= -0.12, p < 0.0001). The heritability of John Henryism was 10% with adjustment for hypertension and 9% with further adjustment for age, and not quite statistically significant (p=0.07). **Conclusion**: Unlike a previous study, we did not find strong evidence of heritability of John Henryism. This study failed to reject the null hypotheses at an α level of .05. However, together with previous findings of significant heritability, the results are suggestive. Imputation of missing JH questionnaire responses will be done using STATA; this will increase the number of relative pairs in the analysis and may affect our heritability estimate.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1716W

Assessment of health information technology tools use among cancer genetic counselors. *M. Doerr[1], T. Vu[2], C. Eng[1].* 1) Genomic Medicine Institute, Cleveland Clinic, Cleveland, OH; 2) Cancer Genetics Program, Carilion Clinic, Roanoke, VA.

Purpose: Survey to establish baseline descriptive data regarding the use of health information technology (HIT) within cancer genetic counseling practice. This survey is a joint project of the Familial Cancer Risk Counseling Special Interest Group (Cancer SIG), Practice Issues Sub-Committee and the Health Information Technology Special Interest Group (Health IT SIG) of the National Society of Genetic Counselors (NSGC). This project was awarded an NSGC Cancer SIG Grant Award in March 2012. Methods: The survey was developed by the authors in conjunction with members of the Health IT and Cancer SIGs. The survey was reviewed and approved by the Cleveland Clinic Institutional Review Board (CCF12-365). The survey is hosted by Survey Monkey, and was publicized through eBlast May 1, 2012. It will remain open for a period of six (6) weeks; results in this abstract are therefore preliminary. Results: As of May 11, 2011, there are 121 respondents. Demographics are consistent with NSGC Professional Status Survey (PSS) 2012 for sex, age, race, and work setting for clinical counselors. At this time, Region IV is overrepresented and Region III is underrepresented in the respondent pool compared with the PSS. 61.7% of respondents sole area of clinical practice is cancer genetics. Among all respondents, 80% agreed/strongly agreed that HIT will improve their efficiency and two-thirds of respondents agreed/strongly agreed that HIT will improve their ability to implement new/different practice models. More than three-quarters of cancer genetic counselor respondents report using an electronic medical record (EMR) in their practice. Of those using an EMR, 79.8% maintain a separate patient database and 57.8% keep shadow charts which contain information not archived in the EMR. The majority (70%) are not satisfied with how family history information is maintained in the EMR. Conclusions: Although HIT use is widespread among cancer genetic counselors, satisfaction with EMR features lags.

1717W

Behavior and food-related concerns in children with Smith-Magenis syndrome are similar to Prader-Willi syndrome. *L.V. Barton[1], T.P. York[1], S.H. Elsea[1, 2].* 1) Human and Molecular Genetics, Virginia Commonwealth University, Richmond, VA; 2) Pediatrics, Virginia Commonwealth University, Richmond, VA.

Smith-Magenis syndrome (SMS) is a complex genetic disorder associated with intellectual disability (ID), sleep disturbances and early-onset obesity. Problem behaviors are commonly reported and may be more characteristic of SMS than physical findings, but few reports have examined changes in behaviors across development and how those behaviors influence weight gain. Parents and guardians (n=81) of children with SMS were surveyed to assess problem behaviors and food-related concerns. The Food-Related Problems Questionnaire (FRPQ) revealed significant positive correlations between age and total score and for each subscale, including preoccupation with food, impaired satiety, and composite negative behavior (CNB). SMS mean scores on the FRPQ subscales were compared to two reference populations: adults with general ID and adults with Prader-Willi syndrome (PWS), a syndrome with almost universal early-onset obesity. SMS mean scores were more similar to PWS than ID in the subscales of preoccupation and CNB, as well as the total FRPQ score. The Behavior Problems Inventory (BPI) showed SMS mean scores that were significantly increased relative to the ID reference population, indicating a higher frequency and increased severity of behavior problems in SMS. BPI data show that total behaviors do not increase with age; however, many negative behaviors including destroying things and head hitting and were consistently reported across all ages. Data show a clear increase in food-related concerns with age, with problems reported in SMS more similar to those seen in PWS than general ID. While a description of the behaviors seen in SMS is useful, identification of similarities between SMS and PWS may be especially helpful in developing more effective therapies, particularly for excessive weight. Previous reports have noted excessive weight in both SMS and PWS, but these data are the first to examine the behaviors associated with the weight gain in both syndromes. Similarities identified between SMS and PWS with regard to food-related problems suggest that treatments for obesity in PWS may serve individuals with SMS. For persons with PWS, avoiding obesity is possible if a diet routine is utilized that addresses food-related behaviors before they manifest, typically in the toddler years. These data suggest that a different approach may need to be taken with children with SMS regarding long-term outcomes and therapies to include routines effective for individuals with PWS.

1718W

Genetic counseling future in historicist Middle East culture. *A. Haghighatfard, M. Mahdavi.* Islamic Azad university, tehran, Iran.

Genetic counseling is a process of communication and education that considering expression and transference of genetic disorders . Achieving to this communication is depended to knowledge about culture and traditions of society and psychological situation of consultant. This article is an inspection about social reactions against genetic counseling as a new branch of medical science in historic and historicist middle east societies and their own culture. Public beliefe to “paternal big family”and “god willing destiny”are two major challenges of genetic counseling in middle east. In case of paternal big family middle east people especially arab tribes of persian gulf region believe that familial marriage especially children of two brothers makes stable and honorable family . for example familial marriage rate in Saudi Arabia and Kuwait are 12 times more than north America .even in some arab tribes marriage of a girl needs her cousin's permission. in the other case religious belief to god willing destiny that is based on some Quran sentences (also mentioned in Torah) made kind of religious historicism. In this ideology called “taghdir”, no person can predict about illness or healthy of newborn child; because it is part of god's desigens and theology . Now middle east geneticist are facing with public distrust about genetic counseling ; especially in old generation. Also increasing rate of recessive genetic disorders patients and deacreasing the gene pool diversity . in the absence of state programs for public education, “perceived personal control” has a completely different definition from west societies. It seem so spiritual influence of pioneer muslim clerics and their “Fatwa” (religious decree) could help to improve people's trust to genetic counseling. Beacause for muslims, the only thing that is more important from traditions is religious obligations.

1719W

Pre-conception genetic counseling awareness in a post-secondary educational institute. *A. NEOGI, A. MAHESH, D. JALALUDDIN SHARIFF.* BITS-PILANI DUBAI, Dubai, United Arab Emirates.

Genetic Counseling research has been carried out in several domains and the common target groups for such studies involve known patient families, genetic counselors and the consumer public. Young adults can contribute significantly to studies on health during the preconception stages where existence of known or unknown genetic conditions can affect conception, the pregnancy and the baby itself. Since preconception healthcare can be offered to any individual with reproductive potential, educational institutes have been identified as potential centers for awareness study target groups. Hypertension, diabetes, thyroid disease, obesity and polycystic ovary disorder are common disorders in the Middle East and India that have a significant effect on pregnancy. The study involves discussions with medical professionals. Information brochures have been designed and distributed to clinics and educational institutes. Health is the main focus for the clinic information sheet, whereas for college, both health and career aspects are included. To thoroughly assess the understanding of Genetic Counseling, a standard questionnaire is made based on a rating system (-10, 0, 10) and higher points indicate the presence and/or knowledge of a genetic condition/symptom either in the immediate family or relatives and friends. This is correlated with a checklist of the role of genetic counseling and corresponding points in this indicate the level of understanding of this role. Personal interviews and electronic surveys are carried out for the largely Indian female population of BITS-Dubai. A case study was done based on personal interview results. The results envisaged a generalized increase in awareness about preconception healthcare and genetic counseling for the few common disorders. Preliminary analysis shows no relationship between age of the individual and willingness to take a survey or go through a personal interview. Among the faculty, women above the age of 30 who are aware of the possible presence of a genetic condition are hesitant to consult a genetic counselor because of fear of information overload and lack of ability to manage symptoms in a current pregnancy. Part of this reluctance is attributed to them having incorrect information regarding testing options and unclear understanding of the role of a genetic counselor. However, the majority of the faculty agrees that there needs to be more awareness regarding genetic conditions and the impact on health.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1720W

Genetic Counseling in India - Challenges in a Developing Country. *R. Puri, U. Kotecha, R. Saxena, J. Verma, S. Movva, S. Kohli, I.C. Verma.* Center of Medical Genetics, Sir Ganga Ram Hospital, New Delhi, INDIA.

Genetic counseling in India has a different perspective as compared with the West. Firstly, the average physician has limited training for diagnosis of genetic disorders. Secondly counseling is mostly sought during pregnancy to ensure a normal fetus. Geneticists have to make a primary diagnosis in the absence of the proband, limited availability of medical documents and absence of DNA sample. From a tertiary care genetic center, we illustrate this challenge through three representative cases. A fourth gravida came at 8 weeks gestation for prenatal diagnosis. Her third born boy was well till 3 months age, after which he developed intractable seizures and neuroregression. Parents had noted increased tone, unexplained fever and irritability. He died at 3 years without a diagnosis. No medical documents except a CT Scan report of cerebral atrophy were available. Krabbe disease was suspected. As there was no DNA of the child, the parents were tested for the common 30 kb deletion, and were found to be positive. This confirmed Krabbe disease, making prenatal diagnosis possible. A consanguineous couple had lost two previous children with immune deficiency but no specific diagnosis. In the last child, though immunologic studies were done, diagnosis of the type of immune disorder was not made. A possibility of ADA deficiency was kept on review of the available data and molecular studies had to be done on the parents to confirm this and provide help in the next pregnancy. In the third case, a fourth gravida was referred for prenatal diagnosis at 12 weeks gestation. All 3 children had failure to thrive and neuroregression after 4 months age. Evaluation and bone marrow biopsy in the third child suggested a storage disorder. No confirmatory tests were done. In the absence of a cost effective strategy to establish a diagnosis for the family, a chorionic villus sampling was done after appropriate counseling. Enzyme analysis for GM1, Gaucher and Niemann Pick disease showed deficient activity of beta galactosidase. This not only confirmed GM1 gangliosidosis in the family but also prevented the birth of a fourth affected child. The challenge of counseling in India lies in a diagnosis to be established without examination of the proband and with no medical records, as these are usually disposed by the family after the death of the child. Molecular studies in the parents enable to confirm the diagnosis of the disorder and help to provide prenatal diagnosis.

1721W

A novel GLI2 mutation and discordant holoprosencephaly phenotypes in monozygotic twins. *C. Quindipan[1], M. Van Hirtum-Das[2], S. Saitta[1,3].* 1) Medical Genetics Institute, Los Angeles, CA; 2) Huntington Medical Foundation, Pasadena, CA; 3) Cedars-Sinai Medical Center, Los Angeles, CA.

Holoprosencephaly (HPE) is a developmental anomaly of the forebrain and midface affecting approximately 1:10,000 live births with widely variable features. Chromosomal aneuploidies account for 25–50% of cases, while 18–25% of HPE cases result from a single gene disorder. There are both autosomal recessive and dominant forms of HPE. We report 19-month-old monozygotic twins in which Twin A presents with single central incisor, microcephaly, hypotelorism, and normal development. Brain MRI performed at 14 months of age was normal. Twin B has significant global developmental delay, single central incisor, microcephaly, hypotelorism, failure to thrive, diabetes insipidus with dysgenesis of the corpus callosum and abnormality of the posterior pituitary noted on brain MRI. IVF was utilized in the pregnancy and a single embryo was implanted. Monozygosity was further confirmed by SNP microarray (LabCorp), which showed no pathogenic CNVs, no evidence of uniparental disomy, and excluded long contiguous regions of homozygosity. Molecular analysis of a panel of genes associated with HPE available on a clinical basis: SHH, GLI2, PTCH1, TGIF1, SIX3, and ZIC2, was performed revealing that both twins were heterozygous for a previously unreported C1573T missense mutation in GLI2 (Prevention Genetics). The reported mutation results in an amino acid substitution of Arginine to Tryptophan (R525W). The arginine residue is conserved among non-human primates with the substitution predicted to be pathogenic using PolyPhen-2. Subsequent parental analyses revealed that the father also carries the same GLI2 mutation. He is not affected with cognitive impairment nor does he have a single central incisor or other clinical features. A brain MRI has not been performed on him. He did not report any history of developmental delay or cognitive impairment, and his family history was unremarkable for clinical manifestations of HPE or stillbirths. Overall, we report a novel GLI2 missense mutation in monozygotic twins with discordant phenotypes and in their apparently unaffected father. We cannot exclude the possibility of an additional gene or epigenetic change that may explain the findings in this family that may require further whole exome and/or methylation analysis. There is wide clinical variability in this disorder, complicating recurrence risk counseling. Approximately 1/3 of individuals with microform HPE will have subclinical manifestations.

1722W

Furnishing appropriate information on mitochondrial diseases to patients and their families. *Y. Sato, Y. Goto.* National Center of Neurology and Psychiatry, Kodaira, Japan.

We've been doing genetic counseling for families who have patients with mitochondrial diseases. We clarified problems in the genetic counseling, and developed solutions for some of them. During the period from 2004 to 2011, a total of 38 clients received our genetic counseling. Approximately eighty percent of them were recommended to visit our hospital by their doctors. The clinical diagnoses included mitochondrial disease without particular category (n=13), Leigh syndrome (n=10), mitochondrial encephalomyopathy with lactic acidosis and stroke-like episodes (n=6), pyruvate dehydrogenase complex deficiency disease (n=6), chronic progressive external ophthalmoplegia (n=1), myoclonic epilepsy with ragged-red fibers (n=1) and Pearson disease (n=1). Most of them were concerned about heredity, causes, symptoms, diagnoses and treatments of the diseases. Some consulted us about their genetic testing, family planning and information sharing among their families. We provided various kinds of information and discussed clients' concerns. We recognized some characteristics of genetic counseling for mitochondrial diseases. First, neither healthcare professionals nor families have enough knowledge about diseases. Second, the explanations about genetics are hard to understand because of their complexity. Third, the results of genetic testing are less useful to expect the risk of occurrence and recurrence. Fourth, patients' mothers and her families often feel guilty due to maternal inheritance. It was suggested that appropriate supports according to clients' understandings and situations are needed. For providing information relating to the diseases, we developed the handbook. It contains explanation and illustration in every page, and consists of totally six parts: basic knowledge, diagnoses, symptoms, treatments, social resources and genetics. We expect that healthcare professionals can explain using illustrations, and families can understand using explanations. In addition, we will provide information on the web and hold seminars for healthcare professionals periodically. We also plan to obtain feedback from users and make use of it towards revisions of the handbook.

1723W

A special thalassemia case with probable uniparental disomy. *Y. ZHANG[1,2], J. WU[1], M.Q. MAI[1], A.H. YIN[1,2], H.K. DING[1,2], D.Q. QIN[1], Q.Y. DU[1], J.Q. LIANG[1].* 1) Prenatal Diagnosis Center & Genetic Disease Diagnosis and Treatment Center, Guangdong Women and Children Hospital, Guangzhou, Guangdong, China; 2) Maternal and Children Metabolic-Genetic Disease Key Laboratory of Guangdong Women and Children Hospital.

Background: Frequently occurs in South China, beta thalassemia is a kind of autosome recessive disease which can be diagnosis by gene analysis with many methods. Prenatal diagnosis is feasible if parents' genotype were known. 17 mutants in beta hemoglobin gene and two kinds of large deletion in beta gene cluster are mostly found in Chinese. Objective: The pregnant woman wanted to know the fetus's genotype of beta thalassemia, for she once gave birth to a boy with severe bata thalassemia, who needed blood transfusion regularly to suatain life from 2 years old. and prenatal diagnosis was subjected. Methods: The parent and proband's peripheral blood was first used to extract the genome DNA and to inspect beta hemoglobin gene(HBB) with PCR-RDB (reverse dot blotting). Then Gap-PCR was used to investigate if SEA-HPFH or Gγ+(Aγδβ)0 exist in the parent and proband. Also, gene sequencing was used to inspect the potential scare mutants in HBB. Furthermore, the fetus' genotype was detected with amniotic fluid. Finally, STS markers were used to analysis the strange genetic phenomenon in the family. Results: The mother was a carrier whose genotype was βCD41–42/βA, paternal genotype wasβA/βA, and the fetus' genotype was the same with that of the mother's. However, PCR-RDB and DNA sequencing show the CD41–42 homozygote for the proband(βCD41–42/βCD41–42), SEA-HPFH or Gγ+(Aγδβ)0 was not found in all of them, which contradicts Mendel's genetics. With informed consent for diagnostic genetic testing, paternity testing was carried out and the result was unexclusive. With four STS markers around CD41–42 in beta hemoglobin gene, we found the probalble uniparental disomy in the proband. Conclusion: The fetus was a carrier with CD41–42 homozygote in beta hemoglobin gene; if the proband was uniparental disomy or de novo mutant in beta hemoglobin gene needs further research.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1724W

Understanding the Needs of Parents of Children with a Genetic Diagnosis - can it improve clinical practice? *L. Bryson, J. Dunlop, A. Anderson, J. Berg.* University of Dundee, Ninewells Hospital, Dundee, Angus, United Kingdom.

There has been a very rapid increase in our ability to diagnose genetic conditions in children, using technology such as Next Generation Sequencing and array CGH. Very little is known about the impact of such diagnoses on parents, and the parents' need for support following testing. We, therefore, used a qualitative methodology to explore the effect of diagnosis on parents and identify areas in which additional support would be most useful. Semi structured interviews were undertaken with the parents of 11 children with a confirmed genetic diagnosis causing developmental delay and additional syndromic features. All patients were identified and recruited with ethical approval and consent from within NHS Tayside. The interviews were transcribed and analysed using framework analysis(1). Four main themes were identified from the transcripts: the reaction to having a genetic diagnosis; the support which was available; how parents felt they were managed by the healthcare team and how a genetic diagnosis affects the family. Within each theme, there were multiple sub-themes highlighting both positive and negative effects on the family. A key point that emerged was that parents felt there was a lack of signposting to support, especially before and immediately after diagnosis, and that provision of appropriate information and structured follow up would improve parents' experiences. Our results have allowed us to design a questionnaire, for both geneticists and paediatricians, to ascertain how often key themes are appropriately dealt with during consultations, and to propose a new framework for clinical follow up of parents with a child with a new genetic diagnosis. 1. Pope C, Ziebland S, Mays N. Analysing qualitative data. BMJ. 2000 January 8, 2000;320(7227):114-6.

1725W

A diagnostic assay for predicting medullary cystic kidney disease type 1 (MCKD1). *B. Blumenstiel[1], M. DeFelice[1], A. Kirby[1,2], A. Gnirke[1], A. Bleyer[3], S. Gabriel[1], M. Daly[1,2].* 1) Broad Institute of Harvard and MIT, Cambridge, MA; 2) Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA; 3) Section on Nephrology, Wake Forest School of Medicine, Winston-Salem, NC.

Medullary cystic kidney disease type 1 (MCKD1) is a rare autosomal dominant disorder which manifests as a slow deterioration in kidney function typically becoming symptomatic in the second to third decade in life. Disease progression ultimately results in the need for dialysis and/or kidney transplantation several decades post diagnosis. Until recently, decisive diagnosis of MCKD1 has been hampered by several factors including variable age of onset and symptom variability. A very recent study (Kirby et. al. pending publication) has implicated a single base (+C) frameshift insertion in a variable nucleotide tandem repeat (VNTR) region of the MUC1 gene located on Chromosome 1 as the causal mutation for the disorder. Interrogating and detecting a rare, single insertion event in a high-copy tandem repeat region posed a unique set of challenges. Due to the high wild-type: mutant allelic ratio (~100:1), high GC composition of the repeat (>80%) and variable positioning of the insertion within the VNTR region, many standard commercial genotyping protocols were not feasible. Exacerbating this challenge was the unique nature of the mutation being a single insertion event creating an eight base cytosine homopolymer (8C) against a high-copy background of wild-type (7C) repeats. Here we describe a novel and robust diagnostic assay for detecting the presence of a single mutated copy of a GC-rich 60-baseVNTR sequence in order to positively verify the causal variant as well as identify carrier patients prior to symptom development. We developed a probe-extension assay capable of detecting a single mutant copy within the predominantly wild-type VNTR using selective restriction digestion of wild-type copies and analyte detection by MALDI-TOF mass spectrometry. By genotyping 62 phenotypically affected and 79 unaffected relatives from collected families, we observed perfect segregation of the mutant allele with the disease phenotype. Additional genotyping of over 500 unrelated controls from HapMap CEU, Japanese, Chinese, Yoruba and Tuscan HapMap3 populations showed no evidence of the +C insertion event arising in the larger global population.

1726W

Studding the importance of common α-deletions among β-thalassemia minor individuals in an Iranian population based on their frequencies and CBC indexes. *A. Moosavi[1,2,3], M. Karimipoor[1], S. Zeinali[1], B. ZarbakhshH[1].* 1) Molecular Medicine Department, Pasteur Institute of Iran, Tehran, Iran; 2) Biochemistry.,Tarbiat Moallem University of Tehran,Tehran,Iran; 3) Karaj thalassemia screening center, Alborz Medical University,Karaj,Alborz,Iran.

Beta-thalassemia (thal) is the most common inherited monogenic recessive disorder in IRAN. In the screening carriers an important challenge is co-inheritance of α-deletions. The pathogenesis of thalassemia depends on alpha/beta chain ratio changes, so we can suppose that alpha deletions as a secondary modifier factor can result in misdiagnosis in the screening of beta-thal carriers. In this study the frequency of common alpha deletions in beta thalassemia carriers was determined. The study includes families referred from different primary health care center with hypochromic anemia [MCV<80fl; MCH<27pg] and A2> 3.5. After providing informed consent,genomic DNA was extracted from peripheral blood by salting out method and proteinase K. Allele-specific PCR was exploited for common β-mutations. Then common alpha deletions (-α3.7, -α4.2, -α20.5 and - MED)were screened by multiplex gap PCR in samples with β-mutations.At the end Multiplex polymerase chain reaction (m-PCR) assay was used to detect the $\alpha\alpha\alpha^{anti3.7}$ triplication and $\alpha\alpha\alpha\alpha^{anti3.7}$ quadruplicate in cases that we had co-inheritance of α-deletion in beta thalassemia carriers. **Among 227 β-thalassemia minor individuals we found α-thal mutations in 43 cases: 37 heterozygote -α3.7 (16.3%), 5 homo -α3.7 (2.2%)and 1 -MED (0.44%).There was not any $\alpha\alpha\alpha^{anti3.7}$ triplication and $\alpha\alpha\alpha\alpha^{anti3.7}$ quadruplicate in these samples.** The results show high prevalence of co-inheritance of alpha-thal and beta-thal in our selected population and comparing CBC indexes in these person shows low changes in these cases. Therefore it is highly recommended that physicians and genetic counselors involved in the screening program of beta thalassemia major in the country consider this phenomenon.

1727W

Genomic data interpretation challenges: healthy older relatives as super controls. *M.S. Naslavsky, R.C.M. Pavanello, N.C.V. Lourenço, A. Cerqueira, M. Lazar, M. Vainzof, D. Schlesinger, M. Zatz.* Human Genome Research Center, Biosciences Institute, University of Sao Paulo. Sao Paulo - SP, Brazil.

With the improvement of whole-genome sequencing and genotyping technologies at affordable costs, an increasing body of data from human populations is emerging. However, the interpretation of such a large flow of information represents a great challenge. Since the identification of rare variants is rising in number with large whole-genome and exome sequencing studies, assessing the pathogenicity of novel mutations/variants and their impact in health is of utmost importance. It was recently observed that many genetic variants once predicted to be deleterious were in fact present in healthy individuals. As an example, we recently ascertained a 44-year father of a 10-year-old daughter with bilateral coloboma and hearing loss, in whom a whole-genome CGH revealed a deletion of exons 38–44 in the dystrophin gene. This mutation, which was not found in almost 1000 Duchenne/Becker patients screened in our Center, was inherited from her asymptomatic father, who was further clinically and molecularly evaluated for prognosis and genetic counseling (GC). To assess whether the present case represents a rare case of non-penetrance, and aiming to obtain more information for prognosis and GC, we analyzed healthy older relatives. Mutation analysis revealed that his mother, brother and 56-year-old maternal uncle also carry the 38–44 deletion, suggesting it an unlikely cause of muscle weakness at least in this family. In short, in addition to population-based studies, keeping DNA from older relatives may turn into a valuable source of information. This is particularly relevant regarding the variability of specific mutations penetrance under influence of distinct genetic backgrounds and environmental factors.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1728W

Assessment of cascade testing in families of carriers identified through newborn screening for cystic fibrosis in western Brittany, France. *C. Ferec[1,2,3,4], I. Dugueperoux[1,2,3,4], M.-P. Audrezet[1,2,3,4], P. Parent[5], V. Scotet[1,3,4].* 1) Inserm, U1078, Brest, F-29200, France; 2) C.H.R.U. Brest, Hôpital Morvan, Laboratoire de Génétique Moléculaire, Brest, F-29200, France; 3) Etablissement Français du Sang (EFS) - Bretagne, Brest, F-29200, France; 4) Université de Bretagne Occidentale, Faculté de Médecine et des Sciences de la Santé, Brest, F-29200, France; 5) C.H.R.U. Brest, Hôpital Morvan, Service de Génétique Médicale, Brest, F-29200, France.

Objectives: DNA-based newborn screening (NBS) for cystic fibrosis (CF; {OMIM#219700}) aims to early detect affected children. However, it also leads to the unwanted but inevitable detection of some healthy carrier newborns (i.e. having only one CFTR mutation). If being a carrier has no influence on the development of the newborn, the knowledge of that status is nonetheless important. It provides opportunity for the parents and relatives to know their genetic status regarding CF, as they can benefit from cascade carrier testing. The aim of the study was to measure the use of cascade testing in families of carrier children identified through NBS in an area where CF is frequent (western Brittany, France) and where NBS is implemented for over 20 years. Methods: We reviewed all the carriers identified by NBS in the district of Finistère since the introduction of the IRT/DNA strategy (1991). In parallel, we recorded the genetic tests made in those families in the sole genetic laboratory of our area, over the 1991–2010 period. Results: Among the 200 378 newborns screened over that period, 199 (0.1%) from 195 families, were identified as CF carriers. To date, 148 families (75.4%) have required carrier testing at least once, and 374 genetic tests were performed (mean number/family: 2.5; range: [1–8]). Testing was mainly used by the parents (77.4% of the tests performed and 95.9% of the parental couples tested) and uncles/aunts (13.7% of the genetic tests done). Overall, family testing identified 183 CF carriers and evidenced 5 new 1-in-4 risk couples. Of them 4 requested prenatal diagnosis 5 times. One affected pregnancy was diagnosed and terminated. Almost all the tested couples (151/156 i.e. 96.8%) could be reassured about their genetic status and about the risk of having a child with a disease that remains severe and life limiting, despite important improvements in the management over the past decades. Moreover, considering the CF incidence observed over that time (1/2 569) and the deducted carrier rate (1/26), we estimated that NBS detected only 2.6% of all the expected carriers born over the 1991–2010 period, and only 1.9% of those born since NBS was enlarged to the whole country in 2002, which had led to a change in strategy. Carriers detected by NBS appeared to be well managed in our area and cascade testing, that prevents birth of other CF children in families, seems relatively active. Study approved by our local ethical committee.

1729W

NextGen sequencing of *HEXA*: A more sensitive assay. *J. Hoffman[1], M. Umbarger[2], C. Kennedy[2], B. Bishop[2], P. Saunders[2], G. Porreca[2], E. Strovel[3], M. Blitzer[3], J. Douyard[1], J. Davies[2], S. Hallam[2], V. Greger[2], C. Towne[2].* 1) Division of Genetics, Department of Pediatrics, Tufts Medical Center, Boston, MA; 2) Good Start Genetics, Cambridge, MA; 3) Division of Genetics, Department of Pediatrics, University of MD School of Medicine, Baltimore, MD.

Tay-Sachs disease (TSD), an autosomal recessive neurodegenerative disease, serves as the prototype for ethnic based carrier screening. With a carrier rate of approximately 1 in 25 in Ashkenazi Jews (AJs) and French Canadians, use of highly sensitive methodology is critical in screening for this disease. *HEXA* enzyme analysis has long been considered the gold standard for TSD screening, detecting roughly 98% of carriers from all ethnic backgrounds. Due to poor enzyme stability and effects of certain medications and/or pregnancy on serum enzyme analysis, this method has significant limitations. DNA mutation screening for TSD is widely available, with sensitivity of 89–96% for carriers of AJ descent, but much lower for other ethnicities. We proposed NextGen sequencing (NGS), a rapid and efficient methodology, along with genotyping for the common 7.6kb *HEXA* deletion as a possible screening method for TSD. 74 blood samples and questionnaires were collected at a family conference for TSD and related diseases. Results of serum *HEXA* analysis (enzyme) and NGS + 7.6kb del analysis were compared. Mutations were confirmed with Sanger sequencing. 51/74 participants (69%) had positive enzyme results (46 carriers, 5 with late-onset TSD). 42/51 enzyme positives had a pathogenic mutation, detected via NGS +7.6kb del analysis. 7/ 51 were found to have a previously reported VUS, 1/51 had a novel VUS, and 1/51 had the *HEXA* pseudoallele (false positive enzyme). 2 of the enzyme-negative individuals had the B1 pathogenic allele (not detectable via enzyme). 7/74 (10%) had enzyme-inconclusive results; 4 had no reported family history of TSD and negative NGS + 7.6kb del analysis, 2 were obligate carriers with previously reported mutations, and 1 had a VUS. Overall, NGS + 7.6kb del screening of *HEXA* found a pathogenic mutation, pseudoallele, or VUS in 100% of the enzyme-positive or obligate carrier/enzyme-inconclusive samples, detected 2 carriers that would have been missed by enzyme alone, and detected mutations in 7/51 (14%) which would have been missed using common mutation panels. The VUSs require further laboratory and pedigree analysis to prove pathogenicity. Our data suggest that NextGen sequencing in combination with enzyme may be used as an efficient screening technology for people of all backgrounds, and will provide a more sensitive test than enzyme alone or a mutation panel with enzyme.

1730W

No association between ATXN2 (SCA2) CAG repeat expansion and amyotrophic lateral sclerosis /parkinsonism-dementia complex of the Kii Peninsula, Japan. *H. Tomiyama[1,2], C. Yamashita[1], R. Sasaki[3], Y. Li[4], M. Funayama[1,4], N. Hattori[1,2,4], S. Kuzuhara[5], Y. Kokubo[3].* 1) Department of Neurology, Juntendo University School of Medicine, Tokyo, Japan; 2) Department of Neuroscience for Neurodegenerative Disorders, Juntendo University School of Medicine, Tokyo, Japan; 3) Department of Neurology, Mie University School of Medicine, Tsu, Japan; 4) Research Institute for Diseases of Old Age, Graduate School of Medicine, Juntendo University, Tokyo, Japan; 5) Department of Medical Welfare, the Faculty of Health Science, Suzuka University of Medical Science, Suzuka, Japan.

Objectives. To clarify the genetic background of amyotrophic lateral sclerosis (ALS)/parkinsonism-dementia complex (PDC) of the Kii peninsula, Japan (Kii ALS/PDC). Backgrounds. Etiology and pathogenesis of amyotrophic lateral sclerosis / parkinsonism-dementia complex (ALS/PDC) of Guam and the Kii peninsula remain unclear. Intermediate CAG repeat expansions of ATXN2 (SCA2) have been detected in autosomal dominant parkinsonism. Recently, intermediate CAG repeat expansions of ATXN2 were also recognized as a predisposing factor for ALS. Methods. In 12 patients with pathologically diagnosed Kii ALS/PDC, we performed fragment analysis for PCR products and determined the lengths of CAG repeat expansions of ATXN2. Results. CAG repeat lengths of ATXN2 were normal (22 repeats) in all 12 patients with Kii ALS/PDC. Conclusions. No abnormal CAG repeat expansions of ATXN2 were found in patients with Kii ALS/PDC. Combined with our previous study, we have found no causative mutations in the reported major causative or susceptible genes (25 genes) related to ALS, frontotemporal lober degeneration, parkinsonism, synucleinopathy, TDP-43 proteinopathy, and tauopathy. However, familial clustering and lack of any environment factors suggest that unidentified genetic factors cause Kii ALS/PDC. ALS/PDC warrants exome or whole genome sequencing and other genetic approaches.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1731W

Assigning disease liability of mutations in the CFTR gene improves detection rate of CF carrier screening. *K.R. Siklosi[1], P.R. Sosnay[2,3], F. Van Goor[4], K. Kaniecki[1], M. Corey[6,7], A.S. Romalho[13,14], M.D. Amaral[13,14], R. Dorfman[8], R. Karchin[9], M.H. Lewis[10], H. Yu[4], J. Zielenski[8], J.M. Rommens[8,11], C. Castellani[5], C.M. Penland[12], G.R. Cutting[1,2].* 1) McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University, Baltimore, MD; 2) Department of Medicine, Johns Hopkins University, Baltimore, MD; 3) Perdana University Graduate School of Medicine, Serdang, Malaysia; 4) Vertex Pharmaceuticals Incorporated, San Diego, CA; 5) Cystic Fibrosis Center, Azienda Ospedaliera Universitaria Integrata, Verona, Italy; 6) Child Evaluative Health Sciences, The Hospital for Sick Children, Toronto, ON, Canada; 7) Dalla Lana School of Public Health, University of Toronto, ON, Canada; 8) Genetics and Genome Biology, The Hospital for Sick Children, Toronto, ON, Canada; 9) Department of Biomedical Engineering, Institute for Computational Medicine, Johns Hopkins University, Baltimore, MD; 10) Genetics and Public Policy Center, Berman Institute for Bioethics, Johns Hopkins University, Baltimore, MD; 11) Department of Molecular Genetics, University of Toronto, ON, Canada; 12) Cystic Fibrosis Foundation, Bethesda, MD; 13) University of Lisboa, Faculty of Sciences, Centre for Biodiversity, Functional and Integrative Genomics (BioFIG), Lisboa, Portugal; 14) Department of Genetics, National Institute of Health, Lisboa, Portugal.

Testing for CF carrier status is commonly performed in the US. The American College of Medical Genetics (ACMG) recommends a screening panel of 23 CFTR mutations of ≥0.1 frequency in CF patients. The panel detects approximately 85% of Caucasian CF carriers and fewer of other ethnicities. Optimal screening would detect ≥95% of carriers so that couples in whom one tests positive and the other negative have a residual risk for having a child with CF that is below their population risk. Improving CF carrier detection rate to ≥95% requires a larger set of mutations of known prevalence and disease liability. The Clinical and Functional Translation of CFTR (CFTR2) project was initiated to define clinical consequences and disease liability of CFTR mutations. The CFTR2 database contains 39,545 patients from the US and Europe. Patients in the database harbor 1,045 unique CFTR mutations, 160 of which occur at a frequency of ≥0.01 and together account for 97% of the identified CF alleles; these underwent further analysis. Mutations were evaluated by clinical, functional and penetrance analysis. The mean sweat chloride concentrations of patients bearing the mutation in question in trans with a severe ACMG mutation was used for clinical verification. Analysis of CFTR function was performed by measuring short-circuit chloride current in cell lines expressing missense mutant CFTR. Mutations introducing a premature termination codon were classified as functionally deleterious. Penetrance was evaluated by genotyping fertile fathers of CF patients who are expected to carry a deleterious CFTR mutation and a “healthy” CFTR gene. The presence of another CFTR mutation in the “healthy” gene indicated that the mutation may have reduced penetrance or no deleterious effect on CFTR function. Of the 160 mutations, 122 were deemed CF-causing based on 1) mean sweat chloride concentration ≥60 mEq/L (146/159 mutations with sweat chloride measurements available); 2) experimental evidence of CFTR dysfunction (120/138 tested mutations); and 3) not observed in trans with an ACMG mutation in the fathers (144/160 mutations). The remaining mutations were deemed as mutations of varying clinical consequence (15) and non CF-causing (6) while 17 mutations continue to be analyzed. Summary information for mutations in the list of 160 is included on the CFTR2 website (www.cftr2.org). The 122 mutations verified as CF-causing to date attains a 95% sensitivity for CF carrier detection.

1732W

Mutational screening of C9ORF72, SOD1, TARDBP and FUS genes in an ALS cohort from the North-Eastern part of Italy. *C. Bertolin[1], J. Agostini[2], C. D'Ascenzo[2], F. Boaretto[1], C. Salvoro[1], C. Angelini[2], E. Pegoraro[2], G. Sorarù[2], ML. Mostacciuolo[1].* 1) Laboratory of Human Genetics, Department of Biology, University of Padova, Italy; 2) Department of Neurosciences, University of Padova, Italy.

Background: Mutations in C9ORF72, SOD1, TARDBP and FUS genes appear as the most common genetic cause of familial (FALS) and sporadic (SALS) forms of ALS. Objectives: To assess the frequency of ALS genes mutations in a large cohort of ALS patients referred to the ALS Clinic of the University of Padova. Most of the patients are Italian (about 98%) and originating from the North-Eastern regions (about 85%). Methods: Overall we investigated 327 patients, 39 with reported family history (FALS) and 288 sporadic/isolated cases (SALS). We performed Sanger sequencing of all coding exons of SOD1 and FUS genes and two exons of TARDBP (exon 4 and 6) in all ALS patients. Moreover, with a two-step repeat primer PCR, we searched for the expanded GGGGCC hexanucleotide repeats in the C9ORF72 gene, recently associated with the disease. Results: 14 pathogenic missense mutations were found in 14 unrelated patients: 6 SOD1 (3 in FALS and 3 in SALS), 5 TARDBP (1 in FALS and 4 in SALS) and 3 FUS (1 FALS and 2 SALS). All the mutations but one have already been reported: a novel TARDBP mutation was identified in exon 4. Expanded repeats in C9ORF72 were found in 19 patients (7 FALS and 12 in SALS). From a clinical point of view, we could observe a shorter disease lifespan and an increased dementia comorbidity in carriers of the expanded C9ORF72 hexanucleotide repeats (25% vs 7% of overall sample, p=0.0009). Conclusions: This study is the first mutation screening of the most frequently mutated genes involved in ALS (SOD1, FUS, TARDBP and C9ORF72) performed in an Italian North-East cohort of patients. Overall, our screening identified the molecular causes of ALS in the 30% of FALS and in the 7% of SALS patients. As already reported by other studies C9ORF72 and SOD1 are the most common mutated genes but, in our sample, mutation frequencies are lower if compared with those reported in other Italian ALS cohorts [Chiò et al., Brain. 135:784–93 (2012)]. Recently, other studies pointed out to a relevant correlation among population origins and gene-specific mutation frequencies. In that view, our results could be very useful to direct the genetic test in ALS patients.

1733W

Cree Leukoencephalopathy and Cree Encephalitis Carrier Screening Program: Lived Experiences of Carrier and Non-Carrier Couples. *J. Le Clerc-Blain[1,2,3], H. Denoncourt[4], A. Richter[1,2,3], A.M. Laberge[1,2,3].* 1) Medical Genetics Division, CHU Sainte-Justine, Montreal, Quebec, Canada; 2) Dept of Pediatrics, Université de Montréal, Montréal, Québec, Canada; 3) CHU Sainte-Justine Research Center, Montréal, Québec, Canada; 4) Cree Board of Health and Social Services of James Bay, Chisasibi, Quebec, Canada.

STATEMENT OF PURPOSE: Cree leukoencephalopathy (CLE) and Cree encephalitis (CE) are severe neurodegenerative autosomal recessive conditions with carrier rates of 1/10 and 1/30, (respectively) in the James Bay Cree communities of Northern Quebec, Canada. Since 2006 a regional public health educational and screening program informs the population about CE/CLE and offers carrier screening starting at 14 years of age. OBJECTIVES of this study are to describe the lived experience of couples who participated in the CE/CLE carrier screening program and to assess their level of satisfaction with the program. METHODOLOGY: In collaboration with the local parents' committee (Eeyou Awash Foundation), we took a phenomenological approach to evaluate the lived experience of couples who have gone through the CE/CLE carrier screening program. Nineteen adults recruited in two Cree communities took part in semi-structured interviews that were audio-recorded, transcribed verbatim and coded with QDA Miner. Transcripts were analyzed to extract elements describing the underlying essence of the participants' experience. RESULTS: All participants had previous direct or indirect contact with a child affected by CE/CLE. Most expressed concern over the impact that carrier status identification may have on their family's future generations. Half of the participants had not discussed the screening process or their own results with family members but wondered about the mutation origin in their family. Participants emphasized the importance of having access to prenatal diagnosis and a minority spontaneously expressed interest in preimplantation genetic diagnosis. They articulated the importance of being able to know the foetus' status, thus allowing pregnancy decisions. However, couples were ambivalent about the decision to undergo medical termination of pregnancy. All participants expressed their appreciation of the program and its importance for the community, emphasizing the need to increase its visibility. CONCLUSION: Few studies have focused on population screening participants' experience. The experience of individuals who participated in the CE/CLE carrier screening program reflects the general satisfaction of participants and the importance of the program for the community. This knowledge will allow the improvement of the program as well as serving as a model for development of other population screening programs, particularly in other Canadian First Nation communities.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1734W

Diagnostic Exome Sequencing in Movement Disorders. *E.J. Kamsteeg[1], C.F.H.A. Gilissen[1], K. Neveling[1], H.J.E. de Reuver[1], B.P.C. van de Warrenburg[2], M.A.A.P. Willemsen[3], S. Vermeer[1], H.G. Brunner[1], M.R. Nelen[1], H. Scheffer[1].* 1) Human Genetics, Radboud University Nijmegen Medical Centre, Nijmegen, Netherlands; 2) Neurology, Radboud University Nijmegen Medical Centre, Nijmegen, Netherlands; 3) Pediatric Neurology, Radboud University Nijmegen Medical Centre, Nijmegen, Netherlands.

Exome sequencing has a great potential in genetic diagnostics, particularly if used for heterogeneous diseases. Sanger sequencing strategies in these heterogeneous diseases are known to have a low diagnostic yield per gene tested. We now have implemented a routine diagnostic exome sequencing workflow for several highly heterogeneous diseases using disease(group)-specific gene filters. Here, the experiences of diagnostic exome sequencing in 50 patients with a movement disorder are reported. Our strategy is to sequence the `whole' exome independent of the diagnostic question, and afterwards apply masking filters for specific disease groups. Over 150 genes are represented in the list of movement disorders, which include forms of ataxia, hereditary spastic paraplegia (HSP) and dystonia. An in-house designed interface was built to display only the sequence variants in those genes, and to prioritize or further select putative pathogenic variants. Using this interface, in the order of 500 sequence variants are typically detected in the genes of the movement disorders list. Subsequent selection based on several parameters, such as frequencies, conservation, type of mutation etc., allows the reduction of variants to about 10 putatively pathogenic variants. The results of a cohort of 50 patients will be presented. QC, cases, and pitfalls will be discussed. With the analysis still ongoing, the movement disorders gene package analysis is expected to provide a molecular diagnosis in ~15% of the patients. So far, mutations were detected in the KIF5A (SPG10) in a large family with a dominant form of HSP, C10orf2/TWINKLE in a patient with an isolated complicated form of ataxia, and KIAA0415 (SPG48) in two siblings with HSP. The results so far indicate that exome sequencing tests are appropriate to identify the underlying molecular defects in movement disorders. Also, some mutations were unexpected, which may imply that the clinical spectra of several disorders may widen now exome sequencing has become diagnostically available.

1735W

Whole exome sequencing as an approach to the diagnosis of the ataxic patients. *Y. Yang[1], M. Vatta[1], Z. Niu[1], Y. Ding[2], H. Sun[1], M. Scheel[2], N. Saada[2], W. Liu[2], M. Wang[2], D.P. Sexton[2], A.C. Hawes[2], M.N. Bainbridge[2], P.A. Pham[2], J.G. Reid[2], D.M. Muzny[2], A.A. Braxton[1], P.A. Ward[1], T.M. Balmakund[3], W.K. Chung[4], A. Willis[1], S.E. Plon[1,5], J.R. Lupski[1], R.A. Gibbs[2], A.L. Beaudet[1], C.M. Eng[1].* 1) Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX; 2) Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX; 3) Arkansas Children'ts Hospital Little Rock, AR; 4) Columbia Univ, Div Molec Gen New York, NY; 5) Department of Pediatrics, Baylor College of Medicine, Houston, TX.

The diagnosis of the patient with suspected hereditary ataxia includes a long list of some common and some very rare conditions that have significant overlap and few distinguishing features. We have recently diagnosed two patients with Autosomal Recessive Spastic Ataxia of Charlevoix-Saguenay (ARSACS, MIM 270550) by whole exome sequencing (WES) testing on a clinical basis. ARSACS is a rare genetic disorder characterized by progressive spasticity and horizontal gaze nystagmus, distal muscle wasting and sensorimotor neuropathy, and dysarthria. A sibling has similar phenotype. The first patient is a 16 year old female with a history of cerebellar atrophy and mild spinal cord atrophy, increased tone in the hamstrings and distal lower extremities, demyelinating changes on EMG, ataxic gait and chronic headaches. WES testing identified two heterozygous novel deleterious mutations in the *SACS* gene, a c.6434T>A (p.L2145X) nonsense mutation and a c.2938_2939del (p.M980fs) frameshift mutation. Sanger sequencing confirmed mutations in the proband and found that the father is a carrier for the c.6434T>A (p.L2145X) mutation and the mother carries the c.2938_2939del (p.M980fs) mutation, indicating that the mutations were in trans-configuration in the proband. The second patient is a 4.4 year old female born to consanguineous parents in the Middle East. The patient presented with delayed motor milestones, spastic ataxia, abnormal gait, and a mild white matter abnormality. A homozygous c.11637_11638delAG (p.R3879fs) frameshift mutation in the *SACS* gene was detected by WES in this patient and confirmed by Sanger sequencing; both parents were found to be heterozygous for this novel mutation. Clinical correlation of the both patients's presentations was felt to be consistent with an ARSACS diagnosis. Both patients had previously gone through extensive workups without a definitive diagnosis. Because of the rarity of the ARSACS disease (less than 400 confirmed cases so far), the phenotypic overlap with other ataxia disorders, and the exclusion of the *SACS* gene from current ataxia gene testing panels, diagnosis of ARSACS can be elusive and costly, a challenge faced by most rare genetic disorders. The diagnosis of ARSACS in these patients demonstrates the potential benefit of WES as a diagnostic approach to patients with previously unsolved rare disorders. Our findings may help determine genotype-phenotype relationships and provide information for enhanced genetic counseling.

1736W

A Retrospective Genetic Analysis of Cases Reported using an XLMR/XLID Next-Generation Sequencing Panel. *M. Parra, S. Mexal, R. Hoiness, K. Waller, I. Lu, C. Gau, J.A. Neidich, J. Wei.* Ambry Genetics, Aliso Viejo, CA.

X-linked mental retardation or X-linked intellectual disability (XLMR/XLID) is a phenotypically and genetically heterogenous disorder that occurs at a prevalence of 1/1000 males. Excluding the CGG repeat expansions in the 5' untranslated region of FMR1 that are associated with Fragile X syndrome, XLMR/XLID has yet to show any major mutational hotspots. At least 90 genes on the X chromosome have been associated with XLMR/XLID, accounting for 10–40% of cases, depending on the inheritance pattern (i.e. sporadic or passed on from obligate female carriers). This study provides a retrospective genetic analysis of the approximately 800 cases reported using the Ambry XLMR Next Gen Sequencing Panel™ since March 2010. This panel is a comprehensive gene sequencing assay for 81 genes associated with X-linked intellectual disability (nonsyndromic and syndromic). Herein we report the mutations identified to date across all cases, as well as in our collection of approximately 380 family studies. Also discussed are variants reclassified as polymorphisms following further family study investigation. Finally, we highlight a case study of a NS-XLMR family, wherein a novel variant in HUWE1 (HECT, UBA and WWE domain containing 1, E3 ubiquitin protein ligase) was found to segregate with the disease in a large family with non-syndromic intellectual disability. Overall we demonstrate here the power of family studies in assisting in reclassification of variants and defining mutations.

1737W

Principles and Informatic Infrastructure for Implementing Diagnostic Next-Generation Sequencing for Genetically Heterogeneous Conditions. *S. Topper, V. Nelakuditi, C. Tan, A. Reeder, D. Del Gaudio, S. Das.* Human Genetics, University of Chicago, Chicago, IL.

The diagnosis of many genetic conditions is complicated by extensive genetic heterogeneity. Phenotypically indistinguishable forms of Intellectual Disability, for example, can be caused by mutations in any one of about 100 known genes, each one of which accounts for a very small percentage of patient cases. In many cases a highly parallelized sequencing approach is a more effective and efficient diagnostic strategy than serial single gene sequencing. But handling the large amount of data generated, while maintaining the essential clinical standards developed in low-throughput tests, can present substantial challenges for the diagnostic lab. Here we present an enumeration of those standards as applied to large-panel, next-generation sequencing (NGS) diagnostic testing, a web-based tool to assist in meeting those standards, and data from validation samples for a set of diagnostic tests. These principles are intended to ensure reliable test results that can be successfully integrated into a diagnostic strategy, especially in the likely case of a negative result, an outcome that is still expected in more than 50% of patient samples. The principles are: 1)Data must be processed and evaluated efficiently and reproducibly; 2)Target regions of essential genes must be evaluated completely; 3)Identified mutations must be confirmed by an orthogonal technology; and 4)Pathogenicity of each identified variant must be evaluated independently and in the context of the patient's entire data set. To ensure that these standards are met we developed a database and web-based interface to facilitate a teamwork approach to sample tracking, variant interpretation, annotated data storage, and integrating NGS and Sanger data. In this tool, an `Overview' module presents data quality statistics, a summary of missed regions in the target genes, and sample tracking information. A `Confirmation' module consists of an interface for coordinating Sanger sequencing of NGS data gaps, variants that must be confirmed, and provides a method for integrating Sanger data into the data set. An `Interpretation' module provides an interface for presenting annotation information, tracking variant review status, maintaining notes, and linking to the data sets of other patients with the same sequence change. Using this system we have evaluated data from 50 validation samples, in some cases identifying pathogenic mutations in samples for which no molecular diagnosis had previously been found.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1738W

Clinical Characterization of Patients Undergoing Molecular Testing for Ehlers Danlos Syndrome at The Hospital For Sick Children in Toronto, Canada: Five Year Experience. *L. Dupuis, M.A. Qoqandi, P. Kannu, L. Fishman, D. Chitayat, S. Bowdin, R. Mendoza-Londono.* Division of Clinical & Metabolic Genetics, The Hospital for Sick Children and University of Toronto, Toronto, ON, Canada.

Introduction: Ehlers-Danlos Syndromes (EDS) is a heterogeneous group of disorders. Molecular testing is only available for a small subset of patients with this clinical diagnosis. The reported yield for molecular testing for mutations in the Collagen V gene in patients that meet criteria for classic EDS is 50%. On average our genetics service evaluates four patients per month with presumed diagnosis of EDS. **Objectives:** To determine the frequency of detectable molecular alterations in patients with clinical diagnosis of EDS investigated by genetic testing and explore the genotype-phenotype correlations with regards to diagnostic yield. **Materials and Methods:** Retrospective chart review of patients undergoing molecular testing between 2007 and 2012. Data collected included basic demographics, anthropometric, Beighton score and clinical findings associated with EDS. Descriptive statistics were calculated from the cohort of index patients. **Results:** A total of 52 individuals from 31 index families underwent molecular testing in the 5 year period. Of these 55% were female (17/31), 45% male (14/31). Average age at evaluation was 18 years (SD 12.5). The average Beighton score (BS) was 5.33 (SD 2.1). Mutations were identified in 29% (9/31) of the subjects. Of these seven mutations affected the Collagen V genes, and one each in Collagen I (*COL1A2*c.304C>T,p.Pro102Ser) and Collagen III (*COL3A1* c.2044G>A,Glu682Lys). Mutations found in Collagen V included one missense mutation (*COL5A1* c.1303 C>G Pro435Ala), one nonsense mutation (*COL5A1* c.1123G>T, Glu375Term), one frameshift mutation (*COL5A1* c.2771delG), one exonic deletion (2.85 kb deletion between IVS1 and IVS2 in *COL5A1*) and 3 splicing mutations (*COL5A2* IVS23+5A>G, *COL5A1* IVS34+4T>C and *COL5A1* IVS54+5C>T). **Conclusions:** In our experience the diagnostic yield for molecular testing in patients with EDS appears to be lower than the one reported in the literature. Detailed clinical characterization of these patients may allow clinicians to perform more targeted genetic testing. Identification of subgroups of patients with consistent phenotypes will facilitate the identification of additional genes responsible for this heterogeneous disorder.

1739W

Detection of oipA using PCR method can help to the diagnosis of *Helicobacter pylori* in the dyspeptic patients. *L. Salimzadeh[1], N. Bagheri[2], F. Azadegan[3], Gh. Rahimian[4], A. Taghikhani[4], M. Hashemzadeh[4], H. Shirzad[4].* 1) Department of Immunology, Molecular Research Center, Shahrekord University of Me, Shahrekord, Iran; 2) MSc of Immunology, Molecular Research Center, Shahrekord University of Medical Sciences; 3) MSc Genetic, Cellular and Molecular Reserch Center, Shahrekord University of Medical Sciences; 4) Cell and Molecular Research Center, Shahrekord University of Medical Sciences.

Background and Aim: Helicobacter pylori has been strongly associated with chronic gastritis, gastric and duodenal ulcers, and it is a important risk factor for gastric cancer. Major virulence factors of H. pylori have been described: the cytotoxin-associated gene product (CagA) and the adhesion proteins BabA2, iceA1, iceA2 and the secretory protein, OipA. H.pylori culture is difficult and CLO test gives various false positive and false negative results. Availability of PCR method for detection of bacterial DNA in biopsy specimens using specific primers implies that detection of one frequent virulence factor can help to the diagnosis H.Pylori. Because of many reports about geographic diversity of the prevalence of H. pylori virulence factors, the second aim of this work was to find their frequency in dyspeptic patients of Shahrekord area of Iran. Methods: Gastric biopsy specimens were taken from 438 patients with gastroduodenal problems during endoscopy. Firstly CLO test and PCR using housekeeping genes, glmM and16srRNA was done. After that, samples that were positive at least for two of mentioned testes were selected and PCR analysis of virulence factors was performed. Results: In this study 189 patients were infected with strains of H.Pylori that at least in two mentioned tests (CLO test, glmM and 16srRNA) were positive. Further PCR analysis of virulence factors showed that the frequency of oipA, iceA1, iceA2, cagA and babA2 were 93.7%, 83.1%, 69.8%, 57.7 % and 38.6%, respectively. Conclusion: In conclusion our results indicated that the oipA virulence factor has most association with CLO test and PCR analysis of H.Pylori housekeeping genes. Therefore oipA can be used as an appropriate auxiliary marker in the molecular diagnosis of H.Pylori infection.

1740W

Connecting family branches through recurrent BRCA mutations. *A. Toland[1], M. O'Connor[2], L. Senter[2,3].* 1) MVIMG/Human Cancer Gen, Ohio State Univ, Columbus, OH; 2) Clinical Cancer Gen, OSU, Columbus, OH; 3) Div Hum Genetics, OSU, Columbus, OH.

Mutational analysis of BRCA1 and BRCA2 has been clinically available since the mid 1990's. To date thousands of individuals have undergone genetic testing of these genes due to personal or family history of breast and ovarian cancer. The Ohio State University Clinical Cancer Genetics program has identified over 390 families with a pathogenic mutation in BRCA1 or BRCA2. All but 10 of the probands resided in the state of Ohio. In our patient cohort we identified 256 different deleterious BRCA mutations of which many are unique. However, excluding the Ashkenazi Jewish founder mutations, we observed 8 deleterious mutations 6 or more times in ostensibly unrelated families. These include BRCA1 mutations C61G (8 families), 3875del4 (7 families), 2800delAA (8 families) and Exon13ins6kb (6 families), and BRCA2 mutations 5910C>G (10 families), 7297delCT (6 families) and 6503delTT (6 families). Some of these are likely due to common distant ancestors or mutational hotspots. We reviewed all 51 pedigrees for these families and based on the information provided, no common ancestor could be identified. We observed another 17 mutations 3–5 times. We compared these mutations to the mutations collected worldwide by CIMBA and identified 9 mutations that were not reported as the most frequent. We hypothesized that some of the rarer recurrent mutations observed in our population may be due to different branches of the same family being tested independently without knowledge of previous testing in their family. To test this hypothesis, we examined 57 pedigrees with shared mutations to look for shared reported family medical history or surnames. In 4 of the 9 groups of pedigrees with mutations in common, links were made despite the fact that individuals were not aware that another branch of their family had been tested. Three of these were related through a more distant relative with a shared surname, but one family had first cousins who were tested without knowledge of the other. It is possible that more of our families with shared mutations are related. As more individuals undergo genetic testing based on family history, we propose that this phenomenon will become more common. It is important to obtain release of genetic information from mutation-positive individuals to their family members for testing purposes. In practice, it would be most efficient and cost effective for genetic counselors to try to link pedigrees prior to testing but this is not always possible.

1741W

Low prevalence of large genomic rearrangements in BRCA1/2 found in at-risk Southern California population. *T. Hoffman[1], M. Alvarado[2].* 1) Medical Genetics, Southern California Kaiser Permanente, Anaheim, CA; 2) Medical Genetics, Southern California Kaiser Permanente, Pasadena, CA.

Large genomic rearrangements appear to represent a small but significant percentage of the overall mutation burden found in the BRCA1/2 genes, and initial data suggests the prevalence of large rearrangements appears to vary somewhat depending on ethnicity. In May 2012, NCCN guidelines were revised to include a recommendation to perform rearrangement studies in all individuals undergoing DNA testing for BRCA1/2. To determine the prevalence of large genomic rearrangements in BRCA1/2 in an ethnically diverse Southern California testing population, we included rearrangement studies ("BART") in all patients undergoing BRCA1/2 testing over a 6 month period with a mutation probability of 10% or higher on at least one of the probability models included in the CancerGene software platform. Patients with a mutation probability below 10% who were felt to be candidates for BRCA1/2 testing were offered standard BRCA1/2 testing without BART testing. Our testing population was comprised of individuals of Western/Northern European (~40%), Latin American/Caribbean (~25%), Asian (~10%), African American (~10%), Ashkenazi (~3%), and Middle Eastern (~3%) ancestry. Among patients tested with comprehensive BRCA1/2 analysis without BART, we found a deleterious sequence mutation in 7.7% (40 individuals out of 520). Among individuals whom BART testing was performed, we found a large genomic rearrangement in 1.6% (3 individuals out of 180). Two of the three individuals found to have a large genomic rearrangement did not meet the "Myriad criteria" typically used by the testing lab to determine when to perform BART testing. Thus, large rearrangements represented about 7% of our total mutation burden. We also retrospectively reviewed our department's cumulative experience with BART testing over a longer period (~5 years) during which rearrangement studies were performed in a nonsystematic fashion due to heightened concern for a lesion in BRCA1/2. We found that the number of large rearrangements identified in cases that did not meet the "Myriad criteria" was low. Furthermore, large rearrangements accounted for ~7% of BRCA1/2 mutations in our patients of Latin American/Caribbean heritage, ~6% in African Americans, ~5% in Western/Northern Europeans, ~2% in Asians and 0% in those of Ashkenazi Jewish heritage. Further data may help identify a more efficient testing strategy to find large rearrangements rather than offering BART studies to all individuals undergoing BRCA1/2 testing.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1742W

Clinical whole exome sequencing for the diagnosis of Mendelian disorders: program design, implementation, and first year reporting experience. *C.M. Eng[1], D. Muzny[2], J. Reid[2], M. Bainbridge[2], P. Pham[2], M.R. Bekheirnia[1], J. Beuten[1], M. Hardison[1], Z. Niu[1], R. Person[1], M. Vatta[1], F. Xia[1], A. Hawes[2], M. Wang[1,2], Y. Ding[1,2], H. Sun[1], M. Scheel[1,2], N. Saada[1,2], W. Liu[1,2], A. Braxton[1], P. Ward[1], A. Willis[1], J. Wiszniewska[1], S.E. Plon[1,3], J.R. Lupski[1], A.L. Beaudet[1], R.A. Gibbs[2], Y. Yang[1].* 1) Dept Molec & Human Genetics, Baylor Col Medicine, Houston, TX; 2) Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX; 3) Department of Pediatrics, Baylor College of Medicine, Houston, TX.

Next generation massively parallel sequencing technologies have entered the clinical arena as a new approach to the molecular characterization of challenging patient phenotypes. In October 2011, the Baylor Human Genome Sequencing Center (HGSC) and the Baylor Medical Genetics Laboratories developed and implemented a CAP/CLIA whole exome sequencing (WES) service with clinical interpretation of variant alleles. WES is performed to a mean depth of >100X coverage >95% target bases greater than 20X coverage), in parallel with a dense SNP array for quality control (SNP concordance to genotype >99.8%) and for detection of copy number variation and regions of absence of heterozygosity, annotation of variants, clinical interpretation, and validation of significant findings by an orthologous method (currently Sanger sequencing). Reporting is done in two phases: a focused report which consists of mutations and variants related to the patient's clinical phenotype, mutations in medically actionable genes, mutations in genes for recessive disorders consistent with carrier status, and selected pharmacogenetic traits. An expanded report includes mutations and variants in genes unrelated to the indication for testing including mutations in genes with no current association with a disease phenotype. To date, 194 samples have been submitted for clinical testing, of which 85% are from pediatric-age patients. The majority of probands have complex neurological phenotypes and had extensive prior genetic testing that has not yielded a specific etiologic diagnosis. Of the first 60 WES datasets analyzed, 16 positive results (27%) consistent with the patient's described phenotype were detected,including mutations in: ARID1B (2 unrelated cases), CASK, CBL, FGFR1/FGF8, GLB1, HIBCH, MECP2, NDUFV1, PTPN11, RBM10, SACS (2 unrelated cases), SMC1A, SYNGAP1, and WDR19. Twenty-three mutant alleles (7 previously reported) in these genes were detected including 7 frameshifts, 3 nonsense, 3 splice site, and 10 missense. In addition to these diagnostic findings, medically actionable variants identified included those causing G-6-PD, Marfan syndrome, and arrthymogenic right ventricular dysplasia. We demonstrate the clinical utility of WES as a diagnostic modality with an initial yield of 27% of previously undiagnosed patients having definitive positive findings as well as early evidence of WES as a cost effective and efficient diagnostic approach for patients and families.

1743W

Diagnostic Yield of Single Nucleotide Polymorphism Array Based Genotyping is Equivalent to Metaphase Cytogenetics for Turner Syndrome. *S.K. Prakash[1], D.C. Guo[1], C.L. Maslen[2], M. Silberbach[2], D.M. Milewicz[1], C.A. Bondy[3], GenTAC Investigators.* 1) University of Texas Health Science Center at Houston, Houston, TX; 2) Oregon Health & Science University, Portland, OR; 3) National Institute of Child Health and Human Development, Bethesda, MD.

Turner syndrome (TS) is a developmental disorder caused by partial or complete monosomy for the X chromosome in 1:2500 females. Metaphase karyotype analysis has been the gold standard for diagnosis of TS but is expensive and lacks molecular resolution. We hypothesized that single nucleotide polymorphism (SNP) array genotyping can provide superior resolution in comparison to karyotyping at a lower cost to facilitate rapid diagnosis and genotype-phenotype correlations in TS. We genotyped 183 adult TS patients (European American, average age 32) using 733,000 SNP markers. All cases met diagnostic criteria for TS based on karyotypes (61%) and/or characteristic physical features including short stature, premature ovarian failure and congenital cardiovascular defects such as bicuspid aortic valves (36%) or coarctation (16%). DNA copy number and percent mosaicism were calculated using intensity and B allele frequency values for 18,239 SNPs along the length of the X chromosome. The SNP array results confirmed the diagnosis of TS in 100% of cases. We identified a single X chromosome (45,X) in 11 cases and mosaic 45,X and 46,X,X or 46,X,Y cell lines in 15 cases. Arrays successfully detected Y chromosomal material in all 6 cases. In 58 additional cases (31%), other mosaic cell lines were present including isochromosomes (14%), rings (5%) and Xp deletions (11%). Our array-based models of X chromosome structure were compatible with karyotypes in 98 of 111 of cases with available data (88%). The principal limitations of SNP data were inability to detect X;autosome translocations (3 cases) and discrepancies in mosaic findings for aberrations that are present in fewer than 10% of cells (8 cases). However, we also identified 4 low abundance mosaic cell lines and 13 large copy number variants in 16 different individuals that were not detected by karyotyping. Our data is the first systematic comparison between the two methods and supports the utility of SNP array genotyping to address clinical and research questions in TS.

1744W

High resolution copy number profiling of the X chromosome in clinical diagnosis. *SA. Yatsenko[1,2], S. Madan-Khetarpal[3], U. Surti[1,2], A. Rajkovic[1,2].* 1) OB/GYN & Reproductive Sci, Univ Pittsburgh, Magee-Womens Hosp., Pittsburgh, PA; 2) Department of Pathology University of Pittsburgh, PA; 3) Department of Medical Genetics, Children's Hospital of Pittsburgh of UPMC, PA.

Whole genome array comparative hybridization (aCGH) has become the first tier test in the clinical diagnosis of patients with intellectual disabilities and congenital anomalies. Copy number variations (CNVs) involving sex chromosomes account for approximately 20% of the overall detected genomic imbalances. The human X chromosome contains a significant number of genes implicated in congenital and metabolic disorders, intellectual disability, disorders of sexual development (DSD), and infertility. However in routine diagnostics by aCGH, small X chromosome deletions and duplications involving an individual gene have remained beyond the detection resolution. To identify intragenic, pathogenic and benign CNVs, we designed a custom high-resolution chromosome X microarray. This microarray contains 180,000 oligonucleotide probes to detect abnormalities, DNA gains and losses, with an average 3–5 kb resolution within the entire X chromosome including pseudoautosomal regions PAR1 and PAR2, and an enhanced 500–2,000 bp resolution for each X-chromosome specific gene, as well as probes for 208 autosomal genes implicated in DSD and ovarian insufficiency. Using this microarray we studied 36 male and 31 female individuals with suspected or known X-linked conditions. Overall we detected 147 copy number changes involving the X chromosome, 16 changes were interpreted as pathogenic, 12 CNVs were of unclear clinical significance, and 119 were considered to be benign CNVs. Pathogenic gains and losses ranging from 16 kb to 19.8 Mb in size were identified in eight patients. These imbalances include heterozygous and hemizygous deletions involving genes *IDS*, *STS*, *DAX1*, *BMP15*, and *POU3F4* which are associated with Hunter syndrome, X-linked ichthyosis, DSD, premature ovarian failure, nonsyndromic deafness, respectively. In addition we characterized large chromosome abnormalities in two patients with an extra mosaic derivative chromosome X and a complex chromosome X rearrangement. We detected 62 recurrent benign CNVs (35 gains and 27 losses) seen in at least two male individuals out of 36 studied. These benign CNVs vary from 475 bp to 251 kb in size, and only ~10% of them involve gene coding regions. The application of high-resolution chromosome X array-CGH for clinical diagnosis brings significant advantages in detection of intragenic aberrations beyond the resolution of clinically available routine aCGH analysis in families with known or suspected X-linked conditions.

1745W

MLPA-based strategy for discrete CNV genotyping: CNV-miRNAs as an example. *M. Marcinkowska[1,2], P. Kozlowski[1,2].* 1) Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznan, Poland; 2) European Centre of Bioinformatics and genomics, Poznan, Poland.

Copy number variation (CNV) in the human genome has become well recognized in recent years. CNVs are genomic regions spanning roughly 1 kb to 1 Mb that show variable numbers of copies in compared genomes. It has been estimated that common CNVs account for about 10% of the human genome, covering hundreds of genes, regulatory sequences and other functional genetic elements. Although substantial progress in genome wide CNV analysis has been recently made there is still need for inexpensive locus specific method allowing discrete (with integer resolution) genotyping of selected CNVs in large number of samples. Such method would greatly facilitate association studies of CNVs with various phenotypes. Recently we have developed multiplex MLPA-based strategy that allows discrete genotyping of almost any CNV. Multiplexed CNVs can be selected either based on localization (e.g. proximity to association signal) or based on prior knowledge about their relation to investigated problem. As a proof-of-concept we developed two MLPA assays covering 17 candidate CNVs that overlap with human miRNA genes (CNV-miRNAs). With the use of developed assays we analyzed 96 samples from three HapMap populations CEU, CHB and YRI and we confirmed polymorphism in 8 out of 17 (47%) candidate CNVs. Among identified polymorphisms were both simple bi-allelic polymorphisms and complex multi-allelic polymorphisms containing up to 9 copies of investigated region. Extensive quality control testing, including (i) reproducibility analysis, (ii) comparison with previously published results and (iii) analysis of agreement with Mendelian inheritance and Hardy-Weinberg equilibrium confirmed very high quality of our results. Almost all polymorphic CNV-miRNAs showed substantial differences in terms of genotypes/alleles frequency and distribution in three human populations. It may suggest that these CNV-miRNAs are functional polymorphisms modifying phenotypes that undergo different selective pressure in various human populations. Summarizing, we have developed and validated the strategy for discrete genotyping of CNVs in complex genomes. The main advantages of this strategy include simplicity of the assay design, flexibility in terms of selection of genomic region and low cost (~$5/sample and <$1/genotype). These advantages make presented strategy attractive for large scale genotyping of individual CNVs required in association studies. Acknowledgement: MNiSW N N302-278937 and NCN 2011/01/B/NZ502773.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1746W

Turkish Medical Students Perceived Genetic Counseling Style and the Use of Hypothetical Genetic Testing Across 5 Psychiatric and 3 Non-Psychiatric Disorders. *D. Kinney[1], E.A. Ergul[2], K. Munir[1], O. Kutuk[1], F. Topuz[1], A. Kanik[1].* 1) Boston Children's Hospital, BOSTON, MA., USA; 2) Massachusetts General Hospital, Boston, MA, USA.

Introduction: Through the advances in genetics, genetic testing has become increasingly relevant to health care. Consumption of these tests is still with debate for psychiatric disorders. The purpose of this study is to evaluate medical students' perceived genetic counseling style as well as the use of predictive genetic tests for 5 psychiatric and 3 non-psychiatric disorders in Turkey. Methods: 206 medical students in Turkey were given questionnaires during class hours. Student response rate was 95%. Results: Students (57%male) were with mean age of 22.77(sd=1.66). Overall 50% of males and 30% of females were in favor of nondirective approach. Except Down syndrome, ADHD and schizophrenia, significant gender difference is observed in counseling style. In autism, 96% of females versus 83%males chose the non-directive counseling (p=0.004) while it was 96% females versus 85% males in Huntington (p=0.008).Higher proportion of males(37%) advised to terminate the pregnancy than females(21%) in CF p=0.049, and in antisocial behavior (4%males versus 15%females, p=0.032). Paradoxically, rate of males advising continuation in CF was also higher(38%) than females,p=0.019. Both in FragileX and Huntington, proportion of males choosing continuation of pregnancy over other counseling styles was higher than females (29% males versus 12% females, p=0.019) in CF and (43% males and 20% females, p=0.005) in Huntington. For hypothetical use of predictive genetic testing, 60% of students approved the test for prenatal population screening in autism, schizophrenia and Huntington whereas 10–15% of students refused the use of this test for any reason. 71% students agreed to use it to identify asymptomatic children with family history of Huntington while 44% students approved it for prenatal population screening for antisocial behavior. Conclusion: Medical students varied greatly in attitudes toward genetic testing and counseling. Most notably, most male students and sizable minorities of female students took directive approaches to genetic counseling, advising pregnant women to either abort or terminate their pregnancy (depending on a particular student's opinion). This approach conflicts with NSGC guidelines, which recommend counselors take a non-directive approach. Turkish Medical school curricula might consider educating students about NSGC guidelines, and discuss ethical issues involved.

1747W

Non-optical massive parallel DNA sequencing of BRCA1 and BRCA2 genes: towards the diagnostic setting. *J.L. Costa[1], S. Sousa[1], R. Fernandes[1], L. Cirnes[1], J.C. Machado[1,2].* 1) IPATIMUP Diagnostics, Institute of Molecular Pathology and Immunology of the University of Porto, Porto, Portugal; 2) Medical Faculty of the University of Porto, Porto, Portugal.

The recent developments of bench top massive parallel sequencers have created new avenues for basic research. However, their implementation in a clinical diagnostic setting remains largely unexplored. In this work we present an optimized work-flow to bring this technology into the genetic diagnostic setting. The breast cancer susceptibility genes- BRCA1 and BRCA2 were used in this assay due to their size, polymorphic character, and lack of mutation hot spots. A total of 40 cases were used and divided into: i) training set with 20 cases with sequences previously determined by Sanger method; ii) validation set of 20 cases of unknown genetic diagnosis whose sequences were determined in parallel using an Ion Torrent Personal Genome Machine (PGM) and an ABI 3130 Genetic Analyzer. Genomic DNA was isolated and the two genes amplified using an in house developed multiplex PCR covering all exons and exon-intron boundaries in a total of 77 amplicons. The various steps of the PGM protocol were optimized to maximize coverage and sensitivity. Barcoding was used to increase the number of cases analyzed per chip. PGM sequencing was analyzed using SeqMan NGen3 and Sanger sequencing using Mutation Surveyor. The training set was used to determine optimal sequencing and analysis conditions for data generated using the PGM. The average PGM run consisted of 340000 reads corresponding to a total of 33 Mbp with the 314 chip, and 2500000 reads corresponding to a total of 150000 Mbp with the 316 chip. To detect all variants known to be present in the training set the variant prioritization was defined using the filters: variant quality score>50, depth of coverage >30x and SNP percentage >30%. These conditions were later used in the validation set. At this stage sequence analysis was carried out blind. The sample by sample comparison showed that all 189 variants detected by Sanger sequencing were also detected by the PGM. However, the PGM sequencing resulted in additional systematic calls. These were found to be false positives lowering the specificity of the methodology. This reduction is tolerated as part of the compromise between the higher capacity and scalability for reducing turn-around-time for diagnosis and the screening nature of this methodology that involves the confirmation of all variants called by the PGM using Sanger sequencing. This work demonstrates the potential for mutational screening of BRCA1 and BRCA2 using the PGM in parallel with Sanger sequencing.

1748W

Identifying the genetic cause of primary immunodeficiency diseases (PID): Development of a next generation sequencing (NGS) approach for routine diagnostics. *S.C. Drury[1], C. Bacchelli[2], S. Bibi[1], F. McKay[1], L. Jenkins[1], H.B. Gaspar[3], C.M. Cale[3], K.C. Gilmour[3], N.J. Lench[1].* 1) NE Thames Regional Genetics Laboratory, Great Ormond Street Hospital, London, London, United Kingdom; 2) GOSgene, UCL Institute of Child Health, 30 Guilford Street, London, United Kingdom; 3) Department of Immunology, Great Ormond Street Hospital NHS Foundation Trust, London, United Kingdom.

Primary immune deficiency diseases (PID) are a heterogeneous group of inherited diseases characterised by variable genetic immune defects, conferring susceptibility to recurrent infection. Though rare, these disorders are chronic and debilitating. Many patients present with similar clinical and laboratory findings and a number of genes are often sequenced sequentially before identifying the molecular defect. This can be time and resource intensive. Due to the rarity of many PID, mutation analysis at DNA level by sequencing is not routinely available. The objective of this study was to develop a robust, time and cost effective NGS method for identifying genetic defects in PID patients as part of a clinical diagnostic service. The aim is to increase the number of PID genes analysed and expedite diagnosis. 35 known PID genes were targeted using the Agilent SureSelect^{TM} Target Enrichment System. In a pilot experiment, DNA capture libraries were prepared from 10 patients, representing 15 known mutations (in *FAS*, *CD40L*, *IL7RA*, *JAK3*, *IL2RG*, *RAG1*, *ADA*, *SAP*, *PRF1* and *RAG2*) from 10 different PID genes and sequenced on the Roche GS Junior and Illumina MiSeq platforms. Data analysis was performed using NextGene (SoftGenetics) software. Using the Roche GS Junior coverage of the known mutations was less than the required minimum of 30x (0x–28x) and included a number of homopolymer errors that made the data difficult to interpret. Results obtained using the Illumina MiSeq demonstrated comprehensive coverage (151x–371x) and simple detection of all known mutations. Preliminary results suggest targeted resequencing using MiSeq will provide a fast and cost effective method for routine screening and characterisation of the genetic basis causing PID. We are also comparing the efficiencies of target capture and sequence coverage using the additional methods of Agilent HaloPlex Target Enrichment and whole exome sequencing (n=10 patients) for the same panel of genes. This will help us develop the most efficient and cost-effective strategy for comprehensive mutation identification for primary immunodeficiencies.

1749W

Exome sequencing analysis for diagnostics. *C. Gilissen, M. Nelen, K. Neveling, R. de Reuver, L.E. Vissers, N. Wieskamp, H.G. Yntema, J. de Ligt, M. Rosario, H.G. Brunner, H. Scheffer, J.A. Veltman.* Human Genetics, Radboud University Nijmegen Medical Center, Nijmegen, Netherlands.

Exome sequencing has great potential in genetic diagnostics, particularly if used for strong heterogeneous diseases were diagnostic yield is currently low. The analysis and interpretation of exome data in diagnostics has specific challenges with respect to quality control, the possibility of incidental findings, and allowing for easy and flexible interpretation. We have implemented a routine diagnostic exome sequencing analysis pipeline that addresses these issues. We have set up pre-defined disease gene packages for seven heterogeneous disorders: Blindness (144 genes), Deafness (98 genes), Movement disorders (152 genes), Oncogenetics (115 genes), OXPHOS diseases (207 genes) and Intellectual Disability (ID) (440 genes). Samples are automatically checked for 1) pre-defined quality control criteria such as enrichment, coverage of the respective disease genes, and concordance with known polymorphisms; 2) matching gender based on sequence coverage and the patient database; 3) Concordance of 12 polymorphisms measured by pyro-sequencing. A diagnostic-oriented graphical user interface automatically limits genetic findings to the relevant genomic loci for a disease, and allows interpretation using predefined filter schemes. In the case of ID samples we sequence the healthy parents and automatically apply an additional de novo analysis; the results of which are combined in the user interface and result in a single overview of potential mutations in known ID genes as well as probable de novo candidates. To date more than 150 patients have been analyzed using this analysis resulting in an average of 487, 397, 388, 288 and 301 variants within the disease gene packages and on average only 11, 6, 7, 6, 4 rare non-synonymous variants for respectively Blindness, Deafness, Movement disorders, Oncogenetics, OXPHOS diseases. In conclusion we have designed and implemented an analysis for exome sequencing based genetic diagnosis of genetically heterogeneous disorders, that allows standardized interpretation and minimizes the risk of incidental findings.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1750W

Detection and mapping the breakpoints of heteroplasmic mtDNA deletions by massively parallel sequencing. *F.-Y. Li, H. Cui, D. Gonzalez, J. Wang, V. Zhang, L.-J. Wong.* Molec & Human Gen, Baylor College Med, Houston, TX.

Background: Large Mitochondrial DNA (mtDNA) deletions have been found to be associated with human diseases including Kearns-Sayre syndrome (KSS), Pearson syndrome, diabetes and hearing loss, myopathy, and progressive external ophthalmoplegia (PEO). Inherited mtDNA multiple deletions are usually secondary due to primary mutations in nuclear genes involved in the maintenance of mtDNA integrity. Regardless of the primary molecular defects, the resultant clinical phenotype may be similar among patients with single or multiple mtDNA deletions. The current methods including Southern blot analysis and PCR/Sanger sequencing detect deletions without being able to map the breakpoints. Objective: To develop an analytical algorithm to map deletion breakpoints from sequence reads data generated from massively parallel sequencing (MPS). Method: The entire mitochondrial genome was amplified with one pair of primer followed by MPS analysis, which provides a deep coverage of >20,000X. With the deep and uniform sequence coverage, a deletion can be easily detected. The unmapped sequence reads were re-aligned to the reference sequence to detect the match with half of the read length. Results: Two exemplary clinical cases: one with a single large heteroplasmic deletion, and the other with multiple deletions were studied. In the first case, a blood specimen from a 10 year old boy affected with encephalopathy, exercise intolerance, easy fatigability, and sensorineural hearing loss was analyzed. A heteroplasmic single large deletion of 7,797 bp (m.7638_15434) was detected. In the second case, a muscle sample from a 70 year old man with myopathy was evaluated. Multiple deletions were identified. The deletion breakpoints of both cases were clearly mapped using our analytical algorithm. The single deletion was confirmed by Southern blot and PCR/Sanger sequence analysis of the breakpoints. In the mutiple deletions case, at least 16 out of 23 junction sequences that were mapped by the algorithms have been confirmed by PCR/Sanger sequencing. Conclusion: These cases demonstrate the utility of MPS analysis in the comprehensive diagnosis of mitochondrial DNA disorders, particularly the simultanious detection and mapping of mtDNA deletions.

1751W

Three novel AGL mutations in glycogen storage disease type III in Tunisian families. *A. MILI[1, 2], O. MAMAI[2], I. BEN CHARFEDDINE[2], A. AMARA[2], S. PAGLIARANI[3], S. LUCCHIARI[3], A. SAAD[2], K. LIMEM[1], M. GRIBAA[2].* 1) Biochemistry Department, Faculty of Medicine, Sousse, Tunisia; 2) Laboratory of Cytogenetics, Molecular Genetic and Human Reproduction Biology, Farhat University Hached Hospital, Ibn ElJazzar Street, 4000 Sousse, Tunisia; 3) 3Dino Ferrari Center, Department of Neurological Sciences, University of Milan, Fondazione I.R.C.C.S. Ospedale Maggiore Policlinico, Mangialli, Regina Elena, Milan, Italy.

Glycogen-storage disease type III (GSD III) is an autosomal recessive inborn error of metabolism caused by mutations in the glycogen debranching enzyme amylo-1,6-glucosidase gene, which is located on chromosome 1p21.2. GSD III is characterized by the storage of structurally abnormal glycogen, termed limit dextrin, in both skeletal and cardiac muscle and/or liver, with great variability in resultant organ dysfunction. The spectrum of AGL gene mutations in GSD III patients depends on ethnic group. The most prevalent mutations have been reported in the North African Jewish population and in an isolate such as the Faroe Islands. Here, we present the molecular and biochemical analysis of 25 Tunisian GSD III patients. Molecular analysis revealed three novel mutations: a nonsense (Tyr1148X) and two deletions (3033_3036del AATT and 3216–3217del GA) and five known mutations: three nonsense (R864X, W1327X, W255X), a missense (R524H) and a acceptor splice site mutation IVS32-12A>G.In adition, Two high frequency mutations The discovery of this founder mutation in center east of Tunisia is great importance since a test based on haplotype analysis and sequencing live Will Be Utilized to determined carrier status and prenatal diagnosis, genetic counseling for Improving Malthus GSD III in Tunisia. In addition, the discovery of two mutations at high frequency which shows the presence of a founder effect which is of great importance as a test based on haplotype analysis and direct sequencing will be used to determine carrier status and prenatal diagnosis, improving genetic counseling for GSD III Tunisia. This is the first report of screening for mutations of AGL gene in the Tunisian population.

1752W

Multiplex PCR assay for detection of the 2 most common mutations in the profilaggrin gene (FLG) involved in ichthyosis vulgaris. *G. Pont-Kingdon[1,3], L. Hubley[1], E.E. Baldwin[2], E. Lyon[1,3].* 1) Institute for Clinical and Experimental Pathology, ARUP Laboratories, Salt Lake City, UT; 2) ARUP Laboratories, Salt Lake City, UT; 3) Pathology Department, University of Utah, Salt Lake City, Utah.

Filaggrin is a major protein involved in the formation of the cytoskeleton of the cornified layer of the skin, establishing a barrier that prevents entry of allergens and microbes through the skin, regulates pH, prevents water loss, and possibly provides UV protection. In 2006, McLean's group reported that null mutations in the gene encoding profilaggrin (FLG) are responsible for ichthyosis vulgaris (IV), a common skin condition characterized by dry and scaly skin. IV has an incidence of 1/400-1/250 in the Caucasian population, but has been described in all ethnic groups. Additional research has shown an association between FLG-null mutations and atopic dermatitis (AD), asthma associated with AD, allergic rhinitis, some food allergies, keratosis pilaris, and palmar hyperlinearity. The FLG gene is located on chromosome 1q21 within the epidermal differentiation complex. It is approximately 25 kb long, has 3 exons and contains 10–12 tandem repeats that are each 972 nucleotides long. These repeats show nearly identical nucleotide sequence. Each repeat encodes the 324 amino acid monomer of the filaggrin protein. We designed a hybridization probe assay to detect the two most common mutations in FLG found in Caucasians of European descent: p.R501X (c.1501C>A) and c.2282del4. Both mutations are found in FLG-unit 1. This assay is multiplexed and uses two primer sets with each primer being specific for FLG-unit 1. Appropriate PCR conditions were achieved for the specific amplification of FLG-unit 1. The test was validated using 113 samples with various genotypes verified by Sanger sequencing. Genotypes obtained by melting curve analysis corresponded completely to genotypes observed by sequencing. Variation in the assay, obtained by melting temperature analysis of triplicates (within run) and five independent replicates (between) was less than 0.25° C. This assay should detect in a single reaction up to 80% of individuals of Caucasian descent harboring an FLG mutation.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1753W

***GeneReviews*™ as a source of medically actionable information for coincidentally diagnosed disorders.** *M.P. Adam, C.R. Dolan, A.R. Amemiya, M. Espeseth, T.D. Bird, K. Stephens, R.A. Pagon.* University of Washington, Seattle, WA.

PROBLEM: The clinical use of whole-exome sequencing (WES) to identify the genetic basis of undiagnosed findings in a patient will inevitably identify "coincidental findings" (i.e., pathogenic mutations in genes not relevant to the features that prompted testing by WES). Some have advocated sharing with patients and their health care providers coincidental findings so that patients can benefit from relevant medically actionable information. HYPOTHESIS: *GeneReviews*, an expert-authored, peer-reviewed resource developed and maintained at the University of Washington and freely available at www.genetests.org, can be used to determine if medically actionable information exists for a coincidentally diagnosed disorder. METHODS: First, we defined medically actionable information as that which addresses: (1) the evaluations (laboratory studies, imaging studies, or specialty referrals excluding medical genetics referral) needed to determine the presence and extent of manifestations of a disorder; (2) disease-specific treatment; (3) primary prevention; and/or (4) contraindicated agents/circumstances. Second, 276 of the existing 555 *GeneReviews* chapters were reviewed. We included only those that deal with disorders that could be coincidentally diagnosed by WES by eliminating chapters that deal with phenotype-based overviews (7), disorders diagnosed exclusively at birth (63), and disorders for which the genetic basis is unknown (1) or not detectable by WES (e.g., repeat expansions or microdeletion syndromes) (18). As a result, 187 chapters (representing 365 genes) were included in this study. FINDINGS: (1) Further evaluations (including laboratory studies, imaging studies, and specialty referrals) were recommended to determine the extent of manifestations in 162 disorders; (2) disease-specific treatment existed for 64 disorders; (3) primary prevention existed for 63 disorders; and (4) contraindicated agents/circumstances were identified for 123 disorders. Of the 187 *GeneReviews* chapters meeting criteria for inclusion in this study, 176 were found to have information considered to be medically actionable. CONCLUSION: As WES and whole-genome sequencing move into mainstream medicine, *GeneReviews* will be useful for identifying those coincidentally diagnosed disorders for which medically actionable information exists and for clinicians caring for patients with these disorders.

1754W

Comprehensive diagnostic hearing loss testing on a single platform: The OtoGenome test. *SS. Amr[1, 2], TJ. Pugh[1, 2], E. Duffy[1], L. Farwell[1], S. Gowrisankar[1], A. Lovelette-Hernandez[1], BH. Funke[1], HL. Rehm[1].* 1) Laboratory for Molecular Medicine, Partners Healthcare Center for Personalized Genetic Medicine, Cambridge, MA; 2) Harvard Medical School Molecular Genetics Training Program, Boston, MA, United States.

To date, the detection of the genetic cause of hearing loss has been challenging due the large number of responsible genes, many of which, have not been available for clinical testing. To solve this problem, we have developed the OtoGenome Test, a next-gen sequencing (NGS) assay that covers all 73 known genes for nonsyndromic hearing loss, Usher, Pendred, and Jervell Lange-Nielsen syndromes. The test utilizes molecular barcoding and Agilent in-solution hybrid capture followed by NGS on the Illumina HiSeq. Molecular barcoding allows for pooling of up to 10 samples in a single capture reaction, which reduces test cost. For sequence alignment and variant detection, our in-house informatics pipeline performs read alignment and variant detection using BWA followed by GATK. In addition, a novel copy number variant (CNV) tool has been developed to detect CNVs. For validation, we selected 61 patient samples with known variants to assess analytical performance. Average read depth was 854 reads providing robust variant detection. All known variants were detected including 77 substitutions, 26 indels ranging in size from 1 to 21 bp, and 8 CNVs at 6 different loci. We selected 10 additional cases that had previously tested negative or inconclusive by our OtoChip test (covering 19 genes), to enable a preliminary estimate of clinical sensitivity of the additional 54 genes. This approach allowed us to identify biallelic variants in 3 out of 10 samples (3 CNVs, 1 indel, and 1 missense variant), and monoallelic pathogenic frameshift variants in two samples. The total number of clinically significant variants identified in this cohort is 3 CNVs, 5 indels, and 12 single base substitutions. The detection rate of the OtoGenome is expected to surpass our current OtoChip detection rate, which detects a biallelic cause of hearing loss in approximately 18% of cases based upon data from over 325 cases. The OtoGenome test is expected to increase detection of biallelic cause of hearing loss to as high as 70%.

1755W

Allele drop out and *MECP2* genetic testing. *E. Bettella, R. Polli, E. Leonardi, A. Murgia.* Pediatrics, University of Padua, Padua, Italy.

Allele drop out (ADO) leads to the preferential amplification of one allele, in a heterozygous individual, due a series of reasons among which the occurrence of sequence mismatch within a primer-binding site or a complex DNA motif. The drop out can cause false results during PCR amplification and be responsible of potential misdiagnosis. Recent evidences suggest that DNA sequences that fold in particularly complex secondary structures can hamper DNA replication and transcription and, if altered by mutations, may be responsible of ADO events in PCR amplification. An example of these genomic DNA motifs are repetitive guanine-rich sequences, known as G-quadruplexes, on one stand and the complementary cytosine-rich sequences on the opposite strand, known as i-motifs. We report a case of ADO by which two previously unreported in cis de novo mutations in exon 4 of the *MECP2* gene (c.1137delC and c.1151_1201del50) were detected as seemingly homozygous in a girl with Rett syndrome. Both mutations are located in the *MECP2* WW binding domain and fall in a "frameshift cluster" region where a number of other small deletions have been reported in literature. Under the assumption of ADO, we and adjusted the primer set and the PCR conditions so that we could distinguish two alleles and confirm the presence of a wild type allele. This report confirms that allele drop out is a definite technical risk in *MECP2* molecular analysis; failure in allele amplification may occur in particular for exon 4, due to the complexity of its sequence.

1756W

Screening for and Identification of Pathogenic Mutations in Patients with Inherited Retinal Degenerations (IRDs) Using Targeted Enrichment and Next Generation Sequencing. *M.B. Consugar[1], Z.D. Fonseca-Kelly[1], E.M. Place[1], E. Au[1], S.M. Harper[2], X. Gai[3], E.L. Berson[2], E.A. Pierce[1,2].* 1) Ocular Genomics Institute, Department of Ophthalmology, Massachusetts Eye and Ear Infirmary and Harvard Medical School, Boston, MA, 02114; 2) Berman-Gund Laboratory, Department of Ophthalmology, Harvard Medical School and Massachusetts Eye and Ear Infirmary, Boston, MA, 02114; 3) Department of Molecular Pharmacology and Therapeutics, Loyola University Health Sciences Division, Maywood, IL, 60153.

Inherited Retinal Degenerations (IRDs) are a major cause of vision loss, affecting the photoreceptor cells and pigment epithelium of the retina. Complicating matters, the light-sensitive outer segments of photoreceptor cells are modified primary cilia, and many IRDs are genetically heterogeneous and can be associated with other cilia disorders. To date, mutations in 186 genes have been linked to IRDs, yet account for only 50% of clinically diagnosed cases. Although part of the low diagnostic rate can be attributed to unidentified disease genes, another contributing factor is how IRD patients are genetically tested. Because a majority of diagnostic laboratories perform gene-, exon- and/or mutation-specific testing based on clinical diagnoses, de novo mutations and mutations in genetically heterogeneous genes can result in missed genetic diagnoses. Next-generation sequencing (NGS), combined with targeted enrichment, has become an attractive method for mutation screening in a fraction of the time associated Sanger sequencing. In order to increase the fidelity of IRD genetic diagnoses, we designed a SureSelect targeted exon + UTR enrichment library specific to the 186 known, as well as another 20 candidate IRD disease genes. Following capture, samples were sequenced on an Illumina MiSeq, with average per sample coverage metrics for a 10-sample multiplexed, 2 × 110 bp paired-end run, being: 95x mean Depth-of-Coverage (DoC); 99.5% coverage at 1x DoC; 97.0% at 10x DoC; 93.9% at 20x DoC. To date, 54 probands clinically diagnosed with autosomal recessive Retinitis Pigmentosa (arRP) or Leber Congenital Amaurosis (LCA) have been screened for pathogenic IRD mutations using our targeted enrichment/NGS method. We were able to assign likely genetic diagnoses to 21 patients (~39%), including one case where we identified mutations in two genes likely responsible for the patient's vision and hearing loss, respectively. These results demonstrate the feasibility and advantages of this approach for genetic diagnostic testing of IRD patients. An important advantage of testing for mutations in all IRD disease genes is identifying pathogenic mutations in genes not typically associated with patient phenotypes. Further, identification of patients without mutations in known IRD disease genes could facilitate novel IRD disease gene discovery research. Finally, accurate genetic diagnoses could facilitate application of gene-based therapies to prevent further vision loss.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1757W

Validation of Whole Exome Sequencing assay for clinic service at Baylor College of Medicine. *Y. Ding[1], D.M. Muzny[1], J.G. Reid[1], M. Wang[1], Y. Han[1], H. Dinh[1], D.P. Sexton[1], M.N. Bainbridge[1], A.C. Hawes[1], P. Pham[1], B. Yin[1], D.P.C Ng[2], C.J. Buhay[1], J. Wiszniewska[2], M. Scheel[1], N. Saada[1], W. Liu[1], H. Sun[2], R.A. Gibbs[1], A.L. Beaudet[2], C.M. Eng[2], Y. Yang[2].* 1) Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX 77030; 2) Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030.

The Whole Genome Laboratory (WGL) at Baylor College of Medicine (BCM) is a joint venture of the Human Genome Sequencing Center (HGSC) and the Medical Genetics Laboratory (MGL) at BCM to provide genetic testing by high quality next generation sequencing (NGS). The whole exome sequencing (WES) utilizing capture probe set VCRome 2.1 (HGSC exome design) coupled with Illumina Barcoded Paired-End NGS on the Genome Analyzer II/HiSeq platform has been successfully developed and validated as the initial clinical service at WGL. This lab is currently a CLIA (the Clinical Laboratory Improvement Amendments of 1988)-registered and CAP (College of American Pathologists)-accredited clinical laboratory. The validation processes included platform (Ilumina GAII and Hiseq 2000) validations followed by WES assay validation. Eleven samples with known point mutations and small insertions or deletions (indels) previously identified by Sanger sequencing were included in the assay validation. Two independent parallel runs were performed for inter-run reproducibility assessment. Illumina high density SNP assay was performed in each run for sample identification and quality control purposes. Sequencing data was analyzed through the Mercury analysis pipeline developed at HGSC. Our validation results showed that the capture specificity was ~70% and 95% of the target bases were covered 20X or higher. A concordance of about 99% between WES and SNP array data was achieved. About 70% of the VCRome 2.1 genes (approximately 200,000) were fully covered and 85% were covered at least 95%. Overall, WES sensitivity and specificity were 100% and 98% respectively for point mutations, and 85% and 27% respectively for indels. The low specificity for indels was caused by the less stringent quality filtering criteria applied to the short read NGS data in an effort to maximize the sensitivity of the indel calls. A reproducibility rate of about 98% on two parallel runs was achieved. Our results have indicated that our WES assay has good reproducibility, high sensitivity and specificity for point mutation detections. WES in combination with the Sanger sequencing can achieve a specificity of 100% for point mutations and indels. Through targeted capture and high throughput sequencing technologies, WES represents a cost-effective and robust NGS approach to detect variants in the coding regions of the human genome.

1758W

Advancing from targeted resequencing to whole exome sequencing for improving the molecular diagnosis of congenital disorders of glycosylation. *M.A. Jones[1], B.G. NG[2], P. He[2], M.E. Losfeld[2], S. Bhide[1], D. Rhodenizer[1], E.L.H. Chin[1], M. He[1], H.H. Freeze[2], M.R. Hegde[1].* 1) Emory Genetics Laboratory, Atlanta GA; 2) Sanford Burnham Medical Research Institute, La Jolla CA.

Congenital disorders of glycosylation (CDG) are comprised of over 30 autosomal recessive disorders caused by deficient N-glycosylation. Approximately 20% of patients do not survive beyond five years of age due to widespread organ dysfunction. A diagnosis of CDG is based on abnormal glycosylation of transferrin but this method cannot identify the specific gene defect. In 2010 Emory Genetics Laboratory introduced molecular testing for 25 CDG genes including single gene testing and next generation sequencing (NGS) panels. A total of 118 patients have been referred for testing. Disease causing mutations have been identified in several genes including PMM2, COG1, COG8 and ATP6VOA2. Mutations identified in the gene ALG1 are the most common in patients referred for testing. To date, molecular diagnostic testing identified the likely gene defect in only 10% of individuals. This is likely due to patients having mutations in new genes. In the research setting we are using a combined biochemical and whole exome sequencing (WES) approach for a collection of 39 patients that remain without a molecular diagnosis to identify the causative gene. Our WES analysis pipeline focuses on a library of 250 glycosylation associated genes. Through WES we have identified mutations in two previously unknown genes. We identified one patient with mutations in DDOST and five patients with CHIME syndrome, which has clinical features similar to SRD5A3-CDG that had mutations in PIGL. We will summarize the results of WES for the remaining 33 patients currently under analysis. We will also discuss the clinical utility of WES for diagnostics and our approach to addressing incidental findings in a severe pediatric disorder.

1759W

A novel splice site variant (IVS44-3T>C) in *COL3A1* results in small amounts of exon skipping and complicates genetic counseling in a family with features of hypermobility Ehlers-Danlos syndrome. *D.F. Leistritz[1], U. Schwarze[1], D. Sarroza[1], P.D. Steele[2], C. Boni[2], B.A. Salbert[2], P.H. Byers[1].* 1) University of Washington, Seattle, WA; 2) Geisinger Medical Center, Danville, PA.

In fibrillar collagen genes the -3 position of splice acceptor sites is almost always C or T, rarely A, and never G. We studied a 25-year old woman with multiple joint dislocations and hypermobility and found that she had a sequence variant in one *COL3A1* allele: c.3202-3T>C, IVS44-3T>C. The T at this position is conserved in 25 of 38 vertebrate species and 12 have C in this site. It is not known to be polymorphic or variable in humans. To our surprise, analysis of splicing in her cultured fibroblasts demonstrated that in about 5-10% of *COL3A1* transcripts exon 45 was skipped. This skip results in a small amount of proα1(III) chains that lack 18 amino acid residues from the main triple helical domain. About 15% of type III procollagen molecules would be abnormal if 5% of the chains lacked the 18 amino acids, but we could not detect these molecules during gel electrophoresis. Ehlers-Danlos syndrome type IV, the vascular type, results from mutations in which all of the products of one allele are abnormal. This means that 85% of the type III homotrimers have an abnormal structure. We have identified mosaicism for a *COL3A1* mutation such that about 10% of alleles or more are altered. These individuals are generally not symptomatic and do not have hypermobility. However, mosaicism results in discontinuous patches in which the altered allele is expressed. In contrast, we expect that in this family most cells would have the abnormal splicing pattern. We identified 5 additional members with the variant, all of whom had a similar clinical picture to that seen in the proband. There is one unrelated family with a heterozygous mutation in *COL3A1* that was found to have only joint hypermobility. The underlying basis of hypermobility is not known in most individuals. These findings raise the possibility that certain classes of mutations in *COL3A1* could explain the clinical findings in a small subset of families. In addition, the production of a small amount of abnormal product—enough to potentially lessen the integrity of tissues, but not enough to be disease causing—creates uncertainty and complicates genetic counseling for the proband and those relatives identified with the same *COL3A1* sequence alteration.

1760W

A DNA-based method for detecting homologous blood doping. *I. Manokhina, J.L. Rupert.* School of Human Kinetics, University of British Columbia, Vancouver, BC, Canada.

Homologous (or allogenic) blood doping, in which blood is transferred from a donor to a recipient athlete, is an easy, inexpensive and rapid way to increase the oxygen carrying capacity of blood and therefore endurance performance; however, the method can be detrimental to health and is prohibited by doping control authorities. Current tests to detect homologous blood doping are based on cell surface markers. We propose that a DNA-based test centered on the allele-specific amplification of common insertion-deletion polymorphisms would be more convenient and less expensive than protein-based tests. Methods: We adapted and optimized a highly sensitive allele-specific qPCR test panel of nine polymorphic insertion/deletion markers. Mixed DNA samples prepared from blood dilutions ranging from 1/10 to 1/10,000) were used to determine the sensitivity of the assay. These ratios simulate cases of doping with 500 ml of either whole blood or 90% - 99.9% leucodepleted blood. The informativeness of the assay was tested using blood mixtures from 12 unrelated individuals and determining what percent of the 132 potential combinations could be detected by one or more of the 9 markers. Preliminary tests of the specificity of the assay were done using DNA samples from specific populations who have the potential to be natural chimeras, including mothers (n=7) and dizygotic twins (n=2). Results: The sensitivity of the assay based on detection limit for the secondary blood cell population was found to be in the range of 1/1000-1/10000 dilution for all studied markers. The informativeness of the assay, calculated by the ratio of potentially detectable `informative' blood mixtures to all the blood samples obtained (132), was 99.2%. Amplification of potential chimeric samples from mothers and dizygotic twins did not reveal specific alternative PCR products, suggesting that the secondary cell populations, if they exist, do not reach the threshold of assay detection. Conclusion: We propose that a DNA-based test for homologous blood doping could detect homologous doping even if the doper used blood from which 99% of the leukocytes had been removed. Such a test could supplement or replace the current protein-based tests. Even if the test was not routinely applied, the knowledge that DNA can be stored and evaluated using this method long after the sample was taken would be a major disincentive to athletes considering homologous blood doping.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1761W

Identification of *ARID1B* mutations in intellectual disability/Coffin-Siris patients by clinical whole exome sequencing analysis. *Z. Niu[1], M. Vatta[1], A. Willis[1], Y. Ding[2], H. Sun[1], M. Scheel[2], N. Saada[2], W. Liu[2], M. Wang[2], D.P. Sexton[2], A.C. Hawes[2], M.N. Brainbridge[2], P. Pham[2], J.G. Reid[2], D.M. Muzny[2], P.A. Ward[1], A.A. Braxton[1], A. Scheuerle[3], M.K. Koenig[4], S.E. Plon[5], J.R. Lupski[1], R.A. Gibbs[2], A.L. Beaudet[1], Y. Yang[1], C.M. Eng[1].* 1) Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX; 2) Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX; 3) Tesserae Genetics, Dallas, TX; 4) Dept. of Pediatrics, Division of Child and Adolescent Neurology, The University of Texas Medical School at Houston, Houston, TX; 5) Department of Pediatrics, Baylor College of Medicine, Houston, TX.

Molecular diagnosis of developmental disability, especially for the autosomal dominant type, faces multiple challenges of a fast-growing number of causative genes and a substantial involvement of *de novo* genetic defects. Recent efforts using whole genome or whole exome sequencing have significantly accelerated mutation discovery in developmental disability by screening large amounts of genomic sequence simultaneously at a reasonable cost. Supported by these considerations, clinical whole exome sequencing (WES) analysis provides a favorable diagnostic approach to identify the etiological factors of developmental disability in particular for difficult cases where multiple lines of testing have not identified the causative gene.

Here, we discuss two cases of unrelated patients with developmental disability, where prior comprehensive aCGH array and single gene tests failed to identify an underlying genetic cause. Both patients presented with motor and speech delay, hirsutism, and facial dysmorphism including low and posteriorly rotated ears, downslanting palpebral fissures, and mild coarsening of the facial features. Patient one displayed profound intellectual disability and is non-verbal, while patient two can speak three word phrases and read at first grade level. Brain imaging results were normal for these two patients. Clinical WES analysis performed in the Baylor Genetics Laboratory identified heterozygous loss-of-function mutations in the *ARID1B* gene: a nonsense (p.Q1190X) mutation in patient one and an exon/intron boundary disruption (c.2981_2986+7del13) in patient two. Sanger sequencing analysis on the parental samples demonstrated that these two mutations arose *de novo*. *ARID1B* (also known as *BAF250B*), is a component of SWI/SNF nucleosome remodeling complex. Disease causing mutations in SWI/SNF complex genes (*ARID1A, ARID1B, SMARCB1, SMARCA4, SMARCA2 and SMARCE1*) have been recently discovered in intellectual disability/Coffin-Siris patients (PMID: 22405089, 22426308 and 22426309). To date, most of *ARID1B* mutations in intellectual disability/Coffin-Siris patients are *de novo*. Clinical correlation of both patients' presentations was consistent with an *ARID1B* diagnosis.

In summary, clinical WES analysis identified *de novo ARID1B* mutations consistent with the patients' clinical phenotype.

1762W

National Institutes of Health (NIH) Genetic Testing Registry. *W.S. Rubinstein[1], D. Maglott[1], B.L. Kattman[1], A. Malheiro[1], J. Lee[1], M. Ovetsky[1], V. Hem[1], G. Song[1], V. Gorelenkov[1], C. Fomous[2], J. Ostell[1].* 1) National Center for Biotechnology Information, National Institutes of Health, Bethesda, MD; 2) Office of the Director, National Institutes of Health, Bethesda, MD.

The NIH Genetic Testing Registry (GTR, www.ncbi.nlm.nih.gov/gtr), an online tool designed to help clinicians navigate the landscape of genetic tests, was unveiled by NIH Director Dr. Francis Collins on Rare Disease Day (February 29, 2012). The GTR arose from the 2008 recommendation by the Secretary's Advisory Committee on Genetics, Health and Society (SACGHS) that called for a publicly available, Web-based registry to enhance the transparency of genetic testing and assist efforts in reviewing the clinical validity of laboratory tests. The GTR was developed by the National Center for Biotechnology Information (NCBI) under the oversight of the NIH Office of the Director and with extensive input from health care providers, researchers, testing labs, patients, and other stakeholders. The GTR contains information about biochemical, cytogenetic, and molecular tests for Mendelian disorders and drug responses. Submissions, which NCBI began accepting in the spring of 2012, are voluntary. The test content is a combination of fully registered tests submitted to GTR, information displayed from the GeneTests Laboratory Directory, and information supplemented by NCBI. *GeneReviews* and GeneClinics are accessible through the GTR and will remain available after the GeneTests Laboratory Directory is phased out in 2013. Fully registered GTR tests are identified by a unique GTR accession and version number and have a detailed description provided by the lab. Test providers are responsible for the content and quality of the data they submit. NIH requires submitters to agree to a code of conduct which includes providing information that is accurate and not misleading. GTR users are requested to report any breaches of the code of conduct. Test information is presented alongside well-vetted publications and resources (e.g., practice guidelines) assembled by NCBI, which enables the professional community to identify information gaps and to focus their efforts accordingly. The accessioning of well-described genetic tests facilitates formal evaluation by professionals and publication of these assessments. NCBI has encouraged the American Society of Human Genetics, American College of Medical Genetics and Genomics, Association for Molecular Pathology, and National Society of Genetic Counselors to publish evaluations of the tests in the GTR. Professional evaluations can be displayed in the GTR as with other important publications and would fulfill the charge articulated by SACGHS.

1763W

EuroGentest Clinical Utility Gene Cards - progress and perspectives. *J. Schmidtke, A. Dierking.* Dept Human Gen, Medizinische Hochschule, Hannover 61, Germany.

CUGCs are disease-specific guidelines for the clinically validated use of genetic testing. Each CUGC is authored and peer-reviewed by international experts and published by the European Journal of Human Genetics (EJHG). CUGCs are a work package within EuroGentest, a EU-funded project for the harmonization of genetic testing. Based on the ACCE framework (Analytic validity, Clinical validity, Clinical utility and Ethical, legal and social issues) the CUGCs are dealing with the risks and benefits of genetic test application. Besides analytical and clinical validity, the clinical utility is a significant aspect to define. It is of importance to evaluate a genetic test in the context of the clinical setting and to realize that the laboratory test is only one component of the overall intervention. Due to their concise format CUGCs offer quick guidance to all stakeholders, including clinicians, clinical geneticists, referrers, service providers and payers. All documents are freely available on the websites of EuroGentest, the EJHG, the European Society of Human Genetics and Orphanet and will be revised annually. The analysis of the EJHG download rates shows that between 600 and 1,500 downloads, with an average above 1,000, were counted per gene card and year. Until the end of 2013 at least 300 CUGCs are planned, so far 65 guidelines are published and about 130 are in progress. The first set of CUGC updates is already submitted to the journal and will be published soon. Hampering factors are for example the limited number of experts in the field of rare diseases and the long process from author invitation to completed review. Nevertheless, the current progress and the feedback from the scientific community are positive.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1764W

Mutation Screening of Genes Causing Inherited Immunodeficiency and Bone Marrow Failure using Next-Generation Sequencing. *T.A. Sivakumaran, A. Husami, D. Kissell, K. Zhang.* Division of Human Genetics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH.

Advancements in high throughput sequencing, also known as Next-generation sequencing (NGS), have facilitated the discovery of new disease genes and also the transformation of routine clinical diagnosis of genetic disease. Targeted resequencing using NGS technology offers a unique opportunity to screen for mutations in multiple genes, in parallel, causing heterogeneous genetic disease at lower cost and time. Several recent studies have shown NGS an efficient tool for mutation screening in the diagnostic laboratories. In this study, we evaluated the performance of NGS in detecting the DNA sequence variants in 124 genes causing inherited immunodeficiency and bone marrow failure. Sequence covering the coding, the flanking intronic and regulatory regions of these genes were enriched using RainDance microdroplet PCR and then sequenced on the Illumina HiSeq 2000 sequencer. The raw sequence reads from each sample were subjected to the molecular genetics laboratory's (MGL) NGS analysis pipeline: (1) quality filtering of raw reads, (2) mapping of high quality reads to the Human reference sequence and detecting sequence variants using NextGENe software, (3) finding target bases with < 40x coverage, (4) identifying potential true sequence variants using MGL criteria and (5) confirmation of variants with Sanger sequencing. An average of about 30 million 50 bp reads, ranging from 15,708,900 - 51,719,428, were generated in 21 deidentified samples that were previously screened for mutations in one or more of these genes in the MGL. About 89–95% of the raw sequence reads met our quality criteria in all 21 samples and about 70% of quality-filtered reads aligned to target genes. Of the 591 kb target sequence covering coding and 20 bp of flanking intronic regions, about 97.5% of bases had greater than 40x coverage. In the preliminary analyses of thirteen samples, NextGENE software detected a total of 4720 sequence variants and about 50% of which were classified as potentially true based on MGL criteria. Comparison of NGS results with the available Sanger sequencing results showed about 100% concordance. The results of this study will be presented. In summary, we anticipate that the NGS-based mutation screening of these genes would simplify the genetic workup of an individual with immunodeficiency or bone marrow failure, enabling the effective clinical management of patients and their families.

1765W

Mutation analysis of the SOX2 and OTX2 genes in 235 patients with developmental eye disorders. *N. Smaoui, B. Williams, S. Aradhya, S. Bale, S. Suchi, G. Richard.* GeneDx, Gaithersburg, MD.

Microphthalmia and anophthalmia (MAC) are rare congenital developmental eye disorders, with a prevalence of approximately of 0.5–1.5 per 10.000 birth. MAC can be either isolated or syndromic and is genetically heterogeneous with different modes of inheritance. At least 9 genes are known to cause MAC including SOX2, OTX2, VSX2 (CHX10), PAX6, CHD7, BCOR, RAX, PITX3, and STRA6. SOX2 mutations are most common and are found in up to 20% of MAC. We report mutation analysis of the SOX2 and OTX2 genes in a large cohort of patients with MAC and various developmental eye disorders. Sequence analysis of the SOX2 gene was performed for 235 patients, and additional deletion/duplication testing by exon-level array CGH was completed on 66 samples. Sequence analysis was positive in 13% of samples, while 2 out of 66 sequencing-negative cases had a deletion of the entire SOX2 gene (3%). Most common were mutations introducing a premature stop codon (84%). Two mutations (c.70_89del20 and Y160X) were recurrent, each identified in 16% of unrelated cases. The phenotype was very variable, including isolated unilateral or bilateral MAC, or MAC associated with other developmental eye disorders or systemic malformations. In one familial case, microphthalmia and ocular coloboma was reported in the proband while the father presented with isolated ocular coloboma. Sequence analysis of the OTX2 gene was positive in 4 out 66 (6%) samples that tested negative for a SOX2 mutation. Deletion/duplication analysis of the OTX2 was performed on 39 samples and identified a deletion of the entire OTX2 in one sample (2.6%). OTX2 mutations were associated with isolated MAC only. Our study confirms that SOX2 mutations are the most common genetic cause of syndromic or non-syndromic MAC. The lower rate of whole SOX2 gene deletions (3% in this study versus 10% in the literature) is likely due to prior testing of many patients by genome-wide array CGH, which can identify genomic deletions of the gene locus. Based on our experience and due to the extreme phenotypic and genotypic variability of the developmental eye disorders, mutation analysis of the SOX2 gene should be considered first for molecular diagnosis of these disorders.

1766W

Digitizing clinical genetic literature: a comprehensive, accurate and scalable approach. *M. Sommargren, Locus Development, Inc.* Locus Development, San Francisco, CA.

A major challenge in bringing personalized medicine to the clinic is capturing published genetic information in a scalable and highly accurate manner. We have digitized over 19,000 clinically relevant genetic variants for over 300 conditions in a framework that enables automated interpretation of a patient's genetic data. Our approach combines manual literature curation by genetic experts with commercially available and custom Locus Development software. The result is a clinical-grade annotated database of genetic variants that captures information about mutation pathogenicity and condition risk models. This is a significant improvement over databases that merely list previously observed variants without confirming mapping, effect, and pathogenicity. We follow ACMG reporting recommendations for genetic variants, but we include a new category for variants observed in patients but with uncertain pathogenicity. We have developed pathogenicity criteria to standardize the treatment of reported and novel variants. We employ concise language to report risk conferred by variants that follow autosomal, X-linked, or condition-specific risk models. Our curation-ahead-of-time approach has several advantages over the traditional sequence-and-interpret method of delivering genetic information to patients: (i) it is scalable, (ii) it enables more comprehensive, concise, and accurate test results, and (iii) it improves turnaround time through automated report generation. We will present practical lessons from the curation of these 300 conditions, such as the types and prevalence of reporting errors in the literature, the challenge of defining conditions in an automated way, and the consistent handling of variants of unknown pathogenicity.

1767W

Estimated variant rate for a next-generation sequencing panel of 13 genes associated with hereditary colon cancer. *AJ. Stuenkel, SL. Tandy, JD. Siegfried.* Ambry Genetics, Aliso Viejo, CA., USA.

Next generation sequencing (NGS) is an emerging technology with numerous advantages over traditional sequencing techniques. The movement of NGS panels into the clinical setting provides clinicians and patients with a rapid method of capturing sequence data from multiple genes at a fraction of the cost compared to Sanger sequencing. Numerous factors influence the adoption of NGS technology in the clinic, including an anticipated increased frequency of genetic variants of unknown significance compared to single-gene tests. Recently, the first targeted NGS panels for cancer predisposition became commercially available. One such panel, ColoNextTM, targets sequence data from 13 genes associated with hereditary gastrointestinal and other malignancies. To provide clinicians and patients with estimated variant rates for the ColoNext panel, variant data was compiled from clinical testing of the 13 genes (APC, BMPR1A, CDH1, CHEK2, MLH1, MSH2, MSH6, MUTYH, PMS2, PTEN, SMAD4, STK11, and TP53). Overall, gene-specific variant rates range from 1.65% (SMAD4) to 12.43% (PMS2) with an average of 5.70%. However, when a three-tier variant classification system (suspected benign, unknown significance, suspected pathogenic) is applied, the rate of variants classified as being of unknown significance ranges from less than 1% to 4% (average of 2.04%) per gene. Based on available data, initial variant rates for the ColoNext panel are estimated as follows: frequency of at least one variant (any classification) is approximately 80%, although the majority of these variants are anticipated to be classified as suspected benign or suspected pathogenic. Continued analysis of variant rates from NGS panels will assist clinicians in test selection and allow for appropriate pre- and post-test counseling for patients pursuing these novel clinical diagnostic tests.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1768W

Simple and accurate assays for targeting *CFTR* mutations of specific geographic/ethnic origins by PCR allelic discrimination: still a place for targeted tools. *C. Vasseur[1], A. LeFloch[1], B. Costes[1], A. deBecdelièvre[1], T. Casals[2], C. Costa[1], M. Goossens[1], E. Girodon[1].* 1) Genetics, GH Henri Mondor, APHP and Inserm U955, Creteil, France; 2) Human Molecular Genetics Group-IDIBELL, Barcelona, Spain.

The search for cystic fibrosis (CF) causing mutations with a frequency above 1% depending on the geographic/ethnic origins of individuals has been proposed as a European recommendation (Dequeker et al., Hum Mutat 2009). This is commonly achieved by using a standard kit and, depending on the ethnicities, additional sequencing to the search for specific mutations. For the latter, identification of variants of unknown significance (UVs) is a major concern. It makes genetic counseling delicate, in particular when the UVs are found in CF carriers' partners and in case of fetal bowel anomalies. Hence, we believe that there is still a place for tests focused on specific mutations, even in view of the advent of next generation sequencing (NGS). First, NGS is far from being readily accessible to diagnostic laboratories, and second, and more importantly, it would even worsen the situation by increasing the detection of UVs. Therefore, we developed simple assays targeted on 6 geographic origins (Italy, Spain, Portugal, Turkey, North Africa and Black Africa). They contain 3 to 8 mutations. The assays are based on probe hydrolysis during PCR and fluorescence detection (custom TaqMan® SNP genotyping). Homozygous and heterozygous mutant controls were generated for each mutation, from cloned PCR products or by directed mutagenesis. Validation was done on blinded series of DNA representative of different genotypes and on prospective series of patients' DNAs to be tested in routine analysis. The results were compared to Sanger sequencing. The assays fulfilled the validation requirements in terms of sensitivity, specificity and robustness. They are simple and rapid to implement in diagnostic practice and simplify genetic counseling in the context of heterogeneous populations.

1769W

The CLARITY Challenge: Children's Leadership Award for the Reliable Interpretation and appropriate Transmission of Your genomic information. *C.A. Brownstein[1,4], E.T. DeChene[3], K.C. Flannery[2], S.W. Kong[2], M.C. Connolly[3], I.S. Kohane[1,2], A.H. Beggs[3,4], D.M. Margulies[1,4].* 1) Division of Genetics and Program in Genomics, Gene Partnership, Boston Children's Hospital, Boston, MA; 2) Harvard Medical School Center for Biomedical Informatics and Children's Hospital Informatics Program, Boston, MA; 3) Division of Genetics and Program in Genomics, The Manton Center for Orphan Disease Research, Boston Children's Hospital, Boston, MA; 4) Harvard Medical School, Boston, MA.

Introduction: Despite rapid progress in the speed and performance of next-generation sequencing, there has not been a dramatic advancement in the application of sequencing's insights to everyday patient care. Before personalized genomic medicine can be adopted at scale, several hurdles must be addressed: 1. Conflicting gene variant annotations and classifications in non-standardized databases 2. Inconsistent or non-specific sequencing results and non-interoperable processes 3. Lack of standards concerning individual privacy and data access 4. Results that are not clear or useful to doctors, genetic counselors and patients. The Boston Children's Hospital CLARITY competition is a step forward in creating standardized methods for analyzing, interpreting, reporting and, using genomic information in a clinical setting. Methods: Three families identified by The Manton Center for Orphan Disease Research at Boston Children's have clinical manifestations and pedigrees that suggest a genetic basis for their disorders. Two families have probands with a neuromuscular condition, and the third family has a proband with an autosomal dominant cardiac condition. The genomes and/or exomes of the family members were sequenced and provide the starting point of the Challenge. Participants are tasked with producing an analysis, interpretation, and report suitable for use in a clinical situation. Competitors will be evaluated on 1) the methods by which they analyze and interpret the genome sequences and 2) the ability of the competitors to synthesize the genomic data and produce clinically meaningful reports with actionable results for participants' physicians. The judges are a diverse group of well-respected individuals from the medical and technology industries. Results: 40 teams from academic and commercial entities submitted applications. 33 teams were selected to participate in the challenge. 29 teams signed the participant agreement and accessed the data. Two teams decided not to sign the participant agreement and withdrew from the competition. Contest entries are due in September, and the winner will be announced at the 2012 ASHG annual meeting. Discussion: The CLARITY contest is an international competition to identify and publicize the best practices in clinical genome interpretation and data delivery. Bringing together the best elements of competing pipelines will allow for the development of clear, consistent methods for applying genomic insights to patient care.

1770W

What is a clinical grade genome? *F. De La Vega[1,2], A. Russell[1], M.A. Morales[2], C.D. Bustamante[2], M.G. Reese[1].* 1) Omicia, Inc., Emeryville, CA; 2) Department of Genetics, Stanford University, Stanford, CA.

The precipitous fall in the cost of sequencing spurred by innovations in high-throughout sequencing (HTS) platforms is bringing the use of human exome and genome sequencing closer to the clinic. An important question is whether current sequencing protocols provide data that meets clinical standards of quality. Evaluation of clinical tests requires establishment of analytical validity, with assessment of the false positive/negative rates for the specific conditions/variants being tested. While false positive variant calls can be evaluated through experimental validation, false negatives are more difficult to assess, in particular in the lack of a genome gold standard. Ajay et al. (Genome Res, 2011) studied the performance of HTS on genome-wide coverage, consensus calling, and quality values of calls, suggesting a significant false negative rate. Here we propose to score the known disease variants, e.g. from the HGMD database, in terms of genotype call accuracy, to asses the clinical grade of a HTS genome. To evaluate our metric we analyze a diverse data set: (1) two whole-genome trios of Puerto Rican and Yoruban origin sequenced by Complete Genomics (CGI); (2) the deeply sequenced sample of Ajay et al.; (3) a sample sequenced with both CGI and Illumina technologies by Lam et al., Nat Biotech, 2011; and (4) an exome obtained through the 23andMe Exome Pilot sequenced to ~80X. To eliminate the variance due to bioinformatics pipelines, and to perform joint variant calling in the trios, we realigned and called all the data with the RTG Investigator software. We note that the hg19 reference genome includes 37 HGMD disease alleles that are not reported in variant files unless there is an heterozygote at such positions. We observe an average 2% to 4% false negative for HGMD disease alleles at minimum coverage depth of 6x and 20x, respectively, which corresponds to the depth needed to achieve ~99% genotyping accuracy. We also study the variance in scores among the genomes analyzed, the differences between whole-genome and exome assays, different call quality value cut-offs, and the impact of average sequencing depth, sequencing platform, and joint pedigree variant calling. Our results provide guidance for establishing the analytical validity of HTS exomes and genomes for clinical applications and suggest improved metrics to assess their clinical grade.

1771W

The mutation type and effect of variants underlying human genetic diseases. *J. Westbrook, Y. Kobayashi, Locus Development, Inc.* Locus Development, Inc., San Francisco, CA.

To develop clinically useful interpretations of genomic data, Locus Development Inc. has analyzed published data on the genetic basis of human diseases. At the time of writing, we have curated 18,298 variants associated with 277 genetic diseases from peer-reviewed scientific publications. We critically assessed the pathogenicity of each variant based on its predicted impact on amino acid sequence, experimental data on its functional impact, linkage data from affected families and the known inheritance pattern of the associated disease. The assembled data provides insight into the types of variants underlying genetic disorders. In terms of mutation type, 69.46% of variants were single nucleotide substitutions. The remaining 30.54% of variants were insertions, deletions, duplications or expansions of varying sizes. Deletions large enough to be considered copy number variants were reported for 36.4% of the included genes. In terms of the variants' effects, we found that 39.31% caused frameshifts or introduced premature stop codons in the encoded protein, while 6.20% disrupted canonical splice sites. Splice site variants were biased towards donor splice sites (55.0% donor versus 45.0% acceptor, chi-squared goodness of fit test, $p < 0.001$). In the case of variants with less catastrophic effects on the protein sequence (missense, in-frame, silent, and non-coding variants), we searched for additional evidence supporting their pathogenicity. We found experimental or linkage data indicating that 19% percent of 368 in-frame and 20.15% percent of 8,368 missense variants were pathogenic. We also describe a theoretical framework for determining the likelihood that a missense variant is pathogenic based on its frequency in the population and the occurrence of the associated disease. In addition to variants affecting protein sequence or canonical splice sites, we found evidence that 32 silent and 45 non-coding variants cause genetic disorders. With ever-increasing amounts of human sequence data, careful analysis and curation will be required to separate neutral from pathogenic variants.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1772W

Application of next-generation sequencing technology in the study of mitochondrial diseases. *Y. Ma[1], Y. Cao[1], Z. Chen[2], H. Pan[1], Y. Qi[1].* 1) Central Lab, Peking Univ First Hosp, Beijing, Beijing, China; 2) Department of Human Genetics, UCLA School of Medicine, ULCA Medical Center, Los Angeles,, USA.

Mitochondrial disorders are the most common group of metabolic disorders, with an estimated prevalence of 1 in 5,000. Mitochondrial disorders may present with any symptom, in any organ, at any age and with any mode of inheritance. Mitochondrial disorders can originate from mutations of mtDNA (contains 37 genes, 2 rRNAs and 22 tRNA) and about 1,500 nuclear genes. Many of the mutations in mitochondrial DNA are heterozygosis, therefore, while mtDNA could easily be sequenced by traditional Sanger methods, this technology is inadequate to detect some mtDNA mutations that occur in a small fraction of the total mtDNA molecules. Therefore, we propose to find the mutation in entire mtDNA genome and nuclear genes by next-generation sequencing. 11 mitochondrial encephalomyopathy patients without common mtDNA mutations were collected, mitochondrial genome and all exons of 13 related nuclear genes (including 12 genes encoding complex I and POLG1 gene) were amplified and then sequencing by Illumina sequencing technology. Sequencing data from 11 mitochondrial encephalomyopathies patients showed 6.5×108bp could be called, and 7.7×106bp could be mapped on the reference (89.4%;). After pasting the quality filter, 892 single nucleotide variants (SNV) were detected totally. C11777A mutation in mtDNA was detected in one patient, then it was verified by PCR-RFLP and Sanger sequencing, and it was also found in her mother. This patient was a 10 years old girl, and she was diagnosed as MELAS (mitochondrial encephalopathy, lactic-acidosis and stroke-like episodes syndrome) in clinical. The patient was examined because of left-hand resting tremor, hemi-paresis, and migraine. Elevated lactate levels were measured in the blood. MRI examination showed hyper intense lesions in frontal and parietal lobe and in cerebellar hemisphere. The ratio of C11777A of the girl is 30% in her blood, and 57.3% in her urine. The ratio of C11777A of her mother is 34% in urine, and this mutation could not be detected in her mother's blood. The next- generation sequencing technology has a broad application prospects on detection of pathogenic mutations both in mitochondrial genome and related nuclear genes. C11777A mutation is recognized as a widely seen mutation according to Leigh syndrome, 4 individuals are reported in the PubMed, and we first found the relationship between C11777A mutation and MELAS.

1773W

A combined next generation sequencing and exon-level aCGH diagnostic gene panel for severe developmental delay and epilepsy phenotypes. *N.J. Lench[1], H. Moody[1], A. McTague[2], E. Meyer[2], S. Drury[1], S. Fielding[1], R. Palmer[1], M.A. Kurian[2], R.H. Scott[1].* 1) Regional Genetics Service, Great Ormond Street Hospital for Children, London, United Kingdom; 2) Neurosciences Unit, UCL Institute of Child Health, London, United Kingdom.

Background: A large number of children present with a variety of overlapping severe developmental delay and/or early-onset seizure phenotypes. The number of causative genes continues to grow. In some cases, the phenotype is recognisable and targeted genetic testing is available for a limited number of genes. However, in many cases there is difficulty in defining the phenotype, gene testing is not available and/or does not include copy number analysis. **Objective:** To significantly improve the diagnostic yield in patients with complex developmental delay/epilepsy phenotypes using a combined approach of high-resolution exon-level oligonucleotide microarrays and next generation sequencing. **Method:** 29 genes known to cause developmental delay/epilepsy phenotypes including MECP2, CDKL5, FOXG1, UBE3A, SIP1/ZEB2, TCF4, NRXN1, CNTNAP2, MEF2C, SLC16A2, EHMT1, ATRX, ARX, SCN1A, STXBP1, PCDH19, PLCB1, ARHGEF2, KCNQ2, were chosen for analysis. In order to detect intra-genic insertions/deletions we designed a custom 12plex Nimblegen 135k oligonucleotide microarray. Comprehensive gene sequencing covering all coding exons and intron/exon splice sites was performed using the Haloplex Target Enrichment system (Agilent) in conjunction with the Illumina MiSeq platform. **Results:** DNA samples from 48 children with severe developmental delay and/or early onset seizures of unknown cause were analysed. All had tested negative by routine genome-wide microarray analysis using the Nimblegen CGX-12 135k genome wide array (effective resolution ~200kb). The custom microarray generated high quality and reproducible data with a resolution of ~500bp within target regions. It identified two individuals with likely pathogenic copy number changes - a 3kb deletion including the 3' portion of MECP2 exon 4 and a 50kb gain encompassing exons 6–11 of the CDKL5 gene. Both were confirmed by MLPA and follow up in parental samples indicated that the mutations had arisen *de novo*, confirming likely pathogenicity. We also report the results of our sequencing analysis including the identification of a previously reported pathogenic point mutation in MECP2. **Conclusion:** Our data demonstrate the improved diagnostic yield of this combined approach which offers a cost-effective and efficient method for mutation testing that will improve diagnosis in these challenging and complex phenotypes.

1774W

High-throughput Mutation Screening in Patients with Inherited Retinal Dystrophies. *X. Wang[1], A. Turner[1], M. Brooks[2], H. Rajasimha[2], K. Johnson[3], Y. Fann[3], A. Swaroop[2].* 1) OGVFB, NEI/NIH, Bethesda, MD; 2) NNRL, NEI/NIH, Bethesda, MD; 3) ITBP, NINDS/NIH, Besthesda, MD.

Retinal dystrophy (RD) is a broad group of hereditary disorders with heterogeneous genotypes and phenotypes. More than two hundred genes or loci have been identified as responsible for a variety of retinal dystrophies (RDs). Current available genetic testing for these diseases is complicated, time-consuming, and expensive. It is very important to develop a high-throughput genetic screening tool which can be cost-effective and reliable in mutation identification in RD patients, especially in patients with uninformative family history and/or difficulty in differential diagnosis. We recently developed an assay to detect sequence alterations in genes involved in inherited RDs. In this assay, we developed a primer library which targeted 3071 amplicons from 2078 exons of 184 genes that are involved in retinal function and/or retinal development. We were able to amplify the target regions using the RainDance PCR enrichment method and sequenced the products using next generation sequencing by Illumina's GAIIx. In a pilot study, we have analyzed 16 samples from patients with Retinitis pigmentosa. Preliminary bioinformatic analysis indicated that this procedure was able to cover 99% of target sequence with an average in depth of reads at 245 per base. Data mining identified previously reported mutations in a variety of genes responsible for Retinitis pigmentosa in this cohort of patients. As an example, we were able to identify and validate through Sanger sequencing two *C2orf71* gene mutations in an isolated ARRP patient. The *C2orf71* gene is a newly discovered ARRP gene and would have been very unlikely selected as a target gene if we used a traditional Sanger sequencing algorithmic strategy to study ARRP genes on this patient. These results indicate that this assay is capable to screen for mutations in RP patients, especially in simplex cases with high sensitivity and efficiency, and with a powerful potential in clinical application.

1775W

Predictive Testing for Neurodegenerative Conditions is Often Not Predictive. *C.L. Goldsmith, D.A. Dyment, K.M. Boycott.* Genetics, Children's Hospital of Eastern Ontario, Ottawa, Ontario, Canada.

Predictive testing for adult-onset neurodegenerative conditions is an integral part of a Neurogenetics specialty clinic and comprises approximately 1/4 of our patient referrals. To better understand our predictive patient population we performed a retrospective 5-year review of patients referred for pre-symptomatic testing for Huntington disease (HD), Familial Alzheimer disease (FAD), Frontotemporal Dementia (FTD), and CADASIL. The vast majority of patients were referred for predictive testing for Huntington disease (n =65 of 73). The average age at referral for HD predictive testing was 33 with a range of 17 to 69 years, and there was a marked skewed gender ratio (3F:1M). Surprisingly, of the 65 patients, 12 (18%) were clearly symptomatic at the time of the initial consultation in the Neurogenetics Clinic. In addition, predictive testing was urgently requested in the context of an ongoing pregnancy in 5 (7%) of the HD referrals. Of the 65 individuals who sought predictive testing for HD, only 7 did not proceed with testing. Of those who proceeded with predictive tests for HD, the number who received positive results was close to the expected 50% (26/58). We conclude that the request for predictive testing during pregnancy in our HD cohort likely reflects, in part, the high incidence of HD compared to other neurodegenerative conditions in the population. The most concerning trend that we saw in our HD cohort was the number of individuals who are referred for predictive testing who are already symptomatic. This may be explained by the fact that the majority of these referrals are coming from health care providers who may be less familiar with the symptomology of HD compared with a neurologist or neuropsychiatrist specialized in the field.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1776W

Large scale population analysis challenges the current criteria for the molecular diagnosis of fascioscapulohumeral muscular dystrophy (FSHD). *I. Scionti[1], F. Greco[1], G. Ricci[2], M. Govi[1], P. Arashiro[3], L. Vercelli[4], T. Mongini[4], A. Berardinelli[5], C. Angelini[6], G. Antonini[7], M. Cao[6], A. Di Muzio[8], M. Moggio[9], L. Morandi[10], E. Ricci[11], C. Rodolico[12], L. Ruggiero[13], L. Santoro[13], G. Siciliano[2], G. Tomelleri[14], C.P. Trevisan[15], G. Galluzzi[16], W. Wright[17], M. Zatz[18], R. Tupler[1,19].* 1) University of Modena and Reggio Emilia, Modena, Modena, Italy; 2) Department of Neuroscience, Neurological Clinic, University of Pisa, Italy; 3) Program in Genomics, Division of Genetics, Informatics Program, Children's Hospital, The Howard Hughes Medical Institute, Harvard Medical School, Boston, MA 02115, USA; 4) Department of Neuroscience, Center for Neuromuscular Diseases, University of Turin, Italy; 5) Unit of Child Neurology and Psychiatry, IRCCS "C.Mondino" Foundation, University of Pavia, Italy; 6) Department of Neurosciences, University of Padua, Italy; 7) Department of Neuroscience, salute mentale e organi di senso, S. Andrea Hospital, University of Rome "Sapienza", Rome, Italy; 8) Center for Neuromuscular Disease, University "G. d'Annunzio", Chieti, Italy; 9) Neuromuscular unit, IRCCS Foundation Cà Granda Ospedale Maggiore Policlinico, Dino Ferrari Center, University of Milan, Italy; 10) Unit of Muscular pathology and immunology, Neurological Institute Foundation "Carlo Besta", Milano, Italy; 11) Department of Neurosciences, Università Cattolica Policlinico A. Gemelli, Rome, Italy; 12) Department of Neurosciences, Psychiatry and Anaesthesiology, University of Messina, Italy; 13) Department of Neurological Sciences, University "Federico II," Naples, Italy; 14) Department of Neurological Sciences and Vision, University of Verona, Italy; 15) Department of Neurological and Psychiatric Sciences, University of Padua, Italy; 16) Molecular Genetics Laboratory of UILDM, Lazio Section, IRCCS Santa Lucia Foundation, Rome, Italy; 17) Department of Cell Biology, UT Southwestern Medical Center, Dallas, TX 75390, USA; 18) Human Genome Research Center, Department of Genetics and Evolutionary Biology, Institute of Biosciences, University of São Paulo, São Paulo 05508-090, Brazil; 19) Program in Gene Function and Expression, University of Massachusetts Medical School, Worcester, MA 01605, USA.

Facioscapulohumeral muscular dystrophy (FSHD) is a common hereditary myopathy causally linked to reduced numbers (≤ 8) of 3.3 kilobase D4Z4 tandem repeats at 4q35. However, since individuals carrying D4Z4 reduced alleles and no FSHD and patients with FSHD and no short allele have been observed additional markers have been proposed to support an FSHD molecular diagnosis. In particular a reduction in the number of D4Z4 elements combined with the 4A(159/161/168)PAS haplotype (which provides the possibility of expressing DUX4) is currently used as the genetic signature uniquely associated with FSHD. Here, we analyzed these DNA elements in more than 800 Italian and Brazilian samples of normal individuals unrelated to any FSHD patients. We find that 3% of healthy subjects carry alleles with reduced number (4–8) of D4Z4 repeats on chromosome 4q and one third of these alleles, 1.3%;, occur in combination with the 4A161PAS haplotype. We also systematically characterized the 4q35 haplotype in 253 unrelated FSHD patients. We find that only 127 of them (50.1%) carry alleles with 1–8 D4Z4 repeats associated with 4A161PAS, whereas the remaining FSHD probands carry different haplotypes or alleles with greater number of D4Z4 repeats. The present study shows that the current genetic signature of FSHD is a common polymorphism and only half of FSHD probands carry this molecular signature. Our results suggest that the genetic basis of FSHD, which is remarkably heterogeneous, should be revisited since this has important implications for genetic counseling and prenatal diagnosis of at-risk families.

1777W

Identification of a Common Pan-Ethnic Haplotype Enriched on SMN1 Duplication Alleles Improves the Detection of Silent (2+0) SMA Carriers and Significantly Modifies Residual Risk Estimates. *M. Luo[1], L. Liu[1], I. Peter[1], J. Zhu[1], S.A. Scott[1], G. Zhao[1], C. Eversley[2], R. Kornreich[1], R.J. Desnick[1], L. Edelmann[1].* 1) Dept Gen & Genomic Sci, Mount Sinai Sch Med, One Gustave L. Levy Pl., New York, NY; 2) Northwestern Reproductive Genetics, Inc. 680 North Lake Shore Drive, Suite 1230, Chicago, IL.

Spinal muscular atrophy (SMA) is a common autosomal recessive neuromuscular disorder, and mutations in the *SMN1* gene on 5q13.2 generally involve *SMN1* copy number loss by either deletion or gene conversion events with the highly homologous nearby gene, *SMN2*. Consequently, carrier screening must be performed using dosage sensitive methods that can distinguish between *SMN1* and *SMN2*. The carrier frequency for SMA has been reported to range from 1 in 35 to 1 in 117 depending on ethnicity; however, detection rates vary from 71–94% due to the inability to identify individuals that are silent 2+0 SMA carriers with two copies of *SMN1* on one chromosome 5 and zero copies on the other. We hypothesized that the identification of deletion/duplication founder alleles might provide a general approach to identify silent carriers and improve carrier detection in various ethnic/racial groups. Founder alleles were initially investigated in the Ashkenazi Jewish (AJ) population by microsatellite analysis of a region spanning 3.8 Mb surrounding the *SMN1* locus. A major duplication-specific founder haplotype was identified in AJ individuals with 3 copies of *SMN1*. Haplotype positive duplication alleles were sequenced by both Sanger sequencing of *SMN1* exon 7 through exon 8 and by Next Generation sequencing of the entire *SMN1* gene. Importantly, two tightly-linked polymorphisms in *SMN1*, g.27134T>G in intron 7 and g.27706_27707delAT in exon 8, were detected on the haplotype positive alleles but not in 351 AJ individuals with two *SMN1* copies, making the haplotype highly specific for duplication alleles in this population. The presence of the polymorphisms was further investigated in various ethnicities including African Americans (AA), Hispanics (HI), Asians (AS) and Europeans (EU). In the AS population, the polymorphisms were specific to a subset of duplications, but absent in individuals with 1 or 2 copies of *SMN1*. In the AA, HI and EU populations, the polymorphisms were much more frequent in individuals with *SMN1* duplications than those with 1 or 2 *SMN1* copies. Therefore, identification of these polymorphisms in AJ or AS individuals with 2 copies of *SMN1* is diagnostic of a silent 2+0 SMA carrier, while in AA, HI and EU individuals it significantly increases residual risk. Conversely, the absence of the polymorphisms in all populations significantly decreases residual risk of being an SMA carrier.

1778W

Multiplexing test for molecular diagnosis of skeletal dysplasia (by CGC mutation panel). *P. Tavares[1], A. Lopes[2], L. Lameiras[2], L. Dias[1], J. Sá[2], A. Palmeiro[2], P. Rendeiro[1].* 1) CGC Genetics, Newark, NJ; 2) CGC Genetics, Porto, Portugal.

Introduction: Osteochondrodysplasias, also known as Skeletal Dysplasias (SD) account for more than 450 different genetic diseases with bone involvement but variable clinical characteristics, whose diagnosis is based on clinical examination, radiological findings, histo-pathological and molecular analysis. They represent around 5%; of genetic diseases of the newborn and are a major cause of problems for families and patients. Genetic testing improves clinical diagnosis and is essential for a differential diagnosis. The molecular characterization of genes responsible for SD, is extremely important for establishing a precise diagnostic evaluation, namely during the prenatal period. Here we report our experience using this custom methodology. Method: We developed a multiplex mutation panel (CGC Mutation Panel) that tests 49 point mutations identified in the 6 main genes involved in severe SD: FGFR3, COL2A1, SLC26A2, CRTAP, LEPRE1 and SOX9. With this approach it is possible to achieve a molecular diagnostics of the most frequent and severe forms of SD. Results: We tested 132 cases (48 amniotic fluids, 24 peripheral bloods, 36 DNA samples, 16 cell cultures, 6 CVS, 1 Umbilical cord blood, 1 paraffin block, and 1 tissue sample from a kidney fetal biopsy; one case was tested by CVS and fetal biopsy, confirming the previous result) and 22 had a positive result. 18 cases were positive for heterozygous mutations in FGFR3 gene (8 cases with c.742C>T (p.Arg248-Cys) mutation, 4 cases with the c.1138G>A (p.Gly380Arg) mutation, 3 cases with the c.1118A>G (p.Tyr373Cys) mutation, 2 cases with the c.1948A>G (p.Lys650Glu) mutation and 1 case with the c.2420G>C (p.Term807Ser) mutation), 3 cases with mutations on SLC26A2 gene (2 cases with c.835C>T (p.Arg279Trp) mutation in homozygous and 1 cases with both c.532C>T (p.Arg178Stop) and c.835C>T (p.Arg279Trp) mutations in heterozygous, and 1 case with a homozygous mutation IVS1+2C>A in CRTAP gene. Conclusion: The multiplex mutation panel detects the most common mutations drastically reducing turnaround time (one week after DNA extraction), maintaining accuracy and liability. Results are independent from the type of sample. Faster molecular diagnostic is achieved, allowing early decision-making process in patient management and conduct, being particularly relevant in prenatal situations and future pregnancies.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1779W

A combination of panel and exome sequencing to identify novel human retinitis pigmentosa genes. *F. Wang[1], H. Wang[1], J. Zaneveld[1], K. Zhang[2], R. Koenekoop[3], G. Silvestri[4], N. Solanki[1], J. Cheng[1], I. Lopez[3], H. Ren[3], D. Simpson[4], R. Chen[1].* 1) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA; 2) Institute for Genomic Medicine and Shiley Eye Center, University of California, San Diego, CA, USA; 3) McGill Ocular Genetics Laboratory, McGill University Health Centre, Montreal, Quebec H3H 1P3, Canada; 4) Centre for Vision and Vascular Science, Royal Victoria Hospital, Queen's University Belfast, Belfast, N Ireland, UK.

Retinitis Pigmentosa (RP) is the most common form of retinal degeneration affecting 1 in 4,000 people worldwide (Hamel, C. 2006). RP is a genetically heterogeneous disease known to be caused by mutations in over 50 genes with multiple forms of inheritance. Although RP has been widely studied, causative mutations cannot be found in known RP genes in half of all people affected with the disease. In order to better understand the genetic causes of RP, we examined 106 well-characterized RP patient families, including 33 families with multiple affected members. In order to enrich for patients with novel causative loci, we developed a custom RP capture panel that, along with Next Generation Sequencing, allows rapid screening for patients lacking mutations in known RP genes and other retinal disease genes. A total of 106 RP patients (one affected member per RP family) have been screened with our custom RP capture panel. Variants were called by our in-house variants calling, filtering and annotation pipeline. Sanger sequencing and segregation tests (where available) were performed to validate the results. So far, our analysis indicates 42 patients lacked causative mutations within 163 known retinal disease genes. These patients are likely enriched for novel RP genes and loci. In this enriched cohort, we will use whole exome sequencing in combination with system-wide bioinformatics, biostatistics, and functional analysis in order to identify novel RP genes and loci with the goal of improving molecular diagnosis and treatment of RP.

1780W

Evaluation of *BRCA1* mutations in patients with family history of breast cancer. *S. Rummel[1], C. D. Shriver[2], R. E. Ellsworth[3].* 1) Clinical Breast Care Project, Windber Research Institute, Windber, PA., USA; 2) Clinical Breast Care Project, Walter Reed National Military Medical Center, Bethesda, MD USA; 3) Clinical Breast Care Project, Henry M Jackson Foundation, Windber, PA USA.

Background: Identification of BRCA1 allowed screening of high-risk patients for causative mutations. Current criteria used to identify patients for genetic testing include family history of early-onset breast cancer with or without a family history of ovarian and/or male breast cancer, or a personal history of breast cancer at an early age. Here, we determined the prevalence of BRCA1 mutations in a cohort of female invasive breast cancer patients with a significant family history. Methods: The Clinical Breast Care Project database was queried to identify all female patients with significant family histories defined as being diagnosed <35 years, diagnosed <50 years with at least one primary family member with breast and/or ovarian cancer, or diagnosed <50 years with at least two secondary family members from the same parental branch from two generations diagnosed with breast and/or ovarian cancer. Genomic DNA was isolated from blood and exonic regions of the BRCA1 genes were amplified and sequenced on an ABI3730xl. Sequence data was analyzed using Sequencher 4.10.1. Results: Of the 115 women with significant family histories, three had known deleterious BRCA2 mutations and one had a mutation p53 with a family history of Li-Fraumeni syndrome; these patients were excluded from further analyses. The most common (49%) group of women with significantly family histories are those diagnosed between 35–50 years with a single primary family member with either pre- or post-menopausal breast or ovarian cancer. Luminal A was the most common tumor subtype (52%), followed by triple negative (32%), luminal B (12%) and HER2-enriched (4%). Of the eleven causative mutations detected, all were in patients with triple negative tumors: three diagnosed <35 years, six diagnosed 36–50 years with at least one primary relative with breast cancer and two diagnosed 36–50 with at least two secondary family members with breast cancer. Conclusion: Presence of causative BRCA1 mutations did not segregate with early age of onset or number of relatives with breast cancer. Rather, only patients with triple negative breast disease harbored BRCA1 mutations. These results have important genetic counseling and risk reduction implications as the association between BRCA1 mutations and triple negative breast cancer suggest tumor subtype may be added to the current criteria used to identify patients who should be offered testing for BRCA1 mutations.

1781W

Measuring clinician satisfaction with disclosure of genetic susceptibility test results: Findings from the REVEAL Study. *S.L. Everhart[1, 3], B.M Wood[2], K. Valverde[1], L. Medne[3], J.S. Roberts[4], C.A. Chen[5], K.D. Christensen[4], R.C. Green[6], The REVEAL Study group.* 1) Acradia University, Glenside PA; 2) University of Pennsylvania, Philadelphia PA; 3) Children's Hospital of Philadelphia, Philadelphia PA; 4) University of Michigan School of Public Health; 5) Boston University; 6) Harvard Medical School.

Clinician satisfaction has been shown to correlate with patient satisfaction thus measuring satisfaction can improve healthcare outcomes. Development of a reliable scale to measure clinician satisfaction with disclosing genetic risk information for common diseases will help practitioners and researchers identify best practices for communicating susceptibility test results. This paper describes the psychometric properties of a satisfaction scale used by four physicians and five genetic counselors after disclosing risk estimates to 280 first degree relatives of Alzheimer disease patients in a multi-centered trial. A confirmatory factor analysis (CFA) was performed to validate whether four previously found factors associated with clinician satisfaction in primary care were the same for genetic susceptibility testing. Data was then analyzed with an exploratory factor analysis. A 19-item scale was identified with a 4-factor structure assessing satisfaction with (1) the interpersonal connection between the clinician and patient, (2) the process of obtaining patient information, (3) the intensity of patient demands, and (4) perception of patients' needs and understanding. Each identified factor had an Eigenvalue greater than one and was found to be internally consistent with Cronbach alphas of 0.90, 0.76, 0.83, and 0.75 respectively. Contrary to the CFA of primary care office visits, appropriate use of time fell out of the analysis in the measure of genetic susceptibility testing . This scale can be useful in determining the satisfaction of clinicians in Alzheimer's disease susceptibility testing and related assessment procedures. As discovery of risk factor genes is ever increasing, this study can act as a foundation for future studies of clinician satisfaction when disclosing genetic susceptibility results.

1782W

Reactions to Direct-to-Consumer BRCA Test Results. *U. Francke[1, 2], C. Dijamco[1], A.K. Kiefer[1], N. Eriksson[1], J.Y. Tung[1], J.L. Mountain[1].* 1) 23andMe Inc, Mountain View, CA; 2) Dept Genetics, Stanford Univ Med Ctr, Stanford, CA.

Direct-to-consumer (DTC) genetic health reports provided online include low-effect disease risk information based on published genome-wide association studies, carrier status for recessive disorders, variants affecting drug response and a few high-impact Mendelian dominant mutations. Concerns expressed in the literature and position statements by medical societies postulate that genetic information should not be provided DTC because consumers will misunderstand it, positive test results will cause panic and inappropriate actions, and negative test results will lead to inappropriate actions such as foregoing cancer screening. To assess these claims, we carried out semi-structured interviews, collecting empirical data on the harms and benefits experienced by consumers in response to the most actionable test with proven clinical utility currently available DTC: for 3 mutations relatively common among women with Ashkenazi Jewish ancestry that predispose to hereditary breast and ovarian cancer (HBOC): BRCA1 185delAG and 5382insC, BRCA2 6174delT. 32 mutation-positive and 31 mutation-negative participants, matched for age, sex and ancestry, were asked about their emotional response to the report, with whom they shared the data, and the actions they have taken or plan to take. 59% of cases and 35% of controls were unaware at the time of purchase that the 23andMe service includes testing for mutations predisposing to HBOC. In each group of 16 males (M) and 16 females (F), 6 carried a BRCA1 and 10 a BRCA2 mutation. 4F had been diagnosed with HBOC, and 5F and 2M had prior knowledge of their mutation status. Of 10F and 14M who learned for the first time that they carried a BRCA mutation, 6F and 4M stated they were completely surprised. No participants recalled their mutation status incorrectly. Asked to rate their emotional response, no cases were extremely upset, 3F/1M were moderately upset, 3F/6M were somewhat upset, 8F/8M were neutral and 1F/1M were relieved. Of 31 controls, 7F/8M felt neutral and 11F/5M relieved. Results were shared most often with family and friends, followed by genetic counselors and primary care physicians who initiated repeat testing and specialist follow-up. Non-carrier controls did not report actions such as foregoing cancer screening. We were struck by the snowball effect within families leading to identification of many additional mutation carriers, and the burden expressed by male carriers realizing the risk for sisters and daughters.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1783W

Targeted gene capture of BRCA1 using short, explicitly non-gene-specific, peptide nucleic acid baits. *SV. Tavtigian[1], K. Tao[1], R. Bell[1], J. Rosenthal[1], W. Kohlmann[2].* 1) Oncological Sciences, Huntsman Cancer Institute, University of Utah, Salt Lake City, UT; 2) High Risk Cancer Clinics, Huntsman Cancer Institute, University of Utah, Salt Lake City, UT.

Introduction. Some implementation of massively parallel sequencing will eventually almost certainly replace capillary-based Sanger sequencing for clinical mutation screening. Because targeted capture protocols are, for limited genes sets, considerably less expensive than whole genome sequencing, these are likely to make their large-scale clinical debut first (e.g., Ambry Genetics' Next-gen Cancer Panels). However, testing of some important susceptibility genes by this strategy is precluded by existing composition of matter patent claims. Purpose. As a naïve strategy to circumvent such patent claims, we examined the suitability of biotinylated 14-mer peptide nucleic acid (PNA) baits for targeted capture of selected sequences prior to massively parallel sequencing. Methods. Biotinylated 14-mer PNA baits were prepared for exon 2 and the 3' end of exon 11 of BRCA1. Baits were selected explicitly to perfectly match two or more sequences in the human genome. Following initial studies of hybridization temperature, their capture efficiency was compared to that of an Agilent targeted capture reagent (designed to capture the coding exons of 28 known and candidate breast and colon cancer susceptibility genes) by Q-PCR. The 14-mer PNA baits and Agilent baits were then used in parallel to capture DNA from libraries made from carriers of a known mutation in BRCA1 exon 2 (187delAG) and two mutations located near the 3' end of BRCA1 exon 11 (p.E1250X and p.Q1313X). The enriched libraries were then sequenced on an Illumina HiSeq2000. Results. At a hybridization temperature of 57° and bait concentrations between 10 nM and 100 nM, 14-mer PNA baits achieved capture efficiencies equivalent to those of the Agilent reagent. Moreover, sequencing of libraries captured under these conditions easily found the known mutations. Conclusion. Biotinylated 14-mer peptide nucleic acid baits could be used for targeted capture in preparation for massively parallel sequencing. Because these baits are less than 15 bp long and can be designed explicitly to target genomic sequences in addition to their primary gene target, they may provide a strategy to circumvent composition of matter claims in certain 1990s-era gene patents.

1784W

Predictive testing for idiopathic scoliosis: Are orthopedic surgeons and pediatricians ready and willing? *A.M. Laberge[1,2,3], A. Moreau[3,4,5].* 1) Dept of Pediatrics, Faculty of Medicine, Université de Montréal, Montreal, Canada; 2) Medical Genetics Division, CHU Sainte-Justine, Montreal, Canada; 3) CHU Sainte-Justine Research Center, Montreal, Canada; 4) Dept of Biochemistry, Faculty of Medicine, Université de Montréal, Montreal, Canada; 5) Department of Stomatology, Faculty of Dentistry, Université de Montreal.

STATEMENT OF PURPOSE: This study is part of a larger project on the genetics of pediatric scoliosis whose goal is to identify genetic and biochemical markers of adolescent idiopathic scoliosis (AIS) and develop a test that could be used to categorize patients into different risk groups. Our purpose is to explore the attitudes of orthopedic surgeons and general pediatricians towards the clinical use of a predictive test for AIS. OBJECTIVES: 1) Identify potential barriers for the use of a predictive test for AIS by clinicians. 2) Identify the needs of clinicians prior to test implementation. METHODOLOGY: Focus groups were conducted with orthopedic surgeons from centers across Canada (1) and general pediatricians (3) in three different centers across Quebec. Sessions were audiotaped. Content analysis was performed. RESULTS: In both groups, there was more interest in the use of a predictive test for AIS as a prognostic test than a susceptibility test. The main barrier is the reluctance of clinicians to guide patient management based on a test result if it goes against their clinical risk assessment: the test will need to have very high positive and negative predictive values for them to trust its results. Also, high cost could be a deterrent, unless it is found to have major impact on patient outcomes. When asked about what information the test report should provide, orthopedic surgeons and general pediatricians disagree about the format of risk estimates (quantitative vs qualitative), the level of detail required in test interpretation and the inclusion of recommendations. Both groups emphasized the need for clear evidence-based algorithms that integrate other known risk factors to guide patient management in different clinical scenarios. Clinicians disagreed about who the target population should be and had different perceptions of what is a high risk of scoliosis. Clinicians expressed concerns about submitting children and adolescents to the risks of unnecessary surveillance or inversely to falsely reassure some families based on test results. CONCLUSION: Orthopedic surgeons and general pediatricians are interested in a predictive test for AIS, but have different views on how it should be integrated in their practice. Many barriers and needs need to be addressed, especially the need for clear evidence of clinical utility. The focus group with orthopedic surgeons was conducted in collaboration with the Canadian Paediatric Spinal Deformities Study Group.

1785W

***Myocilin* genetic screening in severe glaucoma cases and cascade testing in the Australian and New Zealand Registry of Advanced Glaucoma.** *E. Souzeau[1], K.P. Burdon[1], B. Usher-Ridge[1], A. Dubowsky[2], J.E. Craig[1].* 1) Flinders University, Adelaide, SA, Australia; 2) Institute of Medical and Veterinary Science, Flinders Medical Centre, Adelaide, SA, Australia.

Purpose: Glaucoma is the leading cause of irreversible blindness worldwide and affects 3% of the Australian population over the age of 50. The early stages are asymptomatic and it is estimated that half of the cases are undiagnosed. Glaucoma is treatable, and severe vision loss can usually be prevented if diagnosis is made at an early stage. *Myocilin* mutations are identified in 2-4% of primary open-angle glaucoma (POAG) cases, the most common type of glaucoma, and are inherited in an autosomal dominant fashion with high penetrance. Genetic testing to identify asymptomatic at risk individuals is a promising strategy to reduce the prevalence of glaucoma blindness. **Methods:** POAG cases with severe visual field loss were recruited through the Australian and New Zealand Registry of Advanced Glaucoma. Clinical information and blood were collected from participants. Cases were tested for *Myocilin* mutations by direct sequencing and results were provided back to the participants. Predictive genetic testing was made available for relatives of participants found to carry a *Myocilin* mutation. **Results:** Over 1100 participants with severe vision loss from POAG were recruited. The prevalence of *Myocilin* mutations was significantly higher in the cohort with severe vision loss (4.2%) than in a comparative cohort of non-severe cases (1.6%, p=0.02). The mutation detection rate was also higher in individuals with young age at diagnosis, higher intraocular pressure, and strong family history of POAG. Cascade testing was requested by 55 relatives, of whom 29 were found to carry the familial mutation. Twenty-three reported being unaffected, and 91% of them had never seen an ophthalmologist before being tested. The strongest motivations for undertaking predictive genetic testing were to take appropriate interventions and to provide information to children. Individuals who chose to have predictive testing perceived strong benefits, either medical or emotional, in being tested. **Conclusions:** *Myocilin* mutations are more prevalent in phenotypically selected POAG cases. Their identification in glaucoma affected individuals provides the subsequent opportunity to screen their at-risk relatives, and to reduce glaucoma blindness through early management and intervention. Understanding the motivations of individuals undergoing predictive genetic testing helps clinicians and genetic counselors to addressing the relevant issues and providing adequate support to their patients.

1786W

Parental Awareness, Attitudes, and Experiences of Genetic Testing in Autism Spectrum Disorders. *A. Rupchock[1], M. Cuccaro[1], M. Alessandri[2], J.M. Lee[1].* 1) University of Miami, Miller School of Medicine, John P. Hussman Institute for Human Genomics, Miami, FL; 2) University of Miami, Department of Psychology, Coral Gables, FL.

Autism spectrum disorders (ASDs) are one of the most prevalent neurodevelopmental disorders. Studies suggest that up to 20% of individuals with autism have an underlying genetic etiology to account for their autism, and many professional organizations are recommending chromosomal microarrays as a first tier test. Despite these recommendations there has been little research into the psychosocial needs, expectations and experiences of parents whose children receive genetic testing (GT). The aim of this study was to explore the attitudes, beliefs and experiences of GT among parents of children with autism. Key informant interviews were conducted with parents of children with an ASD recruited through a comprehensive outreach and support program serving people with autism and related disabilities. Eight semi-structured interviews were conducted and digitally recorded. The interviews surveyed thoughts and emotions about GT, reasons for pursuing or not pursuing GT, and the general implications of GT. The data were manually coded by two independent researchers and a final consensus was reached. Half of the respondents reported that their child has undergone GT. Three key themes emerged across almost all respondents: seeking answers or a “cure” for ASD, family planning options, and utilizing GT as a means to tailor treatment. Another prominent theme was the confusion about GT and its benefits. The utilization of genetic testing appeared to rely on professional guidance (e.g. “it was never offered to me”). Regarding the utilization of microarrays, several themes emerged: potential usefulness in treatment, assisting in early diagnosis, concerns over access to and cost of GT, genetic discrimination (e.g. labeling with autism for life, information used to deny service), and concerns about use as a screening tool in pregnancy. By eliciting parental attitudes and experiences with GT, we were able to identify several key themes which may be used to inform future research on this topic.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1787W

Genetic Testing in Autism: Parental Attitudes and Expectations. *M. Cuccaro[1,2], K. Czape[1,2], J.M. Lee[1,2], A. Rupchock[1,2], E. Bendik[1,2], M. Alessandri[3], S. Hahn[1,2].* 1) John P. Hussman Institute for Human Genomics, University of Miami Miller School of Medicine, 1501 NW 10th Avenue, Miami, FL 33136; 2) John T. Macdonald Foundation Department of Human Genetics, University of Miami Miller School of Medicine, 1501 NW 10th Avenue, Miami, FL 33136; 3) Department of Psychology, University of Miami, Coral Gables, FL 33124.

Clinical practice guidelines now stipulate that genetic testing, including chromosomal microarray testing (CMA), be offered as a first-tier test for children suspected of having an autism spectrum disorder (ASD. The enthusiasm in such applying genomic technologies to ASD is offset by concerns about the potential psychosocial and experiential impact of such testing on parents of children with ASD. To assess possible concerns and attitudes about use of genetic testing in ASD, we surveyed parents of children with ASD about their knowledge of, experiences with, and expectations for genetic testing in ASD. Developed by clinical psychologists and genetic counselors familiar with ASD, a 54-item University of Miami IRB approved survey, was electronically sent to the Center for Autism and Related Disabilities (CARD) client registries. Participants were restricted to parents who were a part of one of three CARD registries in Florida as of January 2012. We received completed surveys from 397 parents of individuals with ASD (80% White; 52% Hispanic). The majority of respondents agree that genetic testing is useful in healthcare (89%) and are interested in finding out if genetic factors are the cause of their child's ASD (84%). Although nearly all parents were familiar with the terms “chromosomes” and “genes”, fewer were familiar with “chromosomal microarray” (CMA). In contrast to the 6–8% empiric recurrence risk typically quoted for ASD, the median risk estimation by parents was 50%. A number of respondents stated that their child had been tested for Fragile X syndrome (35%); far fewer had karyotype (13%) or CMA (14%). Among those who had CMA, 68% stated they still have questions or don't understand the test result. Interestingly, 76% of parents reported that their treating physicians did not refer them to or discuss genetic services. Tests were most often ordered by neurologists (68%). and few had ever been referred to a geneticist (21%). Surprisingly, nearly half of the respondents (49%) were not sure if their child would benefit from seeing a geneticist. The most common barrier to genetic testing cited was cost and/or insurance issues (64%). Overall, these preliminary data show that only a small subset of children with ASD are offered comprehensive genetic testing and referral to genetics services as recommended by practice guidelines. Our data support the need for enhancing parental knowledge of and access to genetic testing and clinical genetics services.

1788W

Underestimated intimate issues after bilateral prophylactic mastectomy with breast reconstruction in healthy BRCA mutation carriers. *A. Tibben[1,7], J. Gopie[1], C. Seynaeve[3], M. Ter Kuile[4], M. Menke-Pluymers[5], R. Timman[6], M. Mureau[2].* 1) Clinical Genetics, Leiden University Medical Centre, Leiden, Netherlands; 2) Department of Plastic and Reconstructive Surgery, Erasmus University Medical Center, Rotterdam, the Netherlands; 3) Department of Medical Oncology, Family Cancer Clinic, Erasmus University Medical Center-Daniel den Hoed Cancer Center, Rotterdam, the Netherlands; 4) Department of Psychosomatic Gynaecology and Sexology, Leiden University Medical Center, Leiden, The Netherlands; 5) Department of Surgical Oncology, Family Cancer Clinic, Erasmus University Medical Center-Daniel den Hoed Cancer Center, Rotterdam, the Netherlands; 6) Department of Medical Psychology and Psychotherapy, Erasmus University Medical Center, Rotterdam, The Netherlands; 7) Department of Clinical Genetics, Erasmus University Medical Center, Rotterdam, The Netherlands.

Purpose: Decision-making regarding bilateral prophylactic mastectomy and breast reconstruction (BPM-IBR) for healthy BRCA1/2 mutation carriers is not easy and the outcome can be potentially burdensome for the intimate relationship. Therefore, in the current analysis, the impact on body image, sexual and partner relationship satisfaction was prospectively investigated in women opting for BPM-IBR. Patients and Methods: Fifty healthy women undergoing BPM-IBR completed questionnaires at baseline (T0), and most were followed-up 6 months (T1) and after finalized breast reconstruction (median 21 months, range 12–35) (T2). If a statistically significantly change in time was detected with paired t-tests, the outcome was predicted using linear regression analysis. Results: Sexual and partner relationship satisfaction did not significantly change in time. Body image was significantly less positive at T1. After the breast reconstruction process, 39% of the women reported that their breasts felt not pleasantly, 31% was not satisfied with their breast appearance and 23% felt embarrassed for their naked body. Both before and after BPM-IBR, about 30% of women indicated they felt uncomfortable when touched by their partner. An improved body image was predicted by less preoperative cancer distress, a better body image preoperatively, and no severe complications of surgery. Conclusion: BPM-IBR was associated with adverse impact on body image and intimacy, although satisfaction with the overall sexual and partner relationship did not significantly change in time. The psychosocial impact of BPM-IBR in unaffected women should not be underestimated. Therefore, psychological support should ideally be integrated both before and after BPM-IBR for respective women.

1789W

Risk of peripartum acute aortic dissection in women with ACTA2 mutations. *E. Regalado, D. Guo, D. Milewicz.* Internal Medicine, University of Texas Health Science Center Houston, Houston, TX.

Pregnancy is an important risk factor for acute aortic dissection in women with a genetic predisposition to thoracic aortic disease, but there is a paucity of data except for women with Marfan syndrome. We performed a retrospective study of 53 women with ACTA2 mutations who had 137 pregnancies past the 24 weeks of gestation to assess the risk of aortic dissection related to ACTA2 mutations. Only 8 (6%) of these pregnancies were complicated by an acute aortic dissection. Eight women (age range, 29–43 years) carrying 6 missense ACTA2 mutations (1 R39C, 1 R39H, 1 M49V, 3 R149C, 1 G275A, and 1 S302A) suffered either a Type A aortic dissection (6 women) or a Type B aortic dissection (2 women) as early as 28 weeks of gestation and up to 2 weeks postpartum. Type A dissection was lethal in three women, while four women underwent successful emergency repair of the aorta and one woman with Type B dissection was medically treated. Aortic imaging at the time of dissection was available in two women and indicated the maximal ascending aortic diameter to be 37 and 39 mm. Five of these women including one with a Type B dissection were noted to have preeclampsia or a history of hypertension but were not on blood pressure medication prior to the dissection. While a preconceptional aortic diameter of 40 mm or less is generally considered to carry a low risk of aortic dissection during pregnancy, we report here at least 2 women with ACTA2 mutations who had dissection with minimal aortic dilatation and 2 women who had Type B dissections that are usually not predicted by aortic diameter. These findings emphasize the need to counsel women with ACTA2 mutations about the risk of aortic dissection during pregnancy as well as the risk to offspring and reproductive options. In women with ACTA2 mutations who proceed with pregnancy, careful management of blood pressure should be done, especially during the third trimester and a month after delivery, along with imaging of the thoracic aorta.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1790W

Translational impact of VEGFA variants on the prediction and follow-up of ocular complications in pseudoxanthoma elasticum. *E. De Vilder[1], MJ. Hosen[1], LM. Martin[2], BP. Leroy[1,3], JM. Ebran[4], PJ. Coucke[1], A. De Paepe[1], OM. Vanakker[1].* 1) Center for Medical Genetics, Ghent University Hospital, Ghent, Belgium; 2) Department of Dermatology, CHU Angers, France; 3) Department of Ophthalmology, Ghent University Hospital, Belgium; 4) Department of Ophthalmology, CHU Angers, France.

Among the phenotypic characteristics of pseudoxanthoma elasticum (PXE), a rare ectopic mineralisation disease affecting skin, eyes and cardiovascular system, the ocular complications (neovascularisation, subretinal haemorrhage and vision loss) result in important morbidity. In the absence of genotype-phenotype correlations with the causal gene ABCC6, it has been suggested that variants in the VEGFA gene, encoding vascular endothelial growth factor, may influence the ocular phenotype. The aim of this study was to evaluate the clinical usefulness of such VEGFA variants in PXE counseling and ophthalmological follow-up. The VEGFA coding region, introns and promotor were analysed in 66 molecularly confirmed PXE patients with mild, respectively severe retinopathy. Three categorizing methods were applied to define severe retinopathy: i) visual acuity (VA) <10/10, ii) VA < 5/10 (legal blindness) and iii) VA <5/10 and/or multiple anti-VEGF injections. Significant association of 4 VEGFA SNPs with severe retinopathy was found only when using the third categorizing method which also takes into account anti-VEGF treatment, even with (near) normal VA. As such, the identification of these SNPs in a patient does not necessarily imply vision loss with significant impact on activities of daily living but rather an increased necessity for anti-VEGF treatment. This subgroup of patients might thus benefit from a more strict follow-up, emphasizing the importance of regular self-testing, and a more pro-active initiation and duration of anti-VEGF treatment. Whether this will eventually change the outcome will have to be evaluated prospectively.

1791W

Familial occurrence and associated symptoms in a population of individuals with non-syndromic craniosynostosis. *J. Greenwood[1], S. Boyadjiev[2], K. Osann[1], P. Flodman[1], V. Kimonis[1].* 1) Division of Genetics, Department of Pediatrics, University of California, Irvine, Irvine, CA; 2) Section of Genetics, Department of Pediatrics, University of California Davis, Sacramento, CA.

Craniosynostosis (CS) is a common cranial malformation occurring in approximately 1 per 2,000 live births. The majority occurs as an isolated defect (non-syndromic), and is hypothesized to be a multifactorial trait with heterogeneous genetic and environmental factors. It is believed that each of the five suture categories (sagittal, coronal, metopic, lambdoid, and complex) denotes a distinct disease, each with varying associated phenotypes and occurrence risks. This large-scale study was the first to analyze familial occurrence and phenotype associations by suture type in a specific subset of all mutation negative (FGFR1–3 and TWIST) non-syndromic CS patients. We analyzed the family histories of 647 patients and the symptoms of 186 patients with identified idiopathic non-syndromic CS. The incidence of CS in first-degree relatives and suture specific phenotypes were evaluated. We found a significantly increased incidence ratio for CS in first-degree relatives (SIR=56) compared to expected based on general population rates. An increased incidence ratio was observed for all suture categories except mutation-negative coronal CS. We observed occurrence rates in relatives of patients to be highest in metopic CS (5.3%), followed by complex CS (3.3%), sagittal CS (3.0%), lambdoid CS (1.4%), and coronal CS (0.4%). There was a significantly higher occurrence in siblings (4.8%) than in parents (1.9%), and in male relatives (4.2%) than in female relatives (1.7%) of affected patients (p=0.002, p=0.009). In comparisons of phenotypes by suture type, patients with complex CS, not surprisingly, had the highest frequency of reported symptoms while sagittal had the lowest frequencies (p=0.016). Ear infections (p=0.026), palate abnormalities (p=0.003), and hearing problems (p=0.002) were significantly more common in complex CS patients compared to all other patients. Chronic fatigue was more common in sagittal CS (p=0.010); coronal CS patients had increased frequency of visual problems (p=0.002); and metopic CS patients noted increased frequency of chronic nasal congestion (p=0.046) and cough (p=0.002). Our data suggest a likely genetic component in non-syndromic CS which appears to be suture-specific. The incidence of CS in first-degree relatives is higher than expected based on general population rates and differs by suture type, gender, and relationship of relative. Additionally, the phenotype of each suture synostosis exhibits both unique and shared features.

1792W

Personalized risk prediction in a population-based Biobank: the Estonian experience. *K. Fischer[1], K. Läll[2], L. Leitsalu-Moynihan[1], P.C. Ng[1], T. Haller[1], T. Esko[1], A. Metspalu[1,3].* 1) Estonian Genome Center, University of Tartu, Tartu, Estonia; 2) Institute of Mathematical Statistics, University of Tartu, Estonia; 3) Institute of Molecular and Cell Biology, University of Tartu, Estonia.

Estonian Genome Centre, University of Tartu (EGCUT), has collected DNA, plasma samples and extensive phenotypes on about 52000 volunteer participants. In coming years, EGCUT will provide feedback to the participants on their predicted genetic and environmental risks, with the pilot project starting in fall 2012. In this presentation we discuss advantages and limitations of such personalized predictions, based on the results of statistical modeling of the EGCUT data. We concentrate on five diseases and medical conditions: hypertension, type 2 diabetes, glaucoma, lactose intolerance and celiac disease. The individual risk of each condition has two components: genetic and environmental. We estimate the two components separately from different data sources. The genetic component estimates are based on recent large-scale meta-analyses of genome-wide association studies, where a risk score can be formed from the effects of the top hits. The environmental component is based on established risk factors that are additionally validated in the EGCUT cohort. The proportions of the two components of the risk are different for the five conditions. We will propose graphical methods for communication of the individual risk and its components to participants. In addition, we validate the risk predictions using follow-up data for a subset of the cohort. Here we mainly focus on the question, how much does the genetic component improve the risk prediction, when the environmental risk factors have already been taken into account. While comparing prediction models that use a different number of genetic markers, we demonstrate that a genetic risk score is always a stronger predictor of the medical condition than any of the individual markers alone. Still, for hypertension and type II diabetes, the effect of modifiable/environmental risk factors is substantially larger than the genetic risk. For glaucoma, age is a strongest predictor, but for given age group, genetic risk score will clearly differentiate high- and low-risk individuals. For lactose intolerance and celiac disease, genetic risk score has strongest predictive value. Here we show that as these conditions are under-diagnosed in the Estonian population, direct validation of the prognostic models is complicated. We discuss some indirect validation methods (such as associations with nutrition data and related diagnoses) and illustrate their applicability in the EGCUT cohort.

1793W

Acquired Thrombophilias; The Reproductive Geneticist's Dilemma. *T. Goldwaser[1,2], K. Bajaj[1,2,3], S. Klugman[1,2].* 1) Montefiore Medical Center, Bronx, NY; 2) Albert Einstein College of Medicine, Bronx, NY; 3) North Bronx Health Network, Bronx, NY.

Our Reproductive Genetics Division within the Department of Obstetrics and Gynecology plays a crucial supportive role in identifying patients at risk for inherited diseases. There are certain circumstances when the role of the Geneticist extends beyond genetic disorders. Acquired thrombophilias may be one such circumstance. When working in a busy, urban setting, a history of a remote thrombotic event may not be uncovered in the routine prenatal visit. Additionally, a history of recurrent pregnancy loss (RPL) may not routinely prompt a referral to a Maternal Fetal Medicine or Reproductive Endocrine specialist. When a history of RPL is elicited during a genetics session, testing for the inherited thrombophilias (IT) has been offered. Practice guidelines published in 2011 by the American College of Obstetrics and Gynecology support testing for acquired antiphospholipid antibodies (APLA) for RPL and do not support testing for IT. Current ACMG guidelines, last updated in 2006, state that testing for Factor V Leiden mutations may be considered for RPL. This seeming conflict is a call to Genetics professionals to review current practices and indeed order appropriate testing in the case of RPL. A geneticist working within an obstetrics department is well positioned to test patients for APLA. This additional testing affords the opportunity to alert the OB team that a patient may be a candidate for treatment in cases of RPL or to prevent thrombosis during/immediately after pregnancy. There is strong evidence that aspirin and heparin use improve the live birth rate in patients with APLA and RPL. Meta-analysis and retrospective cohort studies found similar benefits in cases of IT, forming the basis of the current ACMG guidelines. This testing strategy will prevent missed opportunities and ensure safer outcomes for obstetric patients by alerting prenatal care providers of patients at risk. We propose that our Reproductive Genetics department begin testing for both IT and APLA. We present a stepwise testing and counseling protocol for patients with RPL and thrombosis based on a review of the current literature. This protocol can be a useful tool to any Geneticist taking care of obstetric patients.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1794W

Family history and risk assessment in black South African women with breast cancer. *T. Wainstein[1, 2], C. van Wyk[1, 2], A. Krause[1, 2].* 1) Human Genetics, School of Pathology, University of the Witwatersrand, Johannesburg, Gauteng, South Africa; 2) National Health Laboratory Service, South Africa.

Black South African women who have breast cancer are generally diagnosed at a younger age, have more aggressive disease and a poorer prognosis compared to their white counterparts. There is little research regarding inheritance of breast cancer in black South Africans. This study aimed to investigate family history and inheritance of breast cancer in black South African women and evaluate the use and consistency of risk assessment models. A retrospective, file-based analysis of 45 black women diagnosed with breast cancer before the age of 50, was undertaken. Probands were ascertained from the weekly Genetic Counseling Clinic held at the Breast Clinic, Chris Hani Baragwanath Hospital, Johannesburg, South Africa. The probands' personal and family history was analysed. Results indicated that there were few affected women with significant cancer histories (4/45; 9%). Results also suggested that age at diagnosis may not be an appropriate predictor of inherited breast cancer risk in this population. A significant proportion of this sample may have sporadic rather than inherited breast cancer. Three risk assessment tools, (the Claus model, the Tyrer-Cuzick model and the Manchester Scoring System) were evaluated. These showed some degree of consistency, yet each had unique advantages and disadvantages. These tools were designed from data on populations of predominantly European origin; applicability to non-white populations' has not been validated. The true value for this population can only be established once molecular genetic analysis has been performed. This study highlights the necessity of molecular genetic screening in this population to delineate individuals who are truly at increased risk of developing inherited breast cancer. This would clarify which individuals would benefit from cancer risk assessments and various cancer prevention and reduction strategies. Information obtained from this study will direct future research in the South African Black population concerning genetic counseling and testing for inherited breast cancers.

1795W

Support for patients with Young-Simpson syndrome, their families and other peoples concerned - Study of patients and family group meetings. *Y. Yamanouchi[1], T. Nishikawa[2], K. Enomoto[2], N. Furuya[2], S. Mizuno[3], T. Kondo[4], M. Adachi[5], K. Muroya[5], M. Masuno[1], K. Kurosawa[2].* 1) Genetic Counselig Program, Kawasaki University of Medical w, Kurashiki, Japan; 2) Division of Medical Genetics, Kanagawa Children's Medical Center, Yokohama, Japan; 3) Department of Pediatrics, Central Hospital, Aichi Human Service Center, Kasugai, Aichi, Japan; 4) Division of Developmental Disabilities, The Misakaenosono Mutsumi Developmental, Medical and Welfare Center, Isahaya, Japan; 5) Division of Endocrinology and Metabolism, Kanagawa Children's Medical Center, Yokohama, Japan.

Young-Simpson syndrome (YSS) is a rare disorder with only seven cases reported in Japan. Obtaining information about this disorder is thus difficult, and the patients, their families and other people concerned tend to be isolated, with no opportunity to meet those affected by the same disorder. We conducted a survey to determine the efficacy of patient and family group sessions in providing support to these patients and families.We held two “YSS patient and family group sessions” (February 2011 and November 2011, with five families participating in each session) organized by our research team. Doctors and researchers, as well as two certified genetic counselors, attended these sessions. The participants said that they would like to continue having such sessions in the future. They enjoyed meeting other patients and families affected by the same disorder, as they had been trying to obtain medical information about research and treatments. The families actively socialized during the information-exchange time. On the other hand, they also commented on how difficult it is to receive proper medical care in their districts. The participants expressed their concerns that when they visit their otolaryngology (ENT) clinics for conditions such as the common cold, the doctors are not aware of their disorder. This may be because YSS is a pleiotropic disorder characterized by skeletal, endocrine, and neurological abnormalities, and few doctors are aware of it. Creating an information card with details of the disorder, medications to avoid, and contact information for doctors with specialized knowledge of this disorder, as well as allowing families to present it to health care staff will likely enable them to receive necessary medical care. These family group sessions were useful as a form of peer counseling, during which the patients and their families were able to discuss their concerns and worries with each other. It was also an opportunity for them to obtain the latest information about this rare disorder, which was otherwise not readily available. It is essential to have support from various health care professionals, such as specialized doctors who are experienced in managing this particular disorder or medical genetics specialists.

1796W

Duplications Xq28 are questionable for genetic counselling. *C. Costa[1,2], A. Briand[1,2], C. Rothschild[3], E. Bieth[4], J. Melki[5], C. Metay[1,2], M. Goossens[1,2].* 1) AP-HP, CHU Henri Mondor, UF de Génétique, Creteil, France; 2) Inserm et Université Paris 12, U955, Créteil, France; 3) AP-HP, Hôpital Necker, Centre de Traitement des Hémophiles, Paris, France; 4) Hôpital Purpan, Génétique Médicale, Toulouse, France; 5) UMR 788, Inserm et Université Paris 11, Kremlin Bicêtre, France.

Today patients with intellectual disability (ID) are commonly explored by CGH-Array to search for chromosomal microdeletion or microduplication. Recently an Xq28 duplication involving F8 gene in 3 patients with isolated (ID) has been reported (El Hattab AW, J Med Genet 2011). We here describe other Xq28 duplications involving F8 gene in patients presenting two different phenotypes, (ID) or haemophilia A (HA), and we discuss the mechanism of these rearrangements and the difficulties of interpretation for genetic counseling. Severe haemophilia A is frequently caused by inversions involving intronic sequence of the F8 gene (int22h1) and one of two remote, distal copies (int22h2 and int22h3). The inversed orientation of the 2 distal copies suggests that interchromosomal recombination should produce more deletions and duplications than inversion rearrangement. Reciproqual deletion and duplication of F8 gene have been suggested. Though several F8 gene deletions have been described, the few duplication reported are most of the time associated with moderate haemophilia and concern only a small part of the gene. The patients were investigated by CGH-array and further analyzed by MLPA targeted on F8 gene or conversely according to their phenotype. The analysis revealed a telomeric duplication that covers about 0.5Mb, affects several genes and part of the F8 gene from the int22h1 copy to a locus lying between the distal int22h2 and int22h3. These observations seems to support the hypothesis of interchromosomal recombination involving these repeat copies. We propose a mechanism for rearrangement and we underlie the apparently conservation of one F8 gene copy in haemophiliac patients. However the identification of this same duplication in patients with ID without low factor VIII levels and conversely in severe HA patients without ID is questionable. Indeed this same Xq28 duplication involving F8 gene is associated with i)isolated ID or isolated HA, ii)severe form of haemophilia, contrary to the initial hypothesis that F8 gene duplication has no consequences, which then do not justify prenatal diagnosis. These data are important and must be taking account for genetic counselling. Indeed, this is of particular interest when such duplication is identified in pregnant woman with low factor VIII level suggesting of haemophilia carrier and no family history of the disease or conversely in woman with history of ID.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1797W

Surname Leakage from Personal Genomes. *Y. Erlich[1], M. Gymrek[1,2], D. Golan[3], E. Halprein[4,5,7], A.L. McGuire[6].* 1) Whitehead Institute for Biomedical Research, Cambridge, MA; 2) Harvard-MIT Division of Health Sciences and Technology, MIT, Cambridge, MA 02139; 3) Department of Statistics and Operations Research, Tel Aviv University, Tel Aviv 69978, Israel; 4) School of Computer Science, Tel Aviv University, Tel Aviv 69978, Israel; 5) Department of Molecular Microbiology and Biotechnology, Tel-Aviv University, Tel Aviv 69978, Israel; 6) Center for Medical Ethics and Health Policy, Baylor College of Medicine, Houston, TX 77030; 7) The International Computer Science Institute, Berkeley, CA 94704.

Sharing sequencing results without identifiers has become a common practice in genomics. Here, we report that surnames can be recovered from these sequencing datasets. Our approach relies on bioinformatic profiling of short tandem repeats on the Y-chromosome and querying massive Web 2.0 genealogical databases. We demonstrate the feasibility of the technique by recovering the surname `Venter' from Craig Venter's genome. We also find that short read datasets are amenable for surname recovery. Applying this technique to more than 1,000 US males shows a success rate of 13%-16%. Surname leakage can lead to a complete breach of anonymity - in most cases, the combination of a recovered surname with minimal demographic information narrows the identity of the target to a handful of individuals. We combine surname recovery with a simple Google search to fully expose the identity of a large HapMap pedigree. We will discuss the implications of surname leakage and suggest policy guidelines to mitigate the risk.

1798F

Optional Enrollment into Access Controlled Genomic Databases: Decision making in Hispanics versus non-Hispanics. *S. Scollon[1], S. Gutierrez[1], A.L. McGuire[2], S. Hicks[3], S.E. Plon[1].* 1) Department of Pediatrics, Texas Children's Cancer Center, Baylor College of Medicine, Houston, TX; 2) Center for Medical Ethics and Health Policy, Baylor College of Medicine, Houston, TX; 3) Department of Statistics, Rice University, Houston, TX.

Background: With the growth of genomic-scale data from research studies there is an increasing request for participants to consent to the deposition of data into databases. Prior studies suggest most are generally willing to share data, but few studies have evaluated decisions in the context of family-based studies, ethnic variation, or parental willingness to have children re-contacted at age 18 for continued data sharing. Methods: We examined the response over a 2-1/2 year period at one center of adults - primarily parents of childhood cancer patients or young adult survivors - being enrolled into a cancer susceptibility research study which includes genome-scale sequencing. Informed consent was obtained in an individual session over approximately 45 minutes. Each adult subject was asked permission to deposit their own genetic information into controlled access scientific databases and to have their child contacted at the age of 18 for consent. Self-described ethnicity and race was recorded. We tested the hypothesis that consent to deposit data or be recontacted may differ among Hispanic and non-Hispanic adults using a Fisher's exact test. Results: A total of 52 adult participants were enrolled with 56% self identified as non-Hispanic and 44% as Hispanic. Ninety percent (19/21) of Hispanic and 74% (23/31) of non-Hispanic adults agreed to have their genomic information included in databases. Of 41 childhood cancer patients 88% of parents (95% of Hispanic and 81% of non-Hispanic) consented to have their child contacted at the age of 18. Both tests failed to reject the null hypothesis that there was no difference between Hispanics and non-Hispanics in either measure (p= 0.1739, p=0.3433). Conclusion: This study found no significant difference between Hispanics and non-Hispanics who consented to participate in this genetic study with genome scale sequencing in data sharing decisions or permission for re-consent of children at age 18. Overall, even when deposition of data into databases is optional, the majority of adults (81%) agreed to have their genome-scale data included within a database and to have their child recontacted at age 18 (88%) to discuss data deposition. This work is supported by grant U01HG006485 to SEP.

1799W

Development of institutional genomic research infrastructure in an academic pediatric hospital setting to enable large-scale enrollment, education and discovery. *C.M. Clinton[1], S.K. Savage[1], C.A. Brownstein[1], I.A. Holm[1,2], D.M. Margulies[1,2,3,4], W.A. Wolf[1].* 1) Program in Genomics, Boston Children's Hospital, Boston, MA; 2) Department of Pediatrics, Harvard Medical School, Boston, MA; 3) Division of Developmental Medicine, Boston Children's Hospital, Boston, MA; 4) Center for Biomedical Informatics, Harvard Medical School, Boston, MA.

Our objective is to develop an institutional genomic research infrastructure that couples a pediatric Biobank with a new genomic translational research initiative at Boston Children's Hospital to accomplish large scale data sharing and promote discovery. This initiative called the Research Connection, follows the Informed Cohort proposed participant engagement paradigm that was originally described in 2007 by Kohane et al, focusing on an educational, opt-in consenting model that includes the option of individual genetic results return while maximizing the use and potential for the data. Our proposed design includes a tiered consenting structure that allows the participant to engage to the level of their comfort, from sample donation, to participant engagement and general feedback, and finally to individual genetic results return and the potential of enrolling in clinical services. One main driver of this opt-in, full consent model is the proposed changes to the Common Rule and its potential impact on banked clinical samples. The first tier revolves around the creation of a protocol and consent form for the Biobank that will allow for simplified enrollment around the donation of already obtained clinical discard samples and linkage with electronic medical record data. This effort is coupled with a comprehensive education campaign, including web-based education, geared toward both patients and providers to allow for rapid integration of this new initiative into a complex academic environment. The second tier includes participant engagement that allows enrollees to use a web-based platform to allow for Biobank executed contact with enrollees to support investigator needs. The third tier includes an educational and preference setting process for enrollees to be eligible for return of individual genetic research results. Separate research and clinical seed projects also feed into our research infrastructure by including consent form provisions that allow for a de-identified copy of genomic data to be included in a variant database as well as the option to donate project specific data to the Biobank for use in other Biobank-approved studies. We argue that creating a comprehensive, inclusive consenting architecture with appropriate program governance at the launch of a genomics initiative maximizes the return on invested time and funds to benefit the broad academic community, and most importantly respects the autonomy of the pediatric patient population.

1800F

A proposal: a family driven social network model for clinical data sharing and research in intellectual deficiencies and other neurodevelopmental disorders with specific genetic causes. *JL. Mandel[1,2,3], JS. Beckmann[4,5].* 1) IGBMC, Illkirch CU Strasbourg, France; 2) CHU Strasbourg, France; 3) College de France, Paris; 4) Genetique Medicale CHUV Lausanne, Switzerland; 5) Dpt Medical Genetics, UNIL Lausanne, Switzerland.

CGH array has led to a rapidly growing number of genetic diagnoses in patients with intellectual deficiency (ID) associated or not to autism, and this will soon be amplified by detection of mutation in ID-related genes with use of high -throughput sequencing. This will be useful for genetic counseling (when penetrance is known to be high). But the extraordinary genetic heterogeneity of ID will render extremely difficult the determination, for each specific cause (recurrent CNV or mutated gene), of genotype-phenotype correlations and natural history, the estimation of penetrance and expressivity variation, and the organization of clinical trials, except for the most frequent causes (see recent work on 16p11.2 deletion or duplication). Symptomatic treatments will be proposed, with little chance to evaluate whether their efficacy depends on the specific genetic cause. It will be difficult to motivate busy MDs to establish and maintain the wide-ranging databases required for such studies. We wish to propose as an alternative ID databases organised in a social network model, whereby most clinical information would be entered by the family of the patient. Recent papers from 23andMe or PatientsLikeMe have shown that such data can be the basis of useful research. Contacts between families affected by the same genetic cause could be established in an initially anonymous way as for Relative finder in 23andMe, creating gene or CNV specific micro-networks to which interested professionals could be associated, akin to disease specific patients associations. Anonymised data could be accessible to professionals for specific projects approved by a comity composed of health or research professionals and of families representatives. Concerned families could then decide to take part or not in such projects.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1801W

Turkish Medical Students Attitude Towards the Ethics of Genetic Testing and Genetic Information Disclosure in Fragile X Syndrome, Breast Cancer and Huntington Disease. *E.A. Ergul[1], I. Gun[2].* 1) MGH, Boston, MA; 2) Erciyes University Public Health, Turkey.

Introduction: There has been ongoing debate for whether patients with positive genetic test results or physicians who holds this information for their patients should disclose the information to the family members. The same debate appears when western ethics guidelines for minority testing is introduced to oriental cultures. The purpose of this study is twofold: To investigate Turkish medical students' attitude toward genetic information disclosure among physicians and patients for Fragile X Syndrome, Breast Cancer and Huntington Disease, as well as their attitude towards the use of NSGC guidelines for minority testing. Method: 188 Turkish Medical Students in Erciyes University is anonymously surveyed. Survey response rate was 97%. Results: Students (57 % male) were with the mean age of 24.43 (sd= 1.73). Over 80 % of students approved the information disclosure to parents and off-springs but it was only 47 % for siblings and relatives across the 3 disorders. About 30 % of students were indecisive for siblings and relatives. More than half of the students thought it is not their responsibility to give a notice to a colleague if patient refuses to disclose. For minority testing for cystic fibrosis, 78 % of students approved the testing and 50 % students did not agree with NSGC guidelines. For Huntington, more 57 % approved the minority testing and was against the guidelines. Conclusion: Medical students in Turkey with presumed oriental culture orientation approved the genetic information disclosure for parents and off-springs but they were not decisive for siblings and relatives. More than half of the Turkish medical students approved the minority testing for two specific genetic disorders (Huntington and Cystic Fibrosis) and they refused the NSGC minority testing guidelines. Medical curricula in Turkey should cover both the NSGC guidelines and cultural orientations in Turkey.

1802F

Hope and hype in sight: multi-stakeholder perspectives on ocular gene therapy. *S. Benjaminy[1], T. Bubela[1], M. Bieber[1], I. MacDonald[2].* 1) Public Health Sciences, University of Alberta, Edmonton, Alberta, Canada; 2) Ophthalmology, University of Alberta, Edmonton, Alberta, Canada.

Nascent ocular gene therapy (GT) clinical trials are raising patient hopes for a treatment for choroideremia (CHM)—a degenerative retinopathy leading to blindness. GT trials raise the practical necessity of communicating about the risks and benefits with patients in light of a perceived `cure'while avoiding the sensationalism that has historically undermined this field of translational medicine. We hypothesize that patients in the context of a potential treatment will differentially interpret communications about GT in response to media sensationalism. We examined communications about ocular GT in two contexts: (1) word frequency analysis surveying tone of GT coverage in US, UK and Canadian newspapers (1990–2010) (nUS=5438; nUK=1522; nCA= 948) using search algorithms that captured GT synonyms in Lexis Nexis and Factiva databases; and (2) analysis of clinical communications about a CHM GT clinical trial. We conducted semi-structured interviews with 20 CHM patients and 14 clinicians responsible for their care about expected therapeutic benefits, perceived risks, and timeline for clinical implementation. We analyzed interviews using a grounded theoretical approach. Media representations of ocular GT described a range of outcome goals including finding a CHM cure. On the other hand, clinicians emphasized that the goal of early-phase GT trials is better described in terms of a treatment rather than a cure. Patient expectations were more conservative than experts estimated, centering on a hope that CHM GT will halt the progression of vision loss rather than reverse it. While clinicians believed that patients were predominantly concerned with gaining access to GT rather than ensuring its safety, our analysis revealed salient patient worries about GT safety (ex. risk of accelerating vision loss). Interviews revealed discrepancies between patient and clinician timeline expectations for clinical implementation of CHM GT, illuminating patient urgency to access GT within the therapeutic window of opportunity. Our findings about conservative patient visual benefit expectations and safety concerns suggest that patients are well-informed and process media representations critically to avoid sensationalism about CHM GT. Nevertheless differences remain between patient perspectives and clinician understanding. Bridging this gap will promote responsible communications about CHM GT and facilitate the translation of this biotechnology.

1803W

Social and ethical implications of families' experiences of pediatric genetics: Key findings from a UK qualitative study exploring families' expectations, their responses to diagnosis and their approach to managing the uncertainties of genetic investigations. *J. McLaughlin[1], M. Wright[2], E. K. Clavering[1], E. Haimes[1].* 1) Policy, Ethics and Life Sciences Research Centre, Newcastle University, Newcastle upon Tyne, Tyne and Wear, United Kingdom; 2) Northern Genetic Service, Newcastle Hospitals Foundation Trust.

Statement of purpose. In the UK children with neurodevelopmental problems with no clear explanation are often referred - usually by paediatricians - to National Health Service (NHS) genetic services. That referral will lead to consultations where the normal procedure is to take a full medical and family history, examine the physical features of the child and from this to come to a differential diagnosis. Samples may be taken for chromosome and/or DNA analysis. Little is known about the perspectives of families or their experiences of the multiple aspects of genetic investigation and diagnosis. This is increasingly important to consider in the light of new diagnostic technologies - e.g. Array-CGH analysis - which bring with them increasing ethical challenges of understanding and sharing findings of uncertain significance for the child's future development and life expectancy. How does a parent or child make sense of a diagnosis which is so rare and new it is a sequence of numbers and letters rather than providing `a name'? This presentation will explore the lessons that can be learnt from a study of families referred to a NHS genetic service using multiple different diagnostic technologies. Methods Used. The study is a qualitative 3 year study funded by the Economic and Social Research Council which followed 26 families for up to 2 years during their time being seen by a genetic service via interviews and observations of clinic consultations. Summary of results. The presentation explores how families find the long duration sometimes involved in genetic investigations difficult and unexpected; how diagnostic findings may not mean the same thing to families and geneticists; and how the fact that any genetic diagnosis is uncertain about what it can predict about a child's future is both challenging and enabling. Understanding families' responses, and how they are linked to the expectations they have when they first enter a genetic clinic, is important for provision of the correct support throughout, from the initial proposal to refer to the genetic service, through to support at the post-investigations stage.

1804F

Driving research forward, but in what direction? Scientist perspectives on goals and groups invested in autism genetic research. *L.C. Milner, M.K. Cho.* Center for Biomedical Ethics, Stanford University, Stanford, CA.

In the past decade, autism spectrum disorders (ASDs) have shifted from relative obscurity into one of the most prominent issues in American healthcare. The presence of ASDs in public and political discourses results from multiple factors, including a sharp increase in prevalence rates (current estimates of ASDs populations suggest that ≥1% of children lie on the ASD spectrum) and the emergence of a formidable ASD advocacy movement instrumental in raising public awareness about ASDs and promoting research efforts. The rise of ASD in both public and research arenas suggests it as an important area of focus for future research efforts. ASD is considered to be primarily biological in origin, a position reflected in current research efforts focusing heavily on genetic contributions to ASDs. Both federal and nonprofit sources have contributed sizable funding and research resource (e.g., genebanks) efforts to promote ASD genetic research. Despite these highly collaborative efforts, the ultimate goals of ASD genetic research remain ambiguous, as do the expectations and motivations of groups participating in this type of research. Considering the investments of researchers, funders and participants in this endeavor, two questions arise: what values are represented in ASD genetic research produced, and whose values do they represent? To explore these issues, we interviewed genetic researchers studying ASD. In these interviews, we assessed researcher goals and motivations, as wells their impressions of funding agencies and participant groups directly involved in their research projects. Data drawn from these interviews suggest that researchers hold a wide variety of goals for research outcomes, from broad and social (e.g., better understand ASD to reduce stigma) to explicit and clinical (e.g., develop ASD biomarkers for screening). When discussing other groups involved in research, most researchers emphasized specific funders and families as `driving research forward' through their deep commitment to and involvement in research efforts. However, researcher impressions of the motivations and expectations of these groups varied substantially, accentuating the diverse and indeterminate nature of collective goals for ASD genetic research. Future work is needed to understand how these goals and motivations differ between stakeholder groups, as well as how these differences can be effectively addressed within the broader ASD research community.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1805W

Experiential process of securing free prior informed consent for genetics research from an indigenous population in the Philippines. *C.D. Padilla[1,2,5], A.L. Sur[2], M.T. Guia-Padilla[3], M. Baluyot[1], E.M. Cutiongco - de la Paz[1,2,5], S. Padilla[3,4].* 1) Institute of Human Genetics, National Institutes of Health, University of the Philippines Manila; 2) Department of Pediatrics, College of Medicine, University of the Philippines Manila; 3) Anthrowatch Philippines; 4) Department of Behavioral Sciences, College of Arts and Sciences, University of the Philippines Manila; 5) Philippine Genome Center, University of the Philippines System.

Introduction: Of the 110 ethno-linguistic groups in the Philippines, 35 are Negritos who are known by various names, i.e., Agta and Atta in North Luzon; Aeta, Ayta, Alta and Arta in Zambales and Central Luzon; Dumagats in Sierra Madre; Ati or Ata in Panay-Negros; Ata and Abiyan in Bicol, Batak in Palawan; Mamanwa and Ata Manobo in Mindanao; Iraya Mangyan in Mindoro and Remontado in Sibuyan Island, Romblon. Negritos are most vulnerable to extinction due to intermarriage with other ethnic groups, assimilation to mainstream culture and death due to illnesses. The conduct of genomic research needs approval from the National Commission on Indigenous Populations (NCIP), a government agency mandated to protect the interest of the indigenous peoples (IPs). An ongoing project on genomic diversity entailed a 2-year process in securing consent from NCIP which prompted the preparation of modules. **Objective**: To present the experiential account of developing a culturally sensitive protocol of securing free prior informed consent (FPIC) for a genetics research from the IPs. **Methodology**: 1) Consultation with NCIP; 2) identification of a social development organization and other major players that deal directly with IPs; 3) development of culturally sensitive modules in securing FPIC; 4) pilot testing of modules among lay and IPs; and 5) use of the protocol in orienting major players. **Results and Discussion**: After a series of consultations with NCIP and non-government organizations dealing with IPs (AnthroWatch and Koalisyon ng mga Katutubong Samahan ng Pilipinas), the protocol which was developed included inputs from the IPs. The 2-hour module utilizes picture analysis, lecture and interactive method in the orientation. The module includes: 1) Introduction of participants; 2) Warm- Up Exercise: Binukid “body parts” song which aims to arouse their curiosity on the topic; 3) What are the parts of the body? - participants are asked on their concepts of different parts of the bodies; 4) Guessing game, an interactive activity that demonstrates that important body parts are always not visible to the eyes, i.e., chromosomes, genes and DNA. The signing of FPIC by the Tribal Council and the IP participants (signature/thumbmark) ends the module. **Conclusion**: Genomic research on IPs require a special process for FPIC. Although a tedious process, the presented protocol is a model procedure that can be used by researchers involving genomic research.

1806F

The essential ethical considerations for using human genome sequencing as a diagnostic tool in clinical medicine. *A.D. Zellers.* Institute for Consultative Bioethics, Pittsburgh, PA.

In a few years it may cost patients only a few hundred dollars to have their genome sequenced. Genetic information provides valuable diagnostic information. However, the ethical implications of genome sequencing bring to light some important considerations for the use of genetic information in health care. This presentation will address three central ethical issues that health care practitioners and providers must consider when deciding whether to adopt human genome sequencing as a diagnostic tool or how to use it if it is adopted. The first ethical consideration surrounds the type of information a sequenced genome provides. The patient will be provided with information about their genetic health, which can be very useful in helping the person determine how to conduct their life. However, there are no therapies to addresses many of these genetic problems. Currently, there are no gene-transfer therapies available in Western medicine. Patients will be provided with potentially devastating information and have no medical means or available therapies to address their condition. The second ethical consideration revolves around risk probabilities and interpreting the information revealed. Take for example an individual who has their genome sequenced and is told they have a 10% chance of developing Alzheimer's disease. What does that 10% actually represent? Is it a 1 in 10 chances, or perhaps a 1 in 10 chance if a series of other genetic factors are expressed? The answer is unclear and it is difficult to describe the true value of genetic probabilities to patients. The ethical considerations surrounding risk probabilities must be explored. The final ethical consideration addressed in this presentation pertains to the family members of those individuals who have their genome sequenced. Since families share genetic material, the results of one individual's test will have implications for others in the family, particularly for biological siblings and children. When genetic information is being discussed the familial implications must be considered as well as the privacy of both the patient and family members. Privacy, confidentiality and family dynamics must be considered when using genetic information in diagnostics.The purpose of this presentation is to clarify and begin analyzing these ethical concerns and lay the groundwork for further meaningful discourse on the appropriate use of genome sequencing in clinical medicine.

1807W

The evaluation of low template, degraded and mixed DNA profiles in forensic science. *D. Balding.* Inst Gen, Univ College London, London, United Kingdom.

The genotyping of forensic short tandem repeat loci has been developed over the past decade so that it can successfully be employed with only tens of picograms of DNA, corresponding to the DNA from a few cells. The resulting profiles are subject to many stochastic phenomena, including allelic drop-out and drop-in (sporadic contamination), exaggerated stutter peaks and other artefacts. Problems with the handling and interpretation of such complex and noisy evidence has generated controversy, including scathing criticisms by judges, for example in an important Northern Ireland terrorist bombing case, and by academic commentators, for example the report on the DNA evidence in the Italian case of Knox and Sollecito. The incredible power of low template DNA evidence to link someone with a crime-scene based on only touch or even breath has made it compelling to prosecution authorities in many countries, yet the problem of how to convey its evidential value fairly to courts has still not been adequately resolved. I will describe recent progress in developing adequate probability models of degradation, drop-out and drop-in, in handling appropriately the parameters of these models and in formulating hypotheses for a fair assessment of evidential weight. I will also describe the implementation of these developments in the *likeLTD* software for evaluating low-template DNA profile evidence. This R software is freely available at https://sites.google.com/site/baldingstatisticalgenetics/software and is being used in court cases in the UK and US.

1808F

Legal considerations regarding participant privacy in the genomic era of research on the African continent. *G.E.B. Wright[1], P. Koornhof[2], N. Tiffin[1].* 1) South African National Bioinformatics Institute, University of the Western Cape, Cape Town, South Africa; 2) Dept of Mercantile Law, University of the Western Cape, Cape Town, South Africa.

Rapid advances in next generation sequencing (NGS) are making medical genomic research more readily accessible and affordable, including the sequencing of patient genomes and exomes in order to elucidate genetic factors underlying disease. Over the next five years, the Human Heredity and Health in Africa (H3Africa) Initiative (www.h3africa.org), funded by the Wellcome Trust (UK) and the National Institutes of Health Human Health (USA), will contribute greatly towards sequencing of numerous African patient genomes and exomes by NGS. Funding agencies and journals often require submission of genomic data from research participants to databases that allow open or controlled data access for all investigators. Access to such genotype-phenotype and pedigree data, however, needs careful control in order to prevent identification of individuals or families. This is particularly the case in Africa, where many researchers and their patients are inexperienced in the ethical issues accompanying whole genome and exome research; and where an historical unidirectional flow of samples and data out of Africa has created a sense of exploitation and distrust. In the current study, we analyzed the implications of the anticipated surge of NGS data in Africa and the subsequent data sharing concepts on the protection of privacy of research subjects. We performed a retrospective analysis of the informed consent process for the continent and the rest-of-the-world and examined relevant legislation, both current and proposed. We investigated the following: (i) issues pertaining to informed consent, including a lack of guidelines for performing culturally-sensitive NGS research in Africa and poor availability of suitable informed consent documents; and (ii) issues pertaining to data security and subject privacy whilst practicing data sharing, including difficulties in conveying the implications of such concepts to research participants in resource limited settings. We conclude that, in order to meet the unique requirements of performing NGS-related research in African populations, novel approaches to the informed consent process are required. This will help to avoid infringement of privacy of individual subjects as well as to ensure that informed consent adheres to acceptable data protection levels with regard to use and transfer of such information.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1809W

Insurance Coverage of Preventive Services for Individuals with Genetic Predispositions to Cancer. *A. Prince.* Cancer Legal Resource Center, Los Angeles, CA.

Advances in genomic medicine bring great promise of personalizing patient care and saving patients' lives. Only with access to preventive services and knowledge about the consequences of genetic predispositions can an individual make informed medical decisions which may vastly improve treatment options and health outcomes. For many people, getting complete access to preventive services is difficult, if not impossible. For this reason, the Cancer Legal Resource Center (CLRC) and Facing Our Risk of Cancer Empowered (FORCE) created a survey for individuals with genetic predispositions to cancer in order to measure insurance coverage of preventive services. This presentation or poster will summarize the results from the survey to show current trends in insurance coverage for preventive services. Between June 2011 and October 2011, 232 individuals participated in this survey. The survey reflects a national audience. The vast majority (93.3%) of those completing the survey had a family history of breast, ovarian, or colon cancer. Additionally, 92.8% had taken a genetic test for cancer. Of those surveyed, 86.2% had a BRCA1 or BRCA2 mutation and 6.5% had other hereditary cancer mutations. Of those individuals who requested services from their insurers, 82.7% had genetic testing covered, 87.2% had preventive screenings covered, 83.9% had prophylactic surgery covered, and 83.3% had chemoprevention covered. Through ongoing survey analysis, we are finding that certain subgroups of individuals are more likely to have insurance coverage of preventive services than others. We are comparing insurance coverage of these preventive services with other collected information, such as state, age, income level, cancer diagnosis, type of insurance coverage, and participation in appeals processes. If an insurance company denies an individual coverage for a preventive service, the best recourse for the individual may be to appeal the denial. Therefore, this presentation will also cover information on appeals processes in insurance.

1810F

The implementation and impact of student athlete sickle cell trait screening: a survey of Division I California universities. *L. Mar[1], N. Lovick[1], C. Hartshorne[2], M. Treadwell[2], J. Youngblom[1].* 1) California State University, Stanislaus, Turlock, CA; 2) Children's Hospital & Research Center Oakland, Oakland, CA.

Background: In 2010, the National Collegiate Athletic Association (NCAA) mandated sickle cell trait (SCT) testing of all incoming Division I student athletes. Although typically a benign condition, under extreme conditions, SCT has been associated with an increased risk for exercise-related sudden death. NCAA protocol states that athletes be tested for SCT, show proof of prior testing, or waive out of screening. Not included in the mandate are standardized guidelines for schools to follow. Purpose: The aim of this study was to investigate the mandate's impact and implementation at California Division I universities. Methods: Athletic trainers from each of the 24 California Division I universities were invited to participate in an online survey which included questions on demographics, trainer perspective, and the education, testing, and management of student athletes. Data analysis was performed with Chi-square and Fisher's exact tests. Nine post-survey telephone interviews were also conducted. Results: Forty athletic trainers from thirteen of the 24 universities (54%) completed the survey. The majority (92.9%) felt that SCT screening was beneficial and 81% felt that there were no associated risks with the screening. A minority of trainers reported being informed of the impact of SCT on family planning (35.7%). Similarly, a minority of respondents (26.8%) reported that athletes were educated on the impact of SCT on family planning. Less than half (38.1%) reported students receiving pre-test counseling, while 64.3% reported students receiving post-test counseling, primarily by team physicians and/or trainers. Trainers reported differences in the management of athletes with and without SCT. Statistical differences were found between trainers representing public and private universities. Interviews highlighted the divergence of university practices, with few universities mandating that athletes undergo testing. Conclusion: Results of the study revealed a variety of practices being followed, with inconsistencies in the education and counseling of athletes. Trainers and athletes may benefit from more well-rounded educational and counseling resources. This study illustrates the importance of standardizing practices and the potential role for genetic counselors in promoting the education, autonomy, and informed decision-making of student athletes surrounding the issue of SCT testing.

1811W

Direct to consumer genetic testing: What information for which consent? *J. Atia, E. Rial-Sebbag, V. Anastosova, A. Cambon-Thomsen, F. Taboulet.* Inserm and Université de Toulouse III Paul Sabatier, UMR 1027, Toulouse.

Direct to consumer genetic testing (DTC-GT) is a recent phenomenon that poses new ethical and legal issues. It could be considered as a hybrid enterprise between medical practice and purely commercial ventures which is carried around the world via internet. This new context requires an adaptation of the concept of informed consent within the scope of consumer law. Traditionally, informed consent is the process by which a consumer can be actively involved in the commercial ventures. The true informed consent depends on the information that service provider make available for consumers. This professional's duty to inform is controversial concerning websites that offer DTC-GT. Indeed, the information available on the websites makes one believe that the genetic results give the power to control consumer's health and indirectly their life. Nevertheless, after the analysis of some DCT-GT consent forms, it is clear that the results of the test have only informational purpose and, in most of the cases, no clinical validity or utility. Thus, we question whether or not, the information available on some of DTC-GT websites really inform the consumer about the purpose of the test, the risks related to the results and their data protection. Moreover, the companies develop genetic research studies while information about the future use of the samples and associated data is generally missing. Our non- exhaustive search, completed in May 2012, involved 62 websites offering online genetic testing. First, we selected among them those websites proposing DTC-GT with a medical purpose and had a freely available consent form (3) and we analyse the information they provided. . Secondly, we observed if this information corresponded to norms and good practices in consumer and medical law. In conclusion, we make some proposals referring to the report for the management of direct to consumer genetic testing prepared by the Human Genetics Commission and Scientific Technology Options Assessment - STOA.

1812F

Genetic discrimination and the relevance of the United Nations Convention on the Rights of Persons with Disabilities. *A. de Paor.* Centre Disability Law & Policy, Natl Univ Ireland, Galway, Galway, Ireland.

In an age of genetic wizardry, genetic testing is a double- edged sword. With the success of the Human Genome Project, rapid advances in gene discovery and personalised medicine, genetic information is becoming widely available. However, the deluge of genetic information also raises ethical, social and legal issues that can not only threaten advancing science, but can also result in the discriminatory use of genetic information, and violation of genetic privacy. If this information is too widely available it exposes putative persons with disabilities (which, with the aging process, includes most of us) to overt and covert discrimination. In looking towards regulation of genetic discrimination at European Union level, the United Nations Convention on the Rights of Persons with Disabilities (UN CRPD) is of relevance and gives added impetus to the argument that EU level regulation of this area is both preferable and necessary. The UN CRPD prohibits discrimination in a range of fields including employment. The Convention commits the EU to higher standards of non- discrimination with respect to persons with disabilities. The EU ratified the Convention in December 2010 (the first time the EU has ratified an international human rights treaty). There is now a responsibility on the European Institutions to reflect on the case for some form of legislative or other response to the issue of genetic discrimination. This may be a unique opportunity to take a proactive approach towards regulating genetic information for EU citizens in the 21st century. A concerted, uniform approach by the EU would contribute towards guaranteeing fundamental rights and principles, ensuring the effective operation of the internal market and ensuring against any potential obstacles to the realization of these rights. Such an approach is welcomed, not only from a legal perspective, but also from a political and social perspective, and may succeed in facilitating the further advancement of genetic science, and eliminate any fears of genetic discrimination that may prevent individuals from undergoing genetic testing for the benefit of their health. Any EU legislation in the field of genetic discrimination could be regarded as proactive implementation of the Convention. This paper will discuss the need for appropriate regulation of genetic information in light of advancing science, and the relevance of the UN CRPD in working towards regulation at EU level.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1813W

Representing the Indigenous Community: Researcher and Physician voices spark debate over Newborn Screening of CPT1A in Alaska Native and First Nations populations. *N. Garrison.* Stanford Center for Biomedical Ethics, Stanford University, Stanford, CA.

A gene variant present in some Alaska Native (AN) and First Nations (FN) populations has been identified as a potential contributor to infant death and other health problems; however, a lack of consensus exists on the health effects of the variant. AN and FN populations have a high prevalence of the P479L variant of carnitine palmitoyltransferase 1A (CPT1A). CPT1A is involved in oxidation of long-chain fatty acids, important for energy metabolism. Previously, 20 mutations in CPT1A have been associated with reduced enzyme function and reduced energy metabolism in other populations. Anecdotal evidence suggests that 2 copies of P479L are associated with reduced energy metabolism, seizures, and sudden infant death. However, concerns have been raised about the validity of the P479L variant association with sudden infant death: studies to date have involved small numbers, and it is unclear whether an association, if present, is causal. This genetic trait occurs in >50% of most northern indigenous coastal communities; however, 98% of these individuals show no symptoms. With the 2002 adoption of Tandem Mass Spectrometry in expanded newborn screening (NBS), CPT1A was included on the list of conditions and disorders for screening in Oregon and Alaska, but not Washington or British Columbia (BC). A community conversation with AN and FN populations in 2011 sparked debate over the utility of NBS for the CPT1A trait and its impact on the community. Alaskan public health efforts brought attention to the P479L association with decreased energy metabolism and sudden infant death to AN families and local health aid workers. In response, FN parents in BC questioned why newborns in their communities were not screened. Through interviews with physicians, researchers, NBS experts, and policy makers (some of whom are indigenous), I identified a range of opinions about NBS of CPT1A and the interpretation of biomedical research results for the community. This presentation traces the history of CPT1A research, explores NBS policy development, examines community concerns, and highlights views influenced by different health care systems in the US and Canada. While some respondents argue that lives could be saved through identification with NBS, others argue that harms (i.e. over-medicalization, stigma, and undue stress) outweigh any potential benefits. This study provides a deeper understanding of engaging responsible research practices from the laboratory to the community.

1814F

Next Generation Sequencing of Pluripotent Stem Cell Lines: Ethical, Social and Policy Considerations. *R. Isasi.* Centre of Genomics and Policy, McGill University, Montreal, PQ, Canada.

Stem-cell genomics promises to improve our ability to prevent and treat disease by providing cells for organ transplantation and cell therapies. It will also enable the creation of successful model systems for drug discovery and contribute to the understanding of the processes of human cell differentiation for the treatment of several diseases including cancer. For instance, the use of next-generation sequencing (NGS) technologies in induced pluripotent stem cell (iPSCs) lines is expected to greatly contribute to the development of new testing methods for drug efficacy, toxicity and safety for individuals, thereby expediting progress towards personalized medicine. The challenges posed by stem-cell `omics' efforts are not constrained only to scientific hurdles, the field is also confronted by an array of ethical, social and policy concerns as pertains for example, informed consent, feedback of incidental findings, privacy and the governance of research. While, some of these concerns are not completely new, the old answers might not suffice given the scale of challenges. NGS methods are now routinely applied to iPSCs lines, allowing fine, detailed, genotypic information of the cell lines at high resolution. Given that such iPSCs are (often) derived from living individuals - including paediatric populations - are contemporary mechanisms for protecting the autonomy, privacy and confidentiality of donors sufficient? Are current governance procedures adequate? In an era of stem-cell genomics what is the appropriate role for regulatory bodies? This presentation will provide an overview of some key ethical issues surrounding applying next generation sequencing techniques to stem cell research. Issues surrounding informed consent, privacy and confidentiality, data security,feedback of research results and governance of research will be addressed.

1815W

Research Policy of the "Genome Science" Project in Japan. *J. Minari[1], T. Shirai[2], G. Yoshizawa[1], K. Okada[1], N. Yamamoto[1], K. Kato[1,3,4].* 1) Biomedical Ethics and Public Policy, Osaka University, Osaka, Japan; 2) Research Administration Office, Kyoto University, Kyoto, Japan; 3) Institute for Research in Humanities, Kyoto University, Kyoto, Japan; 4) Institute for Integrated Cell-Material Science (iCeMS), Kyoto University, Kyoto, Japan.

A new large-scale project, "Genome Science," started in Japan in 2010. It is 5-year project funded by the Ministry of Education, Culture, Sports, Science and Technology (MEXT). The aim of the project is to support the genome researchers requiring the advanced data. The project has three support programs: "Large-Scale Genome Information Production Program;" "Medical Genome Science Program;" and "Bioinformatics Analysis Program." These programs provide the integrated services of high-throughput DNA sequencing, gene expression profiling, epigenetic profiling, metagenome analysis and/or bioinformatic analysis. Among them, "Medical Genome Science Program" involves the use of human specimens and sequencing of the whole genome and/or the exome, and thus has to take into account various ethical, legal and social implications (ELSI). Our unit called "Research Unit for the ELSI of Genomics" is a part of the project. It was created to establish the governance of the human genome research. The research governance refers to a mechanism of scientific research conducted within the social context. We address ELSI of human genome research in the project under that vision. As the first step, we have designed a model informed consent document used in "Medical Genome Science Program". This process involved policy making of personal genome research, and naturally requires the consideration of key issues such as data sharing, return of results, withdrawal of consent, biobanking and the use of previously collected samples. To deal with these issues, we examined and analyzed the current situation in Japan, by collecting and comparing the informed consent documents utilized in Japan. Referring to these documents, we made the draft of our informed consent document through repeated discussion from multifaceted perspectives with executive members of the project. Currently, we are trying to share the policy of the informed consent document with other large genome projects in Japan. In the next stage, we are focusing on data sharing and release policy. Through the project, we accelerate the deposit of personal genomic data to national database in Japan such as DNA Data Bank of Japan (DDBJ) and National Bioscience Database Center (NBDC). This presentation will highlight our policy of informed consent and data sharing, experienced in "Genome Science" project.

1816F

Cultural Influences on Genetic Testing in East Asia. *H. Numabe[1, 2].* 1) Department of Mecial Ethics and Medical Genetics, Kyoto University Graduate School of Medicine, Kyoto, Kyoto, Japan; 2) Department of Clinical Genetics, Kyoto University Hospital, Kyoto, Kyoto, Japan.

In East Asian countries such as Japan, Korea, and China, genetic testing has been carried out using similar methods with accuracy comparable to western countries. However, the implementation of testing is influenced by various factors such as population policy, economic situation, and the medical system of each government, as well as the bioethical view of each individual. Especially, culture plays a substantial role in genetic medicine and reproductive medicine. It is thus imperative to take cultural influence into consideration when giving genetic counseling and genetic testing to patients and clients in Japan and in the rest of East Asia. We made the questionnaire survey of Japanese university students on genetic diagnoses, and reported the result at the ASHG Meeting in 2008. The questionnaire was consist of properties of the student, considerations to spiritual matters, acceptability of an artificial abortion, the time of the beginning of human life, and considerations to various genetic testing. In Japan, cceptability of genetic testing correlates closely with student's recognition of the time of the beginning of human life. Students who recognize the life start from the fertilization or within 15 days after fertilization are not acceptable to prenatal and pre-implantation tests. On the other hand, students who recognize the life start after 22 weeks of gestational age or after birth are acceptable to them. Acceptability of an artificial abortion has a similar tendency, but over 50% of students who recognize the life start from the fertilization or within 15 days after fertilization are acceptable to an artificial abortion. We planned the similar questionnaire survey in other East Asian countries. We make an interim report.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1817W

The Need for an ELSI Watch of Epigenetic Information. *E. Rial-Sebbag, A. Soulier, C. Delpierre.* UMR U 1027, Inserm, Université de Toulouse - Université Paul Sabatier Toulouse III, 37 allées Jules Guesde, 31073 Toulouse Cedex - France.

All questions relating to the various uses of Genetics at large have been envisaged with regards to their ethical, legal and social issues (ELSI) in particular for the production and utilization of genetics information. This information is usually seen as sensitive by law and requires a particular ethical and legal framework. Known about since the 60's, epigenetics is an innovative field of studies. It bridges the gap between two complementary influences on individual traits and characteristics: environment and genetics. New scientific findings indeed reveal how medical treatments, but also diet and behavior, can affect the expression of genes, that is to say the development and the health of an individual and also of his offspring. Depending on these characteristics it should have been easier to define and analyze the ELSI challenges of epigenetic information in the same way as genetic information. In this paper we make the hypothesis that this position cannot be relevant to this innovative field. Although these two sets of information can be considered as health and sensitive data with a strong familial impact and shared ELSI implications, we will discuss in this paper the following aspects which are totally different with regards to their ELSI challenges: cumulative nature of the risks (risks accumulate during life), targeted populations (vulnerable), equity in access to health care, medical actions (diagnosis, prevention or treatment), environment (social inequalities in health) and protection of privacy. For each of these points various pieces of existing legislation (Research law, health law) should apply as well as ethical principles (Justice, Social inequalities), (Rothstein and al. Nat. Rev. Genet. 2009 April; 10(04):224) but for some of these challenges no rules have been adopted yet. Hence, less developed than for genetic information, we will push for the implementation of an ELSI watch of epigenetic information at the very beginning of the development of this scientific knowledge. In particular, we will suggest taking into account the existence of causal chains linking the social to the biological in public policy and we will propose some points to consider for the development of adequate policies.

1818F

Pervasive Sequence Patents Cover the Entire Human Genome. *J.A. Rosenfeld[1,2], P. Zumbo[3,4], C.E. Mason[3,4,5].* 1) IST/Division of High Performance and Research Computing at the University of Medicine & Dentistry of New Jersey, 185 South Orange Avenue NJMS/C630, Newark, New Jersey, 07103, USA; 2) American Museum of Natural History, Sackler Institute for Comparative Genomics, Central Park West at 79th Street, New York, NY, 10024, USA; 3) Department of Physiology and Biophysics, Weill Medical College, Cornell University, New York, NY, 10065, USA; 4) HRH Prince Alwaleed Bin Talal Bin Abdulaziz Alsaud Institute for Computational Biomedicine, Weill Medical College, Cornell University, New York, NY 10065, USA; 5) The Information Society Project, Yale Law School, New Haven, CT, 06520, USA.

There is substantial uncertainty about the scope and eligibility for patents on genetic sequences, as evidenced by the litigation concerning patents held by Myriad Genetics for testing for cancer associated genes (BRCA1/BRCA2). Claims from the BRCA patents provide intellectual property on sequences as small as 15 nucleotides (15mers). Here, we use an analysis of the human genome to show that the “15mer-rule” is too broad. We observe that 100% of human genes have 15mers that match at least one other gene, with segments of the average gene matching 364 other genes. Additionally, when we examine the patent landscape, we find that between 21% and 100% of the genes in the human genome have been claimed in patents, depending upon the alignment stringency. This indicates that existing genes sequences, surprisingly, are continually being patented even after the human genome has become “prior art.” These data also suggest that current legal rulings around short, non-unique nucleotide sequences create extensive problems for patent law. Thus, a re-examination of existing patent claims, specifically by the Supreme Court, is essential to clarifying the law and to bringing personalized medicine forward.

1819W

The revised points of ethical guidelines for human genome research in Japan. *Z. Yamagata[1], K. Muto[2].* 1) Health Sciences, University of Yamanashi, Yamanashi, Japan; 2) The Institute of Medical Science, The University of Tokyo, Tokyo, Japan.

Ethical guidelines for human genome and gene analysis research in Japan were formulated in 2001, by three government ministries: the Ministry of Education, Culture, Sports, Science and Technology; the Ministry of Health, Labour and Welfare; and the Ministry of Economy, Trade and Industry. The present review of the Personal Information Protection Law has taken into consideration advances in genome research such as the implementation of studies involving large amounts of genomic information, diversification of the study design such as genome cohort studies, and the recent ability of sequencing genomic information expeditiously, which is attributable to the advent of next-generation sequencers. The main aspects reviewed are the use of existing samples, method of collection and distribution of resources, the drafting of informed consent so that it will also be applicable to future genome research, and the disclosure of genetic information. With respect to the use of existing samples, institutions that do not possess a correspondence table can handle samples anonymized in a linkable fashion in the same way that they handle those anonymized in a non-linkable fashion. In terms of collection and sale of samples, requirements and administrative procedures have been revised to enable the more effective use of existing samples and other resources. Additional information can be attached to the description of these procedures by obliging organizations (biobank institutions) that collect samples from other research institutions and sell them to different research institutions. These organizations supply samples to user institutions in the linkable anonymized form. Regarding the disclosure of genetic information, the revisions adhere to the basic principle of disclosure with respect to the Personal Information Protection Law, but some disclosure-related issues remain unresolved. New regulations have also been established for safety management measures related to the handling of genetic information, compliance rules for the outsourcing of genetic research, and education and training for researchers and members of ethical review boards.

1820F

Does evidence justify the use of multigene sequencing panels for rare genetic disorders? *D. Allingham-Hawkins, A. Lea, L. Cushman-Spock, L. Wieselquist, S. Levine.* Hayes Inc, Lansdale, PA.

Objective: To assess the evidence behind the use of multigene panels for the diagnosis of rare genetic disorders. Context: Next generation sequencing technologies are advancing at a rapid rate, allowing the generation of a large amount of data in a relatively short period of time without expending significantly more precious laboratory resources. Many rare genetic disorders are genetically heterogeneous with more than one, and sometimes many, different genes contributing to the clinical presentation of the disease. From a laboratory perspective, inclusion of more genes in a testing panel does not present a technical challenge but is there evidence that testing more genes impacts patient outcomes? Methods: Proprietary methodology that combines the ACCE (Analytical validity, Clinical validity, Clinical utility, Ethical, legal and social implications) model for genetic test evaluations with internationally accepted health technology assessment methodology was used. Conclusions were based on published peer-reviewed scientific literature only. Results: The published evidence behind 8 multigene panels for rare genetic conditions including cardiomyopathy, neonatal jaundice and nonsyndromic hearing loss, among others was evaluated. The number of genes tested in these panels ranged from 5 to 66. In some cases (e.g. familial dilated cardiomyopathy), panels offered by different labs had little overlap with respect to the genes included, suggesting a lack of agreement between labs regarding the appropriate genes to test. For some cases, evidence suggested clinical benefit for testing one or a few genes in a panel. However, studies that evaluated the impact of the full panel on patient care were rare and, in most cases, nonexistent. Conclusions: There is currently little or no published scientific evidence of the clinical validity or clinical utility of most multigene panels. This suggests that the development of multigene panels is being driven by technology and laboratory considerations rather than clinical need or patient benefit. Studies are needed to evaluate the clinical impact of multigene panels before their use in patient care can be justified.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1821W

The student-athlete's knowledge, behaviors, concerns and perceptions in regards to mandatory sickle cell trait testing. *N.E. Thompson[1], B.W. Harrison[1], F. Ampy[2], R.F. Murray, Jr.[1].* 1) Howard University, Department of Pediatrics and Child Health, Division of Genetics, Washington, DC; 2) Howard University, Department of Biology, Washington, DC.

PURPOSE: Nearly 20 student-athlete deaths have been associated with sickle cell trait (SCT), a condition occurring among 1 in 12 African-Americans. The National Collegiate Athletic Association (NCAA) mandated that all Division I collegiate athletes participate in genetic testing for SCT, where 40.9% of the student-athletes are African-American. Traditionally, genetic counseling is recommended before/after genetic testing, and is imperative for increasing understanding, allowing for misconceptions to be addressed, and should be incorporated into the NCAA's mandate. We hypothesized that student-athlete's knowledge, behaviors, concerns and perceptions regarding SCT would be factors in determining the need for a standardized protocol for collegiate institutions; addressing stigmas, promoting education and offering genetic counseling. **METHODS**: A 57-item questionnaire was developed and administered to 91 student-athletes recruited before/after team practices, at a student-athlete dormitory, and during SCT testing. **RESULTS**: Significantly, 87.91% of the student-athletes expressed poor knowledge of SCT and 91.2% expressed a moderate or high interest in receiving genetic counseling/education before SCT testing. Overall, they expressed moderate concerns of discrimination and the majority of student-athletes chose not to discuss their SCT status. **CONCLUSION**: A standardized SCT testing protocol is needed due to the lack of SCT knowledge, the high interest in genetic counseling, the concerns of discrimination and lack of uniformity in regards to disclosure. This protocol will assist student-athletes in making informed decisions, while giving coaches/athletic staff the knowledge to protect their student-athletes' well-being. The field of genetic counseling has potential for growth and participation in the process of integrating genetic testing into collegiate sports.

1822F

Primary Care Patients' Experience of and Reactions to Direct-to-Consumer Genomic Testing: A longitudinal qualitative study. *K. Wasson[1], T.N. Sanders[1], N.S. Hogan[2], S. Cherny[4], K.J. Helzlsouer[3].* 1) Neiswanger Institute for Bioethics, Loyola University Chicago, Maywood, IL; 2) Niehoff School of Nursing, Loyola University Chicago, Maywood, IL; 3) The Prevention and Research Center, Mercy Medical Center, Baltimore, MD; 4) Loyola University Health System, Maywood, IL.

Introduction: Direct-to-consumer genomic testing (DTCGT) has generated debate due to the ethical, legal, and social implications of receiving genomic information of uncertain validity and clinical utility, and the ability to obtain it without a qualified health professional. Little data exists on primary care patients' experience of DTCGT. This qualitative study examined the views, attitudes and decision-making considerations of primary care patients. Methods: Twenty patients were recruited from a primary care clinic and offered DTCGT. Participants were interviewed at four points: pre-test, upon receiving results, and three and twelve months post-test. Results were explained by a genetic counselor. Topics included their decision to test, expectations, concerns, with whom they discussed testing, reaction to results, and the impact of testing. Interviews were audio taped and transcribed verbatim. The data were collected and analyzed simultaneously using grounded theory and constant comparison methods. Results: Participants had a mean age of 49.5 years (range 29–63) and were 60% female, 50% African American, all were high school graduates and 30% were college graduates. Five key categories emerged. 1) Motivations for testing included seeking knowledge about oneself, disease prevention or intervention measures, and altruism. 2)Concerns were limited to fear of misuse, privacy breaches, and test accuracy. 3)Expectations were few and many participants were unsure what to expect; others expected disease risk information and hoped for no “bad” results. 4)Responses to testing were mostly positive or neutral. 5)Many participants noted no impact of DCTGT over time. A few expressed increased awareness or reassurance about their health status and a few made positive health behavior changes. More participants predicted they would talk to their physicians than did after receiving results. The majority of participants were pleased to have undergone DTCGT and would recommend it to others. Conclusions: Results are mixed and suggest that DTCGT may not have the predicted burdensome impact on consumers or physicians as many participants reported no impact of results and did not talk to their doctor post-results. A few made positive health behavior changes post-results in light of “good” results. False reassurance remains a concern for those who were relieved by their “low risk” results due to the developing nature and accuracy of DTCGT.

1823W

Whole genome sequencing: will it destroy newborn screening? *C. Allen, D. Avard, B.M. Knoppers.* Center of Genomics and Policy, McGill Univ & Génome Québec Innovation Ctr, Montreal, Canada.

Introduction: As Francis Collins recently predicted, whole genome sequencing (WGS) will almost certainly become part of newborn screening (NBS) in the near future. A 2008 report by the President's Council on Bioethics describes the use of WGS in NBS as an appropriate and inevitable end point in the development of personalized medicine, while the NHGRI and the NICHD are preparing to provide $25 million to fund studies examining this practice. However, in March the ACMG published a report stating that WGS “should not be used as a first-tier approach for newborn screening”. To reconcile these two divergent perspectives, an analytical review of the legal and ethical issues associated with the use of WGS in NBS is needed. Methods: We performed a comprehensive, international review of the literature, guidelines and policies on the use of WGS in NBS. Content analysis identified arguments for and against this prospect, based for example on normative principles, psychological harms and benefits, and the appropriate roles of parents, professionals, and the state. Results: Those who promote the use WGS as a part of NBS programs do so because it will provide physicians with a powerful tool for delivering early individualized disease prevention initiatives, and researchers with an extended opportunity to study the prevalence and etiology of many diseases. It can also help families both with caring for their current children, and in reproductive decision-making. However, the use of WGS in NBS will exacerbate existing legal and ethical issues, and pose several novel ones as well. As WGS vastly increases the complexity and sheer amount of data arising from NBS, the difficulties associated with financing, consent and the notification process, the availability of counseling, educating parents, and providing follow-up care will increase significantly. Challenges unique to genetic screening, such as whether parents should be informed of a child's carrier status for late-onset genetic disease, will also arise. While parents often push for the expansion of screening to maximize information available, many argue that this violates children's future autonomy, limiting their choices as adults and depriving them of an `open future'. Conclusion: WGS will be implemented in NBS within the next decade. Before this occurs, it is imperative to develop guidelines based on empirical evidence to ensure that the advantages of this practice outweigh the risks.

1824F

Patient and provider perspectives on personalized genomic medicine: Qualitative sub-study of a multicenter feasibility study of genomic sequencing in advanced cancer patients. *J.P. Bytautas[1], F.A. Miller[1,2], R.Z. Hayeems[1], S. Ernst[3], H. Hirte[4], S. Hotte[4], A. Oza[5], A. Razak[5], S. Welch[3], E. Winquist[3], P.L. Bedard[5], J. Dancey[6], L.L. Siu[5].* 1) Health Policy, Management and Evaluation, University of Toronto, Toronto, Ontario, Canada; 2) Joint Centre for Bioethics, University of Toronto, Toronto, Ontario, Canada; 3) London Health Sciences Centre, London, Ontario, Canada; 4) Juravinski Cancer Centre, Hamilton, Ontario, Canada; 5) Princess Margaret Hospital, Toronto, Ontario, Canada; 6) Ontario Institute for Cancer Research, Toronto, Ontario, Canada.

Background: In oncology, personalized genome medicine (PGM) can identify somatic mutations in tumours predictive of therapeutic response and variants of uncertain significance, as well as germline mutations. Expectations for PGM are high but evidence for benefit is limited and clinical trials are complicated by end of life hopes and vulnerabilities. Patient expectations and the perspectives of physicians offering PGM should be assessed. **Methods:** We explored expectations and perspectives of patients and physicians involved in a multicenter pilot study evaluating the feasibility and value of high throughput sequencing of metastatic tumor biopsies compared to traditional genotyping in patients with advanced solid malignancies. Eligible patients were over 18, conversant in English, accessible for in-person interviews with exclusions for poor performance status. Eligible physicians were oncologist investigators who recruited patients to the pilot study. With informed consent, we conducted open-ended, semi-structured interviews. Transcripts were analyzed for thematic patterns with techniques of constant comparison. **Results:** 25 of 55 eligible patients completed interviews; most exclusions were for poor health (13). 13 of 17 physicians completed interviews. Patients were hopeful regarding new and targeted treatments; for some, an identified mutation led to trial access but others were disappointed by non-findings or limited access. Patients were willing to receive incidental inherited risk information; it was perceived as either burdensome given disease stage or inconsequential given faced-challenges. Physicians endorsed these technologies as cutting edge but were mindful of elevated patient expectations and the shortage of targeted agents in trials. Physicians felt consent and counselling processes could sufficiently mitigate challenges related to incidental findings. **Conclusions:** Our results suggest that patient and physician perspectives were conditioned by the late stage cancer context. Both groups were enthusiastic about PGM but acknowledged challenges. The potential to generate inherited risk information may pose more psychosocial challenges than physicians anticipate. Information and decision tools are needed to support patients in making decisions about participating in PGM research and receiving inherited risk information.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1825W

Consumer Genomics: Motivations and Intentions. *R.C. Green[1, 2], J. Mountain[3], A. Kiefer[3], T. Moreno[4], E. MacBean[4], S.S. Kalia[1], J.S. Roberts[5], the PGen Study group.* 1) Brigham and Women's Hospital, Boston, MA; 2) Harvard Medical School, Boston, MA; 3) 23andMe, Inc., Mountain View, CA; 4) Pathway Genomics Corporation, San Diego, CA; 5) University of Michigan School of Public Health, Ann Arbor, MI.

Consumer genomics is increasingly being used to provide individuals with personal health-related information. Few data exist, however, regarding the characteristics of such consumers; their motivations for seeking genomic information; and factors involved in their decision-making. We report here on baseline data from the Impact of Personal Genomics (PGen) Study, a longitudinal survey of consumers of two leading personal genomics companies, Pathway Genomics and 23andMe. 1288 participants completed an online survey before receiving their personal genome results. Survey items assessed sociodemographic characteristics, motivations for and attitudes toward testing, and plans for use of personal genome information. The average age of respondents was 47.7 years (range: 19–94). A majority was female (60.5%), Caucasian (83.4%), had at least a college degree (77.0%), and had biological children (52.7%). 54.9% reported being in very good or excellent health. Nearly half were very motivated to pursue testing by a desire to improve their health (45.6%) and/or to create a plan for the future (45.0%). 36.1% indicated limited information about their family health history as a very important factor. 71.7% of participants reported being very interested in their risk of diseases, compared to 74.2% who were very interested in ancestry results. Respondents were most interested in learning their genetic risks for heart disease (67.8%), breast cancer (67.2% among females), and Alzheimer's disease (66.0%). 37.3% reported that they did not consider the possibility of unwanted results before ordering the testing, and 18.4% reportedly did not consider the predictive value of the results. 60.5% of respondents indicated that they strongly trusted the testing company to use their genetic information only for purposes to which they consented. Over half (57.6%) of participants plan to discuss their test results with their primary care provider, while only 6.1% plan to share their results with a genetics specialist. These data suggest that certain sociodemographic groups are more likely to purchase consumer genomics services and that ancestry information may be of equal interest as health-related information. Subsequent PGen surveys will track consumers' psychological and behavioral responses to their personal genome results.

1826F

The Ethical, Legal and Social Implications Program: Impact on Federal Policy. *D. Mathews, B. Drehman.* Berman Institute of Bioethics, Johns Hopkins Univ, Baltimore, MD.

The Human Genome Project (HGP) was launched in 1990, with the goal of mapping the entire human genome. Beyond the scientific issues the project would face, it was recognized that the HGP would also raise many ethical, legal and social issues. In an effort to address this concern, the Ethical, Legal and Social Implications (ELSI) Research Program was created to facilitate research to address the broader societal issues raised by the emerging science. One of the ELSI Program's major goals from the start has been to inform genetics-related policy, be it federal, state, professional or beyond. However, there has never been a systematic, empirical attempt to evaluate the success of the ELSI Program on this goal: this project is a first attempt to do so. Given that different types or levels of policy have different pathways for identifying issues ripe for policymaking, gathering information and formulating policy, this pilot project has focused on the impact of the ELSI Program on federal policymaking. We analyze the products (e.g., manuscripts, reports) of ELSI funding from 1990 through 2011 and look for evidence of these products' influence in publically available federal policy documents. The products of ELSI-funded projects, as listed in multiple public resources, were collected and coded for basic characteristics (e.g., author, title) and content. Publicly available federal materials that may document impact of the ELSI program (e.g., materials from Congress and federal advisory committees) were likewise collected and coded. We then look for evidence of interactions between these two bodies of data. We also use these data to track the focus of ELSI-funded products over time, for example, from fairly broad surveys of the ELSI issues raised by the HGP, which are highly represented in the early years of the Program, to studies on specific aspects of genetic research, which make up a major part of the ELSI-funded products over the past five to seven years. A primary goal of the ELSI Program has been to fund studies that will ultimately inform genetic-related policy. Further, the ELSI Program is used as a model for other areas of emerging science, and is federally funded. It is arguably in the interest of the public, scientists, the bioethics community and the ELSI Program itself, to understand the nature of the body of work the Program has funded and to establish the degree to which that body of work has been designed to and has in fact influenced policy.

1827W

Public perspectives regarding personalized medicine and genomic risk profiling within colorectal cancer screening. *S.G. Nicholls[1], B.J. Wilson[1], S. Craigie[1], H. Etchegary[2], D. Castle[3], J.C. Carroll[4], B.K. Potter[1], L. Lemyre[5], J. Little[1, 6] on behalf of the CIHR Emerging Team in Genomics in Screening.* 1) Epidemiology and Community Medicine, University of Ottawa, Ottawa, Ontario, Canada; 2) Clinical Epidemiology, Memorial University, Newfoundland, Canada; 3) ESRC Innogen Centre, University of Edinburgh, UK; 4) Department of Family Medicine, Mount Sinai Hospital, University of Toronto; 5) School of Psychology & Institute of Population Health, University of Ottawa, Canada; 6) Canada Research Chair in Human Genome Epidemiology.

Introduction: Colorectal cancer (CRC) is a leading cause of cancer death, yet is highly treatable if identified early. Screening programs now operate in a number of jurisdictions, principally using a fecal occult blood test (FOBT) as an initial (first-line) screening measure. Of the total number of CRC cases 15–20%; are thought to be due to multiple low penetrance variants. The inclusion of low penetrance variants in population-based genomic risk profiling ("genomic profiling") may provide more accurate risk information, potentially leading to better risk estimates. This strategy has been proposed as a way to more appropriately triage patients to more personalized types and intensities of screening and follow-up. However, the use of genomic profiling within standard screening depends on public attitudes and consequently it is important to engage publics and respond to concerns when considering these developments. Aims: To identify issues raised by the inclusion of genomic profiling into first-line screening tests for CRC in a population for whom CRC screening is routinely offered. Methods: Qualitative discussion workshops with Canadian residents of Ottawa, Ontario (n=46) and St John's, Newfoundland (n=74) aged 50 years or over. Results: Participants were largely supportive of the move to include genomic risk-profiling within existing CRC screening. Benefits were largely seen to be health related: the potential for early intervention, prevention, and closer monitoring. However, several important issues emerged relating to test validity, cost, and access to results. Discussion: Not only are these generic issues that one can ask about any new intervention, particularly when they are implemented in a publicly funded health system with finite resources, but they are consistent with the ACCE Framework for evaluation of genetic tests, with analytic and clinical validity together with clinical utility uppermost in participants minds. Moreover, access to results - particularly by insurance companies - was clearly articulated as an area of concern. These results indicate that whilst supporting the inclusion of genomic profiling into screening for colorectal cancer, the public see important issues and readily engage with concepts, if not terminology, employed by healthcare professionals when evaluating new health technologies.

1828F

Pursuing pharmacogenomic testing within a nonprofit healthcare system: a comparison of stakeholders' perspectives. *S.M. Fullerton[1], S.B. Trinidad[1], J.D. Ralston[2], J.T. Tufano[3], G.P. Jarvik[4], E.B. Larson[2].* 1) Bioethics & Humanities, University of Washington, Seattle, WA; 2) Group Health Research Institute, Group Health Cooperative, Seattle, WA; 3) Biomedical Informatics & Medical Education, University of Washington, Seattle, WA; 4) Medical Genetics, University of Washington, Seattle, WA.

The second phase of the Electronic Medical Records and Genomics (eMERGE) Network focuses on the incorporation of genomic research findings into clinical research and ongoing clinical care. The integration of genomic information into routine clinical delivery will require careful attention to health system management, the development and adoption of defined clinical decision support tools, and patient and physician education. We therefore conducted a comprehensive needs assessment within the setting of the Group Health Cooperative (GHC), a consumer-governed nonprofit health care system that coordinates care and coverage in the Pacific Northwest. Our qualitative research focused on three major classes of stakeholders: (1) select clinical leadership with responsibilities relevant to health system adoption (n=13); (2) patients, including those who do and do not regularly take prescription medications (n=50); and (3) physicians, including primary care doctors and specialists (n=20). Interviews and focus groups were designed to explore general attitudes toward the use of genomic information in clinical care, as well as the more specific possibility of using genomic information to inform drug prescribing (pharmacogenomics). While there was much interest, across all stakeholder categories, in the use of genomic information to identify better targeted treatments and minimize side-effects, numerous ethical and logistical considerations were also raised. Foremost among these were concerns that widespread genetic and/or genomic testing might have adverse effects on patient-physician interactions. The development of clinical decision support tools that facilitate shared decision-making were identified as a high priority for the next phase of the project's research.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1829W

Public preferences for the use of email and Internet communication in large-scale genetic and biobank research. *J. Murphy Bollinger, R. Dvoskin, K. Kreger, D. Kaufman.* Genetics and Public Policy Center, Berman Institute of Bioethics, Johns Hopkins University, Washington, DC.

The success of genetic biobank research depends on establishing and maintaining effective communication with participants that ensures transparency about study goals and procedures. The ability to efficiently contact participants can facilitate recruitment; consent; education; scheduling; data collection; delivery of study findings; and collecting feedback. The use of email and the Internet could effectively address many of these challenges. Few data are available, however, on public acceptance of Web-based tools for these purposes, and the recruitment biases that might result. We surveyed 3,061 U.S. adults about the use of online communication for seven research tasks. **Methods**: We surveyed a representative sample of U.S. residents (in English or Spanish) about the use of a password-protected website and email in a proposed national cohort study. We compared responses among various demographic groups, including social networkers (SNs)—defined as Facebook users who had posted to their account in the past week. **Results**: Among survey respondents, 56% were Facebook members and 30% were social networkers (SNs). A significantly higher percentage of younger participants, women, and people who took the survey in English were SNs. Overall, email was viewed as most acceptable for administrative tasks, including scheduling appointments (42% supported) and sending survey reminders (60%). Majorities supported the use of a password-protected website for receiving general (56%) and individual findings (52%), completing surveys (51%), updating health status (57%), and learning which researchers were using the data (51%). More respondents opposed the use of the Internet to relay individual results (29%) and update participants' health status (22%) than opposed uses for other tasks. More education, higher household income, and social networking were associated with higher support for the use of online communication for all tasks. Respondents who agreed that “I trust the study to protect my privacy” were more likely to accept use of the Web and email for all tasks. **Discussion**: While 7 in 10 participants supported email or Web use for the tasks tested, studies that exclusively use these tools may alienate older and lower-income participants. The higher acceptance among social networkers, coupled with the growth of social networking, suggests that acceptance of online research communications may increase, especially for tasks that do not involve personal health information.

1830F

Public attitudes towards consumer genetics in Japan, South Korea and Taiwan. *K. Muto[1,2], M. Sato[2], H. Hong[1,2], C. Chang[1].* 1) Dept Pub Policy, IMS, Univ Tokyo, Tokyo, Japan; 2) Graduate School of Frontier Sciences, Univ Tokyo, Tokyo, Japan.

Background: DTC (direct-to-consumer) genetic testing for health purposes have been featured in USA and Europe and stimulating discussions for regulation have been evolved. In contrast, several Shanghai-based companies stated to advertise in Japan to sell DTC genetic tests for non-health purposes, such as tests advertised to identify genes related to an individual's talent, character traits and aptitude. Purpose: We have explored sociocultural beliefs and eugenic thoughts behind these so-called `recreational testings` in East Asia. The same surveys in South Korea and Taiwan have been in progress. Methods: To address broad ethical, legal and social implications arising from such tests, we have conducted a web-based self-administered questionnaires among Japan, South Korea and Taiwan. Results: We have obtained 14,718 responses from Japanese citizens. Response rate was 37%. Results: Advertisement is commonly observed in East Asia though some of the tests have been banned in South Korea. Our survey shows that 3.1% of Japanese respondents has purchased DTC genetic tests, and 86.7% of them hasn't recognized such tests. They show higher interests towards tests for health purposes (23.8–47.2%) than non-health purposes (9.8–18.7%). Younger generation shows higher interests toward DTC genetic tests for their children and partners. Discussions: The trend of our survey shows the same trend from our postal survey conducted in 2007 and 2011. Many parents are deeply concerned about providing education for their children and it's difficult to intervention parental decision to purchase such tests.

1831W

Japanese Networkers ` attitudes toward Personal Genome Services. *T. Kido[1], M. Kawashima[2].* 1) Rikengenesis, Tokyo, Japan; 2) Tokyo University, Tokyo, Japan.

Purpose: This study explores potential early adapters of personal genome services in Japan, focusing on expectations related to the several kinds of future personal genome services. Methods: An online survey of 4,343 Japanese Internet users was conducted to assess 1) use and interest in personal genome services; 2) attitudes toward several kinds of personal genome services; 3) expectations for future personal genome services. Comparative analyses with past research on U.S social networkers and Japanese general populations were conducted to characterize Japanese networkers' attitudes on personal genome services. Results: In our research samples, 89.0% of respondents used Internet everyday, 87.5% registered the social network services (SNS), 16.6% used SNS every day. 1.5% used Personal Genome Testing (PGT) (six percent in U.S), 76.8% did not know PGT at all. 58.2% of responders would consider using PGT on life style related diseases or cancers, 50.8 % consider it on intelligence, and 48.7 % consider it on personality. Conclusions: Although the percentage of usage of PGT in Japan is lower than that of U.S, around 50 % of Japanese Networkers have interests to use it. The networkers seem to have more positive attitudes than general Japanese populations, especially on non-medical personal genome services such as applications on intelligence and personality. More detailed results and discussion will be reported in this poster presentation.

1832F

Genetic risk predictions - Attitude of the Estonian public. *L. Leitsalu-Moynihan[1,2], A. Allik[1], M.L. Tammesoo[1], A. Metspalu[1,2].* 1) Estonian Genome Center, University of Tartu, Tartu, Estonia; 2) Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia.

The Estonian Genome Center of the University of Tartu (EGCUT) is a research institute with a population-based biobank that includes biological samples and comprehensive health information on over 51,000 adults. The three main objectives of the center are to promote the development of genetic research, to collect information on the health of the Estonian population and genetic information concerning the Estonian population, and to use the results of genetic research to improve public health. In the coming years the EGCUT will start providing feedback to the participants on their personalized risks predictions, with the pilot project starting in the fall of 2012. Besides the necessary developments in genomics research and policy making, it is crucial to educate the physicians as well as the public in order to facilitate effective application of genomic information in clinical care. In order to investigate the public's standpoints the EGCUT conducted a survey via TNS Emor polling agency. The sample included 1000 Estonian residents between the ages of 15–74 and was composed to be proportional to the Estonian population structure with respect to age, gender, region, and nationality. In order to get an idea what the potential impact of such personalized risk predictions would be and how it might differ depending on the gender and age of the individual we used Chi-squared test and two-sample test to analyze the results of the survey. Overall, women and younger age groups were significantly more likely to change their health behavior according to the physicians' recommendations, and even more likely so if genetic information would be considered. The results also indicate that there is a disconnect between the perception of healthy behavior and the actual health behavior of the individuals. The results should be considered when educating the physicians on communicating risk predictions.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1833W

Attitude change for four years toward genetic testing of children for common disease risk in Japan. *I. Ishiyama[1,2], K. Muto[3], A. Tamakoshi[4], T. Maeda[5], A. Nagai[1], Z. Yamagata[1].* 1) Department of Health Sciences, School of Medicine, University of Yamanashi, Yamanashi, Japan; 2) Teikyo-Gakuen Junior College, Yamanashi, Japan; 3) The Institute of Medical Science, University of Tokyo, Tokyo, Japan; 4) Department of Public Health, Aichi Medical University School of Medicine, Aichi, Japan; 5) The Institute of Statistical Mathematics, Tokyo, Japan.

[Aim] The swift pace of the development of genetic testing for common disease risk raises expectations and concerns, in the case of minors in particular. The aim of this study was to assess the four-year attitude change of Japanese general public toward genetic testing of children for common diseases which will be able to be prevented or controlled, and to explore factors related to the change, by comparing two nationwide opinion surveys conducted in 2005 and 2009. [Methods] In each survey, 4,000 people (age, 20–69) were selected from the Japanese general population by a stratified two-phase sampling method. They were queried about the following topics in a mail survey: attitudes toward genetic testing for disease susceptibilities of children for common diseases, interest in medical genomic studies, belief in science, level of genomic literacy, images of the genomic studies related to medicine (using SD method), and awareness of the benefits and risks of medical genomic studies. The four-year attitude change and the factors related to the change were examined using logistic regression models. [Results] The response rate was 54.3% (2,171/4,000) in 2005; and 52.5% (2,009/4,000) in 2009. The genetic testing for disease susceptibilities of children for common diseases was favored by 55.5% people in 2005; and 58.8% people in 2009. A difference in the attitudes was observed. The interest in medical genomic studies was increased; belief of "scientific development has more advantage than disadvantage" was increased; the average score of genomic literacy became lower for four years. The positive images toward the medical genomic studies, such as "Good", "Certain", "Like", and "Not scary" were increased. Awareness of the risks such as medical genomic studies may "breed discrimination based on disease or disorder" and "have unexpected harmful effect" were decreased. Interest in medical genomic studies, belief in science, positive images and awareness of risks were significantly related to favorable attitudes in each survey. It was suggested that interest, belief, images (changed positively) and awareness of risks (changed negatively) were the factors related to the outcome of positive attitude change toward genetic testing for children. With the level of public acceptance of genetic testing increases, it will become crucial for peoples to make sense of children's right to autonomy and possible risks associated with genetic testing.

1834F

African-Americans in South Florida's Perceptions about Genetic Research Participation: Opportunities and Barriers. *J.M. Lee[1], M. Gavier[1], S.E. Hahn[1], L. Adams[1], D.A. Caldwell[1], K.G. Murphy[1], C. Jean[1], M.A. Pericak-Vance[1,2].* 1) University of Miami, Miller School of Medicine, John P. Hussman Institute for Human Genomics, Miami, FL; 2) University of Miami, Miller School of Medicine, Dr. John T. Macdonald Foundation Department of Human Genetics, Miami, FL.

Diverse representation in genomics research is imperative to ensure that the translation of findings into genomic medicine leads to health equity. Unfortunately, African Americans (AA) remain underrepresented in genetic research. Although strides have been made to increase participation of diverse groups in genomic research, the factors that motivate and inhibit participation still merit continued exploration. Previous qualitative research has identified that community partnership and the utilization of culture specific recruitment strategies is paramount to increasing participation in genomic research (e.g., Johnson, 2011; Corbie-Smith, 1999). South Florida's AA community is diverse and includes individuals of Caribbean ancestry with potentially unique cultural and community needs. The goal of this research was to identify attitudes and beliefs that influence participation in genomic research among AA in South Florida using a focus group methodology. Focus group participants were recruited from South Florida (Miami-Dade/Broward Counties). The following areas were explored through a structured question guide: 1) Access to health information and knowledge about genetic research, 2) Perceptions and attitudes regarding genetic research, 3) Cultural motivators and barriers to participation in genetic research, 4) Effective communication and outreach strategies. Six focus groups were completed by racially-matched facilitators. One group was dropped due to ascertainment bias, leaving five groups for analysis. A total of 45 participants (16 male, 29 female) participated in the groups. Preliminary analysis revealed several barriers to genetics research participation including: limited knowledge about genetics, fears of being incidentally diagnosed with a disease, mistrust of the intent of research, and economic barriers to participation (e.g., taking off work to participate is prohibitive). A number of strategies were identified including: matching the appropriate advertising to the individual, holding community meetings about health related topics, and using doctors and family elders as messengers regarding research studies. Content from the groups will be analyzed in detail to identify overlapping and unique themes related to this community in comparison with extant research on this topic. However, preliminary analysis supports the need to vary and tailor research recruitment strategies not only to the culture and community of interest, but to the individual.

1835W

Unexpected results in biobank research: Experiences of information delivery and direction to health care after detection of mutations leading to long QT syndrome. *H. Kaariainen[1], E. Kujala[2], P. Alha[1], V. Salomaa[1], S. Koskinen[1], H. Swan[3], A. Haukkala[2].* 1) National Institute for Health and Welfare, Helsinki, Finland; 2) Dpt of Social Research, University of Helsinki, Helsinki, Finland; 3) Dpt of Cardiology, Helsinki University Hospital, Helsinki, Finland.

During 2000–2001 a large Finnish population based cohort Terveys 2000 (Health 2000) including 6334 subjects (aged ≥ 30) was collected for public health studies, including relationships between genes, environment and disease. As some of the subjects were found to have mutations for heritable long QT syndrome, it was decided to inform them about these results. The experiences of the process of information delivery and directing the subjects to health care services were examined by sending a questionnaire to all 27 subjects who were found to have one of four mutations for long QT syndrome and interviewing 5 of them. Of the 17 respondents, 16 had experienced the process of returning the results as mainly positive and useful for them, while one had experienced this unexpected result negatively due to concerns related to informing children. All participants felt that genetic results should be reported back to subjects. The participants had received information on long QT syndrome and its inheritance by a nurse of the cardiology unit. Even though genetic counseling had not been offered, 15 of them reported that they had been sufficiently informed about the disease and its inheritance. The 17 respondents had 35 children altogether. They reported that they had informed all the children, except one, still a minor, about the genetic condition. Detailed results from the questionnaire study and the interview will be presented. Biobank studies raise several ethical concerns including the question of delivering genetic results to study subjects. This was one of the first occasions where participants in a large-scale population-based study were informed about a genetic finding and their experiences were studied. Despite of the long interval between giving the sample and receiving the result, subjects thought positively about the procedure. Subjects' main wish for future studies was that there should be a contact person for later questions and concerns relating to the genetic finding.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1836F

The Gene Partnership: A model for pediatric genetic repositories in the age of the Incidentalome. *I.A. Holm[1,2], C.M. Clinton[1], S.K. Savage[1], P.L. Taylor[2,4], K.D. Mandl[2,3,4], J.P. Bickel[2,3,4], D.J. Nigrin[2,3,4], I.S. Kohane[2,3,4], D.M. Margulies[1,2,3,4], L.M. Kunkel[1,2].* 1) Program in Genomics, Boston Children's Hospital, Boston, MA; 2) Department of Pediatrics, Harvard Medical School, Boston, MA; 3) Center for Biomedical Informatics, Harvard Medical School, Boston, MA; 4) Children's Hospital Informatics Program at the Harvard-MIT Division of Health Sciences and Technology, Cambridge, MA.

In 2007 Kohane et al. defined a new "collaborative clinical research regime" called the Informed Cohort. The Informed Cohort is defined mainly by "the dynamic quality of the subjects' changing involvement over time," which translates into the ability for researchers to contact participants on an ongoing basis, both as necessary and as based on participant-expressed preferences. This paradigm allows participants to be mutual partners in research rather than passive donors of material. We have established the Informed Cohort at Boston Children's Hospital by creating a project called the Gene Partnership. The Gene Partnership consists of a large de-identified repository of genetic and health information and an established communication tool called a personally-controlled health record, which encourages communication and collaboration between researchers and participants, including the dissemination of individualized research results. Participants are offered the opportunity to opt-in to results return. 94% of our participants have requested to receive research results, 5% declining results return, and 1% have not confirmed their choice. The Gene Partnership allows parents to grant adolescent access to their child who is 13 years of age or older so they may also engage in the result return process. Even with the wide adoption of the feedback process, parents of adolescent participants were most likely to deny access to their adolescent children, with 70% of them declining access. Many factors influence this decision from developmental ability to family dynamics. We note that establishing this type of structure in a pediatric setting yields a number of challenges, mainly from an ethical perspective of informed consent and return of research results. Despite these complexities, we argue that this is the future of repository design, as it engages the participant community and enables research staff to manage the intricacy of incidental and research results that will certainly arise from genome-wide association studies.

1837W

The Medical College of Wisconsin Program in Genomics and Ethics: An empirical bioethics venture. *K.A. Strong, T. May, K.L. Zusevics, A.R. Derse.* Center for Bioethics and Medical Humanities, Medical College of Wisconsin, Milwaukee, WI.

In 2010 the Medical College of Wisconsin utilized genome sequencing to diagnose a six-year-old boy's mysterious digestive disease, enabling a life-saving bone marrow transplant and representing the first documented clinical application of whole genome sequencing (WGS). The clinical application of WGS will continue to grow as the technology becomes more efficient, the cost of testing decreases and the potential for WGS to guide treatment becomes more established. In preparation for this increase in clinical use of WGS, the Program of Genomics and Ethics has been established at the Center for Bioethics and Medical Humanities at the Medical College of Wisconsin (MCW). An aim of this program is to bridge the gap between the clinical application of WGS and its ethical implications. Ethical discourse in relation to genetic testing in general tends toward psychological harm/benefit arguments, impacts on family relationships, social discrimination potentiality, and influences on autonomy (particularly in the case of testing minors). Concerns about WGS encompass all of the above, but given the inherent uncertainties regarding potential results, also relate to difficulties around traditional concepts of informed consent, the meaning and disclosure of incidental findings and the impact those results may have on patients and their families. Whilst traditional normative philosophical analysis is important in identifying and addressing the ethical questions, this Program will also utilize empirical methods (eg. interviews, surveys, focus groups) to explore the attitudes, experiences, values and processes involved in decision-making that may differ from normative expectations and provide additional guidance regarding clinical practice. The employment of `empirical bioethics' in this manner can be particularly helpful when exploring ethically contentious technologies where concerns regarding those technologies involve empirically testable claims.

1838F

Evaluation of a scalable method for returning results and genetic findings from genomic research to research participants. *J.Y. Tung, S. Wu, B.T. Naughton, J.L. Mountain.* 23andMe, Inc, Mountain View, CA.

Discussions around the return of research results and incidental findings to research participants have come to the forefront recently, in particular as genome-wide sequencing becomes more common. Studies have shown that most research participants who have donated DNA samples to research would like to receive individual results from studies performed using their samples. While multiple groups have published recommendations on how best to return results, very few research groups have actually returned such data. 23andMe, a personal genomics company, not only provides a service that interprets genetic information for non-scientists but also conducts its own research and shares those results with participating customers. These activities put 23andMe in a unique position to assess methods of returning results and findings to research participants. We describe our approach for returning different types of information and compare it to recommendations made by independent working groups. For example, as recommended by many, our genotyping is performed in a CLIA-certified lab, and we have already developed and implemented a system for evaluating whether a finding should be shared with our participants. In addition, we have an established method for recontacting our participants when new findings become available. We also evaluate differences in information-seeking behavior between individuals who came to our service based on their interest in genetic testing and individuals who joined as part of a research project, and show that these differences vary according to the nature of the information presented and the way the individuals were introduced to the service. Though 23andMe's research participants may not be representative of all research participants, these data still suggest that individuals whose primary motivation for joining 23andMe is to participate in research are interested in learning about the "incidental findings" uncovered in the sample they contributed.

The problem of how to return results and incidental findings to research participants in a responsible, scalable, cost-effective, and updatable way is a daunting one. One possible solution is for researchers to partner with entities like 23andMe that specialize in communicating this type of information. Whatever the channel of communication of results, many lessons can be learned from the work already in progress at personal genomics companies.

1839W

Perceptions of genetic counseling services in direct-to-consumer personal genomic testing. *B.F. Darst[1], N.J. Schork[1, 2], E.J. Topol[1, 2, 3], C.S. Bloss[1].* 1) Scripps Translational Science Institute and Scripps Health, La Jolla, CA; 2) Department of Molecular and Experimental Medicine, The Scripps Research Institute, La Jolla, CA; 3) Scripps Clinic Mecial Group, La Jolla, CA.

Aim of Investigation: To describe consumers' perceptions of genetic counseling services in the context of direct-to-consumer (DTC) personal genomic testing. Methods: We performed a secondary analysis of data from the Scripps Genomic Health Initiative, a longitudinal cohort study originally designed to assess the psychological and behavioral impacts of DTC genomic testing for common disease in a large sample of adults. Analyses presented are based on web-based assessments performed at baseline and long-term follow-up, which was completed an average of 14 months (SD= 1.3) after receiving genomic test results. At long-term follow-up, participants were asked to respond to several items gauging their interactions, if any, with a Navigenics genetic counselor, and their perceptions of those interactions. Participants purchased the Navigenics Health Compass at a discounted rate and genetic counseling was provided at no additional charge. Results: Out of 1325 individuals who completed long-term follow-up, 187 (14.1%) indicated that they had spoken with a genetic counselor. Primary reasons given for not utilizing the counseling service included the perception of already understanding one's results (55.6%), being too busy (20.3%), and unawareness that the service was available (10.5%). Of those who did have contact with a genetic counselor, the most common reasons given for utilizing the service included wanting to take advantage of a free service (43.9%), wanting more information on risk calculations (42.2%), and being contacted by a counselor who therefore prompted the interaction to occur (39.0%). Among those who utilized the service, a large fraction reported that counseling helped them to feel more educated about genetics in general (43.9%) and that it improved their understanding of their results (54.5%). A total of 17.6% reported that genetic counseling made them likely to discuss their results with their physician. Conclusions: Only a small proportion of participants utilized genetic counseling after receipt of DTC genomic testing for common disease. Those who did, however, appeared to find it beneficial with many indicating that it improved their understanding of genetics in general, as well as understanding of their own test results more specifically. Our results suggest that genetic counseling may not be a necessary component of genomic testing for most consumers, but that for some, it may have notable beneficial effects.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1840F

Shared decision-making: Seeking a change in the genetics consult. *S. Adam[1], P.H. Birch[1], A. Townsend[2], F. Rousseau[3,4], J.M. Friedman[1,4].* 1) Dept of Medical Genetics, Children & Women's Hospital, University of British Columbia, Vancouver, BC, Canada; 2) Dept of Occupational Science and Occupational Therapy, W. M. Young Centre for Applied Ethics, University of British Columbia, Vancouver, BC, Canada; 3) Medical Biology, Université Laval, Montreal, PQ, Canada; 4) The APOGÉE-Net/CanGèneTest research and knowledge network in genetic health services and policy.

The physician-patient encounter is increasingly regarded as `a meeting between experts' reflecting a change from a former emphasis on paternalism to recognition of patients' unique understanding of their health. It is a shift in philosophy of health care from beneficence to autonomy and from physician-directed care to shared decision-making. More recently, an explosion of health Internet use has accelerated this transition and the `e-patient' (e.g. educated and empowered) has emerged. This trend is an important dynamic for the genetics community to consider especially in complex ethical and medical situations such as clinical whole genome sequencing (WGS). We conducted two focus groups, one with 8 parents of children with intellectual disability (ID) of unknown etiology; the other with 10 members of the lay public (from diverse educational backgrounds and little knowledge of WGS). Qualitative analysis of the transcripts revealed that both groups expressed support for this construction of expert and engaged patients in the context of receiving clinical WGS results. Perhaps surprisingly, both lay and parent groups held very similar views. They gave various explanations for the need to be “expert” patients: 1) Patient choice was emphasized, especially around the ability to choose the types of WGS results to receive; 2) Personal responsibility was conveyed by many as central to the need to keep informed of medical and technological advances; 3) Advocacy was conveyed by both groups as key to receiving full information and optimal care for themselves and their children. For some, this was associated with a lack of trust in the health care system; and 4) An over-burdened health care system was identified by some as a reason to be pro-active. Problems with complex data and understanding results were identified as obstacles to being an expert patient. However, both groups offered novel educational solutions including websites, discussion groups and other formats to suit a range of learning styles. Our findings indicate that the public wishes to engage in shared responsibility for the enormous amount of data that will be generated by clinical WGS by taking control of their own results. In turn the genetics community needs to recognize and facilitate the learning needs of e-patients, providing a variety of effective multi-media tools as options to support the inevitable, and potentially enormously beneficial, evolution of the medical encounter.

1841W

Personal genomics: A study assessing the safety and health outcomes of receiving whole exome sequencing results. *Y. Bombard[*, 1, 2], K. Schrader[*, 2], E. Glogowski[2], R. Rau-Murthy[2], S. Patil[1], A. Eaton[1], E. Elkin[1], M.J. Massie[3], J. Hay[3], S. Shah[2], J. Vijai[2], N. Kauff[2], Z. Stadler[2], M. Robson[2], K. Offit[2], (* Equal contributions).* 1) Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center; 2) Clinical Genetics Service, Department of Medicine, Memorial Sloan-Kettering Cancer Center; 3) Department of Psychiatry and Behavioral Sciences, Memorial Sloan-Kettering Cancer Center.

In the wake of whole genome and exome sequencing (WG/ES), return of individuals' research results involves numerous challenges mainly due to the large volume of incidental data. Research infrastructures need to include counseling and provision of incidentally-identified disease risk (IIDR). Using a prospective, observational cohort design, we are conducting a feasibility study offering the return of IIDR results identified using WGS/WES to 100 participants from existing research protocols involving cancer survivors and their unaffected relatives. Participants are offered communication of categories of incidental results, using a decision aid based on the binning scheme proposed by Berg et al. This offers return of presumed pathogenic variants within genes represented in http://www.genetests.org, a clinical resource providing, expert-authored, peer-reviewed, genetic disease descriptions informing diagnosis, management and genetic counselling. SNPs associated with pharmacogenomic and common disease risk are also included. Genes and SNPs have been categorized according to their levels of clinical actionability and potential distress. Category 1 of the decision aid includes medically actionable genes and SNPs with established disease prevention or treatment guidelines; Category 2 includes clinically valid genes and SNPs without actionable interventions; Category 3 includes genes associated with reproductive implications; Category 4 refers to variants of unknown significance. Participants select categories for receipt during the pre-disclosure counselling. Participants are strongly encouraged to learn Category 1, and must opt-out of receiving this category, but can opt-in to receive results of Categories 2 and 3. Category 4 is not available for receipt. Variants will be validated and approved for return by N.Y. State Department of Health. The counseling team reviews their IIDR and offers medical recommendations. Participants are followed for 1-week, 3-, 6- and 12- month intervals to determine the safety (self-reported psychological distress), perceived utility, adherence to recommendations, health behaviors and healthcare utilization associated with receiving IIDR from WG/ES. Preliminary analysis to estimate returnable results from our sequenced dataset demonstrates a low number of presumed pathogenic variants occurring within each category (zero to three), which suggests that the return of IIDR results will be feasible in a single session.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1842F

Integrating Exome/Genome Sequencing into Clinical Genetics Practice. *S.M. Jamal[1], H.K. Tabor[1,2], J. Conta[3], J. Yu[1], M.J. Bamshad[1,4].* 1) Department of Pediatrics, University of Washington, Seattle, WA; 2) Seattle Children's Research Institute, Seattle, WA; 3) Seattle Children's Hospital, Seattle, WA; 4) Department of Genome Sciences, University of Washington, Seattle, WA.

Exome sequencing (ES) and whole genome sequencing (WGS) have emerged as powerful tools for identifying novel disease genes, and are quickly being developed in service labs as an alternative to targeted genetic testing. Several widely publicized examples in which ES has been used to facilitate a clinical diagnosis have further stoked enthusiasm for adopting ES/WGS into clinical practice. We are similarly enthused but caution that translation of ES/WGS into clinical settings raises important questions about the circumstances in which ES/WGS is likely to be of greatest benefit, how to effectively obtain consent for clinical testing, how to control expectations and by what means results return should be managed for providers and families. While these questions frame the discussion, there is little empirical data about how to best answer them. This is a major obstacle to integrating ES/WGS into clinical care. To provide the clinical genetics community with guidance on these issues, we systematically reviewed the consenting process and the approach to ES/WGS results return used in a variety of research settings. Based on this review, we developed “core” content and procedures that we think begins to address the challenges of using ES/WGS in a clinical setting and sets the cornerstone for development of “best practices.” Furthermore, we reviewed the suggested indications for use of ES/WGS, informed consent documents and return of results reports from each lab in the U.S. offering diagnostic clinical ES/WGS. The presence of core content varied widely among the consent documents and results reports, and none contained all of the elements that we identified as the minimal required. For example, descriptions of what results would be returned (information from parental samples, incidental findings, adult-onset conditions for pediatric samples, and non-actionable results) and whether participants could opt out of certain kinds of results (e.g., “relationship confirmation” or “cancer”) were highly varied. The disparity of core content suggests a critical need for standardization, transparency and clarity in the information provided for clinicians and patients. This would improve the translation of ES/WGS to a range of clinical settings, improve outcomes and maximize clinical benefits while avoiding confusion, unnecessary expense, and miscommunication about the presence or absence of primary and unanticipated results.

1843W

A proposed approach to the integration of sequence data at a pediatric academic institution. *S.K. Savage[1], C.M. Clinton[1], I.A. Holm[1,2,3], I.S. Kohane[2,5,6], D.M. Margulies[1,2,4,5].* 1) Division of Genetics and Program in Genomics, Boston Children's Hospital, Boston, MA; 2) Department of Pediatrics, Harvard Medical School, Boston, MA; 3) The Manton Center for Orphan Disease Research, Boston Children's Hospital, Boston, MA; 4) Department of Developmental Medicine, Boston Children's Hospital, Boston, MA; 5) Center for Biomedical Informatics, Harvard Medical School, Boston, MA; 6) Children's Hospital Informatics Program, Boston Children's Hospital, Boston, MA.

The Gene Partnership at Boston Children's Hospital is a research effort focused on the disclosure of individual research results from a large genomic repository to promote personal and clinical benefit to participants. The return of results derived from whole exome/genome sequencing (WES) performed for research presents many challenges, in part due to the anticipated diversity of result types. In addition to research results, there are likely to be findings that have high clinical validity and are available in a clinical setting. Some findings may be directly relevant to the participant's clinical phenotypes, whereas others may be related to unanticipated phenotypes. As we are dealing with these issues in the research setting, clinical providers at our own institution are implementing WES as a diagnostic tool, and the same diversity of result types will likely be produced. To maximize the benefits to patients and research participants, it is important to develop consistent criteria for differentiating result types and a collaborative approach for the use and disclosure of data and results derived from WES. To this end we propose an approach for the integration of research results and data derived from WES into clinical care in a way that promotes participants' autonomy. The central element of our proposed approach is a rules-based decision-making model that manages physicians' access to data and results produced from WES for research and addresses our ethical obligation to protect participant privacy, promote autonomy, and facilitate quality care. Under this model, we propose that research results derived from WES that meet a designated threshold of analytic validity and clinical validity (determined through a scientific review process) be approved for return to the medical record. We also propose that participants in the research be given the opportunity to designate their preferences for results return to themselves, as well as their preferences for placing research results in the medical record. To facilitate appropriate clinical diagnostic testing, we also propose that physicians be granted access to the research-derived WES data for targeted reinterpretation, if coverage of the genes or regions that are directly pertinent to the patient's clinical presentation is of high-quality. This approach will allow researchers and clinicians to realize the full potential of DNA sequencing in a manner that is ethical, scalable, and clinically-beneficial.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1844F

Assessing Participant Preferences for the Return of Genomic Research Results. *L. Shahmirzadi[1], M.K. Cho[2, 3], S. Soo-Jin Lee[3], K.E. Ormond[1, 3].* 1) Department of Genetics, Stanford University, Stanford, CA; 2) Department of Pediatrics: Division of Medical Genetics, Stanford University, Stanford, CA; 3) Stanford Center for Biomedical Ethics, Stanford University, Stanford, CA.

Purpose: Advances in whole genome (WGS) and whole exome sequencing (WES) have provided vast amounts of information on a wide range of genomic results, posing practical challenges in determining what information to return to patients and/or research participants. Although several groups have proposed categories for returning genomic research results, there are no consensus guidelines in this area. The purpose of this exploratory study was to assess underlying preferences of participants of genomic research with regards to return of genome data. **Methods:** We performed an exploratory study on individuals enrolled in two WGS or WES studies at Stanford University in 2011–2012. Semi-structured telephone interviews were conducted to assess participants' views for the return of WGS/WES research results. Data were analyzed using grounded theory. **Results:** Four participants, all with a personal or family history of cardiovascular disease, completed interviews to date. Regarding motivations for participation in WGS/WES studies, participants expected that their enrollment would lead to: 1) information that may guide medical management for themselves or family members 2) answers to complex presentations. Participants also enrolled out of a sense of altruism. Overall, participants wanted to learn more about their condition and participate in research even without immediate benefit and believed that benefits outweighed the risks. Three preferred to receive comprehensive results of WES/WES, including all variants that have known medical risks, with and without current treatment, and with uncertain significance. One participant preferred not to receive results of uncertain significance or without current intervention because of the worry it may cause. One participant specifically voiced interest in learning about protective variants in addition to those associated with disease. Participants were also interested in learning other incidental findings that may arise, such as mistaken relationships, pharmacogenomic risks and protective genes. **Conclusion:** Our results suggest that while people with a personal or family history of disease report altruistic motives for participation in genome research, they are interested in receiving their own results and view benefits as outweighing risks. As this is a small study drawing from only one genomic research pool, it is unclear whether these results generalize to those without a history of disease.

1845W

The NHGRI/NIH Clinical Sequencing Exploratory Research Program. *B.A. Ozenberger[1], L.A. Hindorff[1], S.D. Schully[2], J. Boyer[1], N. Lockhart[1,2], L. Lund[1], C. Mahomva[1], A. Felsenfeld[1], J.E. McEwen[1].* 1) National Human Genome Research Institute, NIH, Bethesda, MD; 2) National Cancer Institute, NIH, Bethesda, MD.

In 2010, the National Human Genome Research Institute (NHGRI) undertook an extensive expert consultation to formulate a strategic forecast for genomic research and the emerging application of genomic approaches in medicine, culminating in publication of a vision for genomics research (Charting a Course for Genomic Medicine, NATURE Feb 11, 2011). This strategic plan recognized the potential benefits to patients of comprehensive genomic data that soon will be available to clinicians with the rapid deployment of new DNA sequencing instruments and methods. NHGRI subsequently crafted the Clinical Sequencing Exploratory Research initiative to: 1) leverage the Institute's long-standing experience in genomic sequencing and analysis to ease the adoption of these methods into clinical care, 2) guide the development and dissemination of best practices for the integration of clinical sequencing into clinical care, and 3) research the ethical, legal, and psychosocial implications of bringing broad genomic data into clinical decision-making including, for example, evaluation of the risks and potential benefits associated with the return of incidental findings or information on variants of weak effect. Six grants were awarded in late 2011 [to Baylor College of Medicine, Brigham and Women's Hospital, Children's Hospital of Philadelphia, Dana Farber Cancer Institute, U of North Carolina, and U of Washington], with the National Cancer Institute cofunding two of the awards. The consortium of grantees will expand to additional sites with new awards expected to be issued in 2013. Over the next four years, these groups will generate and incorporate sequence data in the clinical care of patients and examine the relevant ethical, legal, and psychosocial issues. Several Working Groups are collaboratively defining the state of the art and exploring opportunities to coordinate efforts within and outside the consortium in the areas of: Sequencing Standards, Electronic Medical Records, Phenotype and Analysis Measures, Actionable Variants/Return of Results, Informed Consent, Outcomes and Measures for the psychosocial component of this research, and special issues relating to pediatric populations. The anticipated products and results from the Clinical Sequencing Exploratory Research consortium, and its role in NHGRI's broader efforts relating to genomic medicine, will be described.

1846W

Genomic Incidental Findings: Perspectives of Geneticists and Genetic Counselors in the Context of Clinical Sequencing. *Z. Lohn[1], S. Adam[1], P.H. Birch[1], A. Townsend[2], J.M. Friedman[1].* 1) Department of Medical Genetics, University of British Columbia, Vancouver, Canada; 2) Department of Occupational Science and Occupational Therapy, University of British Columbia, Vancouver, Canada.

Clinical whole genome and whole exome sequencing are powerful tools for genetic diagnosis; however, incidental findings (IF) will also be discovered, which may have important medical and social implications. There is consensus in the literature that analytically valid and medically actionable IF should be returned to patients if practically possible, but whether additional findings warrant return is debated. Therefore, we aimed to determine whether geneticists and genetic counselors support the disclosure of various types of IF, and whether there is a difference in professionals' opinions towards the management of IF for pediatric versus adult patients. To explore this, an online cross-sectional questionnaire was distributed to members of the Canadian College of Medical Geneticists and the Canadian Association of Genetic Counsellors in November and December 2011. In total, 210 genetics professionals completed the questionnaire (response rate=42%), which used clinical scenarios to elicited opinions regarding degree of support for IF disclosure. The proportion of professionals who supported the disclosure of IF ranged from 95% for information pertaining to a serious and treatable condition to 12% for information with only social implications. There was discordance surrounding the disclosure of other types of IF including information of unknown significance, a serious and untreatable condition, and a multifactorial condition. Three factors were identified as being important to impacting disclosure decisions: condition-specific factors (e.g., treatment availability), test accuracy, and the evidence for pathogenicity. In regards to differences between pediatric and adult patients, there was significantly less support for disclosure of carrier status or information pertaining to an adult-onset condition for pediatric patients compared to adult patients. Some professionals expressed concern for pediatric autonomy when disclosing IF that are not directly relevant during childhood. These data suggest that genetics professionals might support development of independent IF policies for pediatric and adult patients. In conclusion, the disclosure of IF from clinical genomic sequencing is a complex and controversial issue among genetics professionals. Future research could incorporate more detailed clinical scenarios to fully explore the disclosure of IF from clinical sequencing investigations among genetics professionals.

1847W

Accounting for racial disparities in mortality among children with Down syndrome. *S. Santoro[1], A. Esbensen[2], L. Hendershot[2], F. Hickey[2], B. Patterson[2].* 1) Human Genetics, CCHMC, Cincinnati, OH; 2) Developmental and Behavioral Pediatrics, CCHMC, Cincinnati, OH.

Abstract: Aim: Life expectancies observed among individuals with Down syndrome (DS) are substantially shorter for African-Americans than Caucasians. There is a need to understand how racial disparities in comorbid health conditions contribute to these disparate life expectancies. Methods: Data were obtained from a DS specialty clinic at a Midwestern children's hospital. Medical history, pre- and post-natal history, current medical conditions, and referral patterns were obtained from 831 children with DS in a clinic sample (10.2% African-American). Results: Minimal racial group differences in comorbid medical conditions were observed. Group differences were observed in the frequency of nystagmus and PE tubes. Group differences were observed in the need for ongoing support from ophthalmology and cardiology. No differences were observed in the frequency of congenital heart disease (CHD), possibly suggesting that severity and not frequency of CHD warranted the ongoing cardiology support. Conclusions: The hypothesized racial disparity in comorbid conditions contributing to mortality was only minimally supported. Racial disparities in life expectancy are likely related to other factors warranting investigation, such as differential use of or access to health care services, or severity of comorbid health conditions. How use of health care services is related to severity of health conditions warrants future investigation.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1848W

ELSI 2.0: A new initiative to create an international infrastructure for the ethical, legal and social implications of genomic research. *K. Kato[1], J. Kaye[2], E.M. Meslin[3], B.M. Knoppers[4], E.T. Juengst[5], M. Deschenes[6], A. Cambon-Thomsen[7], D. Chalmers[8], J. De Vries[9], K. Edwards[10], N. Hoppe[11], A. Kent[12], C. Adebamowo[13], P. Marshall[14].* 1) Graduate School of Medicine, Osaka University, Suita, Osaka, Japan; 2) HeLEX, Department of Public Health, University of Oxford, Oxford UK; 3) Center for Bioethics, Indiana University, Bloomington, IN; 4) Centre for Genomics and Policy, McGill University, Montreal, Canada; 5) UNC Center for Bioethics, University of North Carolina, Chapel Hill, NC; 6) P3G, Montreal, Canada; 7) UMR 1027, INSERM, Epidemiology and analyses in public health, Toulouse, France; 8) Faculty of Law, University of Tasmania, Hobart, Tasmania, Australia; 9) Faculty of Health Sciences, University of Cape Town, South Africa; 10) Department of Bioethics and Humanities, University of Washington, Seattle, WA; 11) CELLS, Leibniz Universitaet, Hannover, Germany; 12) Genetic Alliance UK, London, UK; 13) Institute of Human Virology, Garki, Abuja, Nigeria; University of Maryland School of Medicine, Baltimore, MD; 14) Department of Bioethics, Case Western Reserve University, Cleveland, OH.

Genomic science is advancing rapidly and is increasingly international, proactive and coordinated, but so far ELSI research remains national, reactive and fragmented. However, given the exceptional scholarship, policy information and public engagement that have emerged from ELSI's conventional research methodology so far, important achievements are to be expected if ELSI research adopts the same strategic, large-scale and collaborative approaches that have characterized genomic science. This international initiative will develop an infrastructure to make large-scale ELSI research more efficient, effective and economical. The aim of this initiative is facilitate new collaborations, increase output, provide tools to help coordinate endeavours, and to avoid research redundancy. This will facilitate networking, critical reflection and the development of proactive strategies for international ELSI research in genomics. In doing so, it will provide maps of the international ELSI landscape to inform and coordinate future research, but also will also foster a new way of thinking and doing international ELSI research (Kaye, J. et al. *Science*, 336, 673–674, 2012). This paper will outline the pathway taken by ELSI programmes up to this point, identify shortcomings in conventional research approaches and build a new model on the basis of the lessons learnt from globalised genomics research. We will then present a concrete proposal for a research infrastructure and methodology, which will match the scale and output of ELSI research in genomics with that of other research in genomics, before concluding with an invitation to the genomics community to contribute to and build this ELSI initiative.

1849W

The challenges of defining a phenotype for addiction: In the words of addiction scientists. *J.B. McCormick[1, 2], J.E. Ostergren[2], R.R. Hammer[3], H.S. Harvey[2], M. Dingel[4], B.A. Koenig[5].* 1) Divisions of General Internal Medicine and Health Care Policy Research, Mayo Clinic, Rochester, MN; 2) Clinical Translational Science Activities, Biomedical Ethics Research Unit, Mayo Clinic, Rochester, MN; 3) Biomedical Ethics Research Unit, Mayo Clinic College of Medicine, Rochester, MN; 4) University of Minnesota Rochester, Rochester, MN; 5) Department of Social and Behavioral Sciences Institute for Health and Aging, University of California San Francisco.

The possibility of discovering a biological basis for addiction has prompted optimism, an influx of research funding, and a sense that the future of addiction research is fundamentally in neurogenetics research. However, like other frontiers of research, the development of novel treatments and preventive measures has lagged far behind the rate of discovery. In particular, there is disconnect between the complex phenotype of addiction and what we know about genes for susceptibility and drug targets. The utility of genetic research relies on an accurate account of the addiction phenotype. Yet the addiction phenotype remains a fuzzy concept because of its pleiotropic nature and the vast social and environmental factors that can influence the manifestation of the condition. Refining the phenotype of addiction and creating mechanisms for measuring it that can be shared broadly have been a priority for the National Institute on Drug Abuse (NIDA). To explore views on the state of addiction phenotype science, we interviewed 20 scientists conducting research on addiction in 2004–2005. Researchers expressed concern about the seemingly slow translation of findings. They discussed the need to better define addiction phenotype(s) and to develop tools that would enable them to find and measure existing and new phenotypes. In 2011 we interviewed 20 additional scientists conducting addiction research to explore this further. All interviews were analyzed using standard qualitative methods. Some said the Diagnostic and Statistical Manual (DSM) criteria were somewhat useful, but have “weaknesses.” They talked about molecular “pathway alteration” leading “to some phenotypic behaviors.” We heard that phenotype is “an outward expression,” but that it doesn't have to be. When asked whether genotype defines phenotype or vice versa, interviewees said that it is complicated: “these are interdependent entities.” Yet, some talked about how identifying genetic components to addiction would help “parse a lot of these fine differences” of phenotype. Several noted how useful epigenetics would be. Our findings show that addiction scientists acknowledge the fundamental tension between the socio-environmental elements of addiction and the neurogenetic basis, demonstrating a need for more careful evaluation of the relationship between the development of models of addiction, the phenotypes used, and the broader social environment in which addiction occurs.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1850W

Discriminating faces: Visual cues of ancestry. *J. Wagner[1], P. Claes[2], M. Shriver[3].* 1) University of Pennsylvania, Center for the Integration of Genetic Healthcare Technologies, Philadelphia, PA USA; 2) Katholic University of Leuven, KU Leuven, Belgium; 3) Pennsylvania State University, Department of Anthropology, University Park, PA, USA.

Racial disparities exist in health and justice. While there are no biological human races, cultural constructs of race (which have relied upon *perceived* biological differences) are real, many, and variable. Much variation (biological and cultural) is masked by such classifications. Racial discrimination, a known contributor to psychological and physiological health problems, is not an equal opportunity offender but is related to one's appearance and racial identity. How individuals self-identify and are identified by others is shaped to an unknown degree by our expectations of how an individual of a particular ancestry “should” look (e.g., laypersons note familial resemblances; cosmetic surgeons note patients wish to preserve racial identity; researchers report own race bias, other race effect, and cognitive adaptation). The level of discordance between one's appearance, perceived ancestry, and genomic ancestry may affect the formation and stability of one's racial identity; the form, frequency, and severity of discrimination one experiences; and even one's own prejudices. Yet relationships between identity, appearance, and genealogical, genomic, apparent, and perceived ancestry remain poorly understood. Using data from participants recruited in USA and Brazil (race; ancestry; 3D facial photos; skin and hair color; face shape; and genomic ancestry estimates), we investigated perceived ancestry from faces. Using 3D facial photos from 100 participants with mainly European and West African genomic ancestry, false color GIFs were prepared and incorporated into perception instruments. Observers were asked to provide their own information (e.g., race, ancestry, skin color, residence before age 18) before viewing 10 facial images with corresponding questions. In addition to exploring relationships between racial identity, appearance, and genealogical, genomic, and perceived ancestry of 100 research participants, data were analyzed to compare observers' ratings and explore how observers' characteristics influenced perceptions. N=822 observers participated. Perceived ancestry is correlated with face shape and genomic ancestry. Sigmoidal curves show both categorical and biased perception. By better understanding perceived ancestry, we may develop effective interventions to mitigate racial disparities in justice attributable to racism, colorism, and beautyism and may get one step closer to isolating factors contributing to racial disparities in health.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1851F

Integrating medical genetics into primary health care: report of a pilot program in Brazil. *T. Vieira[1, 2, 3, 4], C. Giugliani[5], L. Silva[1,2], L. Faccini[1,2,3,4], J. Leite[1,2], O. Artigalás[1,2,6], M. Lenz[7], M. Rojas[8], R. Giugliani[1,2,3,4].* 1) Medical Genetics Service, Hosp Clinicas Porto Alegre, Porto Alegre RS, Brazil; 2) WHO Collaborating Centre for the Development of Medical Genetics Services in Latin America, Porto Alegre RS, Brazil; 3) Postgraduate Program in Medical Sciences, Universidade Federal do Rio Grande do Sul, Porto Alegre RS, Brazil; 4) INaGeMP - Instituto Nacional de Genética Médica Populacional, Porto Alegre RS, Brazil; 5) Department of Social Medicine, Universidade Federal do Rio Grande do Sul School of Medicine, Porto Alegre RS, Brazil; 6) Children's Hospital, Grupo Hospitalar Conceição, Porto Alegre RS, Brazil; 7) Community Health Service, Grupo Hospitalar Conceição, Porto Alegre RS, Brazil; 8) Medical Genetics Service, Hospital de Clínicas de Porto Alegre, Porto Alegre RS, Brazil - Current affiliation: Medical Department, Genzyme do Brasil, São Paulo, Brazil.

Introduction: The integration of Medical Genetics (MG) and Primary Health Care (PHC) seems to be an alternative to develop actions of prevention and control of genetic conditions, and to facilitate community access to health care based on knowledge about genetics. Therefore, it is necessary that primary care professionals have a basic knowledge about genetics, the main conditions, management measures and prevention of genetics conditions. Aim: To evaluate if the application of an educational program in MG directed to PHC teams can help to integrate concepts and attitudes related to the identification, management and prevention of congenital malformations and genetic diseases. Method: The program was designed by health professional of the Medical Genetics Service/Hospital de Clínicas de Porto Alegre and applied to three Basic Health Unitis (BHU) of Grupo Hospitalar Conceição. The program consisted of meetings that took place at the BHU, which emphasis on the following topics: basic concepts; evaluation of families; teratogenic agents and prevention of birth defects; mental disability; inborn errors of metabolism and neonatal screening; cancer genetics; genetics of common chronic diseases; and genetic and preconception counseling. It was developed support materials for participants and educational material for the community. The program was evaluated through pre and post-test, as well through self-assessment questionnaire applied to the PHC professionals. Results: Forty three professionals participated in the program. In two of the three BHU it was found a statistically significant increase in the number of hits in post-test. Most professionals reported increased knowledge and confidence after the program, said that the program influenced their clinical practice and informed that the information resources and educational material presented during the program were useful. Conclusions: The program provided an approximation between GM and PHC, so that professionals pay attention to the genetic determinants of disease in the community. However it was realized the need for strategies for continuing education and ongoing support to the APS, with the aim of strengthening the integration of the two specialties, bringing to practice the community genetics. Support: March of Dimes.

1852F

The Gene Messenger impact project: a genetics e-learning program for family physicians. *J.C. Carroll[1], R. Grad[2], P. Pluye[2], J. Allanson[3], J. Permaul[1], N. Pimlott[4], B.J. Wilson[5].* 1) Department of Family Medicine, Mount Sinai Hospital, University of Toronto, Toronto, ON, Canada; 2) Department of Family Medicine, McGill University, Montreal, QC, Canada; 3) Children's Hospital of Eastern Ontario, Ottawa, ON, Canada; 4) Department of Family Medicine, Women's College Hospital, University of Toronto, Toronto, ON, Canada; 5) University of Ottawa, Ottawa, ON, Canada.

Context: Patients look to their family physicians (FPs) for guidance making informed choices about genetic testing. FPs express willingness to deliver some genetics services, but are challenged by lack of knowledge, and the rapid pace of genetic discovery. Our recent trial, GenetiKit, demonstrated effectiveness of a multi-component knowledge translation strategy on FPs' genetics referral decisions and confidence in core genetics competencies. We designed an additional project to disseminate and further evaluate one component, a knowledge support service called Gene Messenger (GM). GMs are 2-page reviews of genetic tests/disorders featured in the media. Objective: To determine the value of GMs as a genetics e-learning program for FPs. Design: FPs received 3 email invitations. After completing an online demographic questionnaire, participants received 12 GMs by email, one every 2 weeks for 6 months. Setting: Electronic participation. Participants: All English-speaking, active members of the College of Family Physicians of Canada were invited. Intervention: After reading each GM, participants completed an Information Assessment Method (IAM) questionnaire, a validated and useable evaluation tool for linking delivery of emailed information with continuing professional development. The IAM assesses participants' perceptions of the value of clinical information: cognitive impact, relevance, use, and expected health benefits. Demographic data and IAM ratings of GMs were analyzed and compared. Results: 1402 of 19,060 (7.4%) FPs participated. Most rated GMs as valuable: GMs will “improve practice” (mean=73%); are “relevant for a patient” (mean=79%); if relevant FPs will “apply GM information to patients” (mean=88%); and “expect health benefits” (mean=58%). How each GM would impact practice varied by topic. Overall 93% would like to continue to receive GMs. Findings are limited by participation rate, possibly affected by recruitment method. Conclusions: This appears to be an effective approach for supporting FPs' educational needs for practical up-to-date genetics information. FP feedback will be used to improve GMs, and explore sustainability of this e-learning genetics program.

1853F

Examining the Research Training of Recent Genetic Counseling Graduates. *S. Hahn[1], A. Rupchock[1], E. Bendik[1], E. Burkett[2], E. Heise[3].* 1) Hussman Institute for Human Genomics, University of Miami, Miami, FL; 2) Legacy Health System, Portland, OR; 3) Duke University Medical Center, Durham, NC.

The American Board of Genetic Counseling requires accredited programs incorporate research training in the curriculum, but programs are granted latitude regarding what and how this is incorporated. Previous research shows that among GCs specializing in research, the majority of skills and knowledge necessary for this role were learned on the job, rather than in their training program. In this study, we sought to determine how key skills and knowledge necessary for conducting research are currently included in GC program curriculum using a web-based survey of GCs who graduated within the last five years. Participants were recruited via the NSGC discussion forum, their program directors via the Association of Genetic Counseling Program Directors, and through an email to ABGC diplomats. The survey included questions regarding research training experiences, perceived competency with research skills at graduation, perceived importance of research training, and opinions about GC involvement in research. The survey was taken by 260 participants, representing at least 29 different GC training programs. Since graduating, 37% of respondents have held a position where they ascertained research subjects at least 10% of the time. Eighty-three percent of respondents conducted a research project that involved human research subjects as part of their genetic counseling training. Over 35% indicated research rotations were available as an option as part of their training, and, of those, 66% elected to complete this rotation. Research rotations and thesis requirements were rated as important or extremely important by 67% and 83%, respectively, and 40% of respondents would have liked more experience with human subjects research in graduate school. 96% feel genetic counselors should conduct their own research, and 91% feel counselors should obtain training in their programs so they can conduct their own research. These preliminary data illustrate the varying methods and degrees to which research is incorporated in training programs and may be used to guide the development of core competencies or GC program curriculum.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1854F

Increasing the number of URMs in Genomice Science. *D. Murray[1], D. Whittington[2], R. Gibbs[1].* 1) Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX; 2) Strategic Evaluations, Inc., Durham, NC.

The Human Genome Sequencing Center (HGSC) at Baylor College of Medicine provides training opportunities for under-represent minorities to interest them in a career in the genomic sciences. The HGSC provides exposure to genetic/genomic research and educational enhancement activities through a summer research program (HGSC-G/GREAT) and a post-baccalaureate program (HGSC-PGET). Since the program's inception in 2003, we have collected data on our program participants at the end of their respective program and through follow up evaluations. A Cross-Sectional Survey was used to generate background data and narrative data from the participants. Once collected the data was analyzed using description, analysis, and thematic development methods. Based on our analysis, we have improved our program through developing activities and recruitment strategies that will ensure the success of our participants. Our HGSC-PGET data indicates that over the last three years, our graduate school acceptance rate has increased with 60% of our research post-baccalaureate fellows and 70% of our laboratory technicians being accepted in doctorate programs. We have seen a 40% increase in the number of our summer interns enter Ph.D. programs in the last three years. The changes we have made to our undergraduate and post-baccalaureate opportunities has a positive impact on our ability to achieve our program goals.

1855F

Pharmacotherapy of Inherited Metabolic Diseases PharmD Fellowship Training Program. *J. Utz, W. Whitley.* Adv Therapies, Pediatrics, Univ Minnesota, Minneapolis, MN.

Background: Clinical pharmacy services (CPS), including Pharmaceutical Care (PC) or Medication Therapy Management (MTM), are provided in an increasing number of primary care settings, and have been shown to improve clinical outcomes and reduced overall healthcare costs in managing common chronic diseases. Although raining in provision of CPS for common chronic disease states is required curriculum in colleges of pharmacy in the United States, training in providing CPS in rare diseases has not been available. The orphan diseases known as inherited metabolic disorders (IMD) include conditions that require care from multiple specialty providers and treatment with numerous, costly medications, including orphan drugs. Objectives/ Methods: We instituted a 2-year post-doctoral PharmD training program that provided advanced clinical practice training for IMD, and explored the usefulness of CPS towards improving clinical outcomes for patients with IMDs. Results: During the 2-year training program, CPS were provided for patients with lysosomal storage diseases and PKU in the Advanced Therapies Clinic at the University of Minnesota. Clinical advancements for IMD patients that were realized during this training program included: 1) identifying 3 major types of infusion reactions to enzyme replacement therapies (ERT) through PC/MTM; 2) establishing methods for managing all 3 types of infusion reactions through PC/MTM; 3) instituting medication treatments for gangliosidoses; 4) improving neuropathic pain management in Fabry patients; 5) instituting a formal process for evaluating Fabry disease pain; 6) implementing systematic pharmacogenetic evaluations of medications for IMDs. Clinical research advances and educational experiences included: 1) initiating a clinical trial to study treatment of Fabry disease neuropathic pain; 2) submission of FDA IND for Fabry pain trial; 3) manuscript writing; 4) grant writing; 5) on-site industry training; 6) presenting abstracts at 4 international scientific meetings. Conclusions: The addition of a Doctor of Pharmacy (PharmD) to the clinical team in specialty clinics for rare diseases and institution of formal PharmD training programs for rare diseases enhances clinical program offerings and clinical research progress for rare diseases. Continued growth and expansion of rigorous professional PharmD training programs in rare diseases is warranted for the improvement of clinical outcomes in rare disease states.

1856F

Promoting student engagement and critical thinking in genetics: An evaluation of "Luck of the Draw" game. *M. Godfrey[1], A. Bower[2], K. L. Tracy[2], R. Anderson[1], A. Jameton[1], B. Arobba[3], C. Ryan[2].* 1) University of Nebraska Medical Center, Omaha, NE; 2) University of Nebraska, Omaha, NE; 3) Inter Tribal Buffalo Council, Rapid City, SD.

Students in the United States have fallen behind in international comparisons of math and science proficiency. Some educators have called for greater use of teaching methods that promote student engagement in science and math and that provide opportunities to develop critical thinking skills. One such method involves the use of classroom games. The current study represents initial efforts to evaluate "Luck of the Draw," a game designed to engage students in genetics and encourage critical thinking about issues, such as genetic engineering. Science teachers (N=24) attending a summer workshop responded to open-ended questions concerning their initial impressions of the game; middle and high school students (N= 134) completed questionnaires concerning their experiences playing the game. Results indicated that teachers and students enjoyed the game and identified positive outcomes, including greater engagement and critical thinking. Further, students' responses suggested that the potential problems that teachers thought might be encountered occurred infrequently.

1857F

Genetic Numeracy in College Students: A quantitative analysis. *M. Bergman[1], H. Honore Goltz[2].* 1) University of Washington, Seattle, WA. 9418 24th AV SW, Seattle, WA 98106; 2) University of Houston - Downtown Department of Social Sciences One Main Street, Houston, TX 77002.

Determining the genetic numeracy levels of populations is necessary for determining if, and where, interventions to improve genetic numeracy are needed. To date, very few quantitative studies of genetic numeracy have been performed (Lea et al., 2011). In this paper we assess quantitative levels of numeracy; identify differences in subgroups, if present; and determine if these differing levels impact behavior. This study is based on secondary analysis on data from the Perception of Genetic Risk in Sexual and Reproductive Decision-Making (PGRID) by College Students dataset (Honoré, 2008), a 138-item web-based survey instrument. The results of the study (N=2,568) imply that people believe that risk relationships with their parents follow inheritance patterns of autosomal dominant genes. Overestimating risk was another finding in this area of investigation. For example, many respondents believed that their inherited risk of disease from their parents was 70%. This was also the third most common response for inherited risk of disease from siblings. These estimated risk relationships implies that these people have poor understanding of genetic risk at it relates to their own family. Another implication is that many respondents do not believe that parents and children have reciprocal risk relationships. Education, especially didactic genetic information, seems to be a predictor for genetic understanding (Biological science degree OR = .73; Taken a genetics course OR = .76; Taken a course with genetic information OR = .70). Religiosity also seems to play a role in determining genetic numeracy (Services attended per week OR = 1.16) so respecting these differences in risk perception is important for public health genomics practitioners. Respondents in this study reported that they were increasingly unwilling to marry or have children perceived as having a genetic disorder (70% - Date; 58% - Marry; 45% - Have Children), but these trends do not seem to relate to their numeracy level. Study findings provide support for creating educational interventions that improve understanding of the mathematics of inheritance and the recognition of all the factors that can affect disease. Finally, improving patient/ provider communication relating to the patient's perception of risk is an important step in ensuring that the information received can impact patient behavior.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1858F

The Growing Stronger Organization and the Acondroplasia - Achondroplasia Blog: a parents' initiative to share knowledge and foster alliances towards potential therapeutic approaches for achondroplasia. *M. Kaisermann.* Consultant Physician, Rio de Janeiro, Rio de Janeiro, Brazil.

Background: Achondroplasia (ACH) is the most common cause of dwarfism. Despite the fast progress in the understanding of the mechanisms of this disorder, there is little research in potential therapies for ACH, either due to lack of official focus on this kind of disorder or because of its low prevalence. On the other hand, there are many web based sources available where interested people can find data about ACH. A single search in Google for *achondroplasia* retrieved 737K entries (Jun 2). However, with a number of few relevant exceptions, information is frequently outdated, misleading, fragmented, not referenced or is written in technical language, which can be difficult for non-specialists to understand. **Objective**: To describe the *Growing Stronger Initiative* (www.growingstronger.org, GS) and the blog **Acondroplasia-Achondroplasia** (www.tratando-acondroplasia.blogspot.-com). Discussion: **GS** is a non-for-profit organization founded by parents with the mission of improving the medical care of little people through supporting research. **GS** is establishing alliances to fund research towards therapeutic approaches for the several kinds of dwarfism and specially ACH, with the main goal of preventing or reducing the medical complications associated with this condition. The blog **Acondroplasia-Achondroplasia** is an independent, trilingual web based source of information about ACH directed to non-specialists, parents and affected individuals. It is linked to **GS** and offers articles in Portuguese, English and Spanish comprising reviews of genetics, molecular biology and potential pharmacological approaches to treat ACH. Articles are referenced and illustrated preferentially with web based accessible and validated scientific data, including direct links to peer reviewed articles, abstracts in Pubmed and animations provided by reliable sources (e.g.: Nature). It offers links to the main ACH Investigators' webpages; searches in Pubmed and ClinicalTrials.gov; links to relevant papers and to non-for-profit organizations devoted to support achondroplasia and / or other causes of dwarfism. **Conclusion**: By integrating the current information about ACH into a dynamic web based source containing adapted wording for the non-specialists, **GS** and the **Acondroplasia-Achondroplasia** blog can help interested individuals to improve their knowledge about the condition. Finally, they also aim to increase the research in new therapeutic approaches for ACH by directly funding research.

1859F

Genetics Home Reference, Ten Years In: Where We Are Now. *S.M. Morrison, M.L. Cheh, S. Calvo, H. Collins, J. Fun, K. Greenberg, L. Forman-Neall.* Lister Hill National Center for Biomedical Communications, U.S. National Library of Medicine, National Institutes of Health.

Genetics Home Reference (GHR) is an online resource from the U.S. National Library of Medicine (http://ghr.nlm.nih.gov/). This website provides an overview of genetic conditions and associated genes and is intended as a resource for patients, families, and the general public needing information about genomic medicine. GHR also provides background information including a primer (Help Me Understand Genetics), a glossary of genetic and medical terms, and links to numerous other quality resources. The “Spotlight” feature highlights important observances and discoveries in the field of human genetics and draws attention to useful learning tools and clinical resources. The GHR website was launched in 2003 with 19 condition summaries and 16 gene descriptions. Today, GHR offers easy-to-read summaries of more than 720 genetic conditions, more than 960 genes, around 65 gene families, all human chromosomes, and mitochondrial DNA. New summaries are added regularly. During the last ten years of development, GHR has grown in scope and content in order to better serve our audience. This presentation describes the improvements to the website, including new features, changes in focus, and new procedures for development and management, along with the rationale for their inclusion. We provide a discussion of future goals for the project and ways we plan to meet those goals. This work was supported by the Intramural Research Program of the National Institutes of Health, National Library of Medicine.

1860F

Investigating a conceptual change strategy to improve student understanding of basic genetics concepts in undergraduate non-science major and nursing student populations. *B. Bowling[1], M. Glassford[1], S. Barnes[1], S. Borgman[1], E. Reilly[2], T. Beery[3], C. Huether[4].* 1) Dept Biological Sci, Northern Kentucky University, Highland Height, KY; 2) University of Kentucky Markey Cancer Center, Lexington, KY; 3) College of Nursing, University of Cincinnati, Cincinnati, OH; 4) Dept Biological Sci, University of Cincinnati, Cincinnati, OH.

Replacing student misconceptions in genetics with the correct scientific conceptions can be difficult and may not be possible via conventional teaching strategies. Conceptual change texts (CCTs), brief passages that introduce a misconception, refute it, and explain the correct conception, are a possible intervention for correcting such student misconceptions. This study involved testing of CCTs to address the following misconceptions in undergraduate level courses: 1) DNA is different in different types of cells within one individual, 2) the terms gene, DNA, and chromosome can be used interchangeably, and 3) all mutations have a negative effect. Students enrolled at two Midwestern universities in both a non-science majors' general biology course and genetics for nursing course were randomly assigned to the control or experimental groups. Students in the experimental group received a CCT on the identified misconception, while students in the control group received information that discussed a topic from the course, but was not associated with the misconceptions. One hundred twenty students responded to three open-ended questions regarding the basic genetics concepts at the beginning of the course, at the time of reading the CCTs, and during the final week of class. The responses were scored using rubrics previously determined to be valid and reliable. While the post-course responses were on average higher for students completing the CCTs, statistically significant increases from pre- to post-course were observed in both the control and experimental groups. For all three questions, the post-course responses remained deficient with the average response including less than 50% of the key components of a comprehensive answer. Two additional aspects of the study will allow further investigation of the impact (or lack of impact) of the CCTs. A portion of the students are participating in a one-year post-course assessment to test the durability of their learning. In addition, content analysis will include identifying common themes in the responses and misconceptions held by the students throughout the different time points.

1861F

Alignment and assessment problems in the undergraduate genetics curriculum. *M. Dougherty[1,6], T. McElhinny[2,3], B. Bowling[5], J. Libarkin[2,4].* 1) Education, ASHG, Bethesda, MD; 2) Department of Geological Sciences, Michigan State University, MI; 3) Department of Zoology, Michigan State University, MI; 4) Center for Integrative Studies in General Science, Michigan State University, MI; 5) Department of Biological Sciences, Northern Kentucky University, Highland Heights, KY; 6) Department of Pediatrics, University of Colorado School of Medicine, CO.

`Backward design' is a model of curriculum development that relies on identifying learning outcomes and defining what constitutes evidence of learning before planning the teaching. Although backward design is widely considered best practice, it is often overlooked by university faculty and may help explain the inconsistency between faculty teaching behavior and the genetics concepts they claim are most important. This paper will review the state of genetics instruction in the United States from the perspective of backward design, with particular attention to the goals and assessments that inform curricular practice. An analysis of syllabi and leading textbooks indicates that genetics instruction focuses most strongly on the structure and function of DNA and Mendelian genetics. At the same time, a survey of faculty indicates that other concepts, such as the application of genetics to society or the environment, are viewed as equally or even more important than certain foundational concepts. This disconnect suggests a need for more explicit goal setting prior to curriculum development. Preliminary analysis of existing assessments, specifically concept inventories developed for higher education, indicates that assessments are poorly aligned with faculty goals for instruction and need to be modified into more valid and reliable measures of student conceptions.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1862F

YouTube as an educational resource for visual and kinesthetic learners: A study of DNA replication animations. *J.M. Lind.* University of Western Sydney, School of Medicine, NSW, Australia.

Background: Learning styles can be broadly classified into visual (diagrams), aural (listening), read-write (information displayed as words), and kinaesthetic (real/simulated events). Lecturing is a common mode of content delivery within undergraduate genetics teaching, which traditionally caters for aural and read-write learners. Students and educators can now easily access resources that assist visual and kinesthetic learning via websites such as YouTube, a video sharing website. This study aimed to determine how the medium of YouTube has been utilised in the demonstration of genetics fundamentals by way of content upload and content viewing, with a focus on DNA replication animations. **Methods:** A search of www.youtube.-com using the criterion “DNA replication animation” was performed. Videos uploaded between the inception of YouTube (February, 2005) and December 31, 2010, with primary content about DNA replication were included. Only short videos (less than 20 mins) were included. Date of video upload, number of times viewed, and length of video were collected. Data for videos uploaded multiple times were combined to get a total number of views per video, using the first upload occurrence as the date uploaded. A general linear model was used to test for associations, with a significance level of p=0.05. **Results:** A total of 127 unique videos, viewed a total of 21,264,244 times, were included in the analysis. The first video was uploaded on May 3, 2006. As of June 4, 2012, the number of views per video ranged from 15 to 2,891,443, with an average of 41,787 views per video, per year. The number of new videos uploaded each year ranged from 17 in 2006 to 33 in 2009, with a mean of 25.4. There was a significant association between the time since upload and the number of views, with an additional 360.8 views for each additional day the video had been online (95%CI 220.6–501.1, p<0.001). There was no association between length of video and number of views. **Conclusions:** YouTube has been extensively utilised to demonstrate the fundamentals of human genetics. It has enabled students and educators to easily access resources for visual and kinesthetic learners to study processes occurring at the molecular level. YouTube content can enhance student learning by appealing to a broad range of learning styles.

1863F

Genetic Counseling in India: Current certification courses and a need for the development of a Master's training program. *K. Shah.* SEATTLE CHILDREN'S HOSPITAL, SEATTLE, WA.

Non-communicable diseases such as diabetes, cardiovascular disease, chronic respiratory disease and cancer are becoming significant in India since their incidence is increasing. While infectious diseases remain the leading causes of death in the earlier years of life, injury and non-communicable chronic disease predominate in older age groups. With the hope of the potential decline in communicable diseases, the emerging importance of genetic disease is inevitable. The access to clinical genetic services in India is mostly confined to the urban areas. Genetic counseling has been traditionally conducted by physicians in India. While some genetic testing centers or institutions based in a university setting have implemented genetic counseling workshops or courses for their clinicians; an official certification/diploma course in genetic counseling has only been recently initiated (as of January 2012) in India as a full time 1 year course. The two new genetic counseling certification programs in India could aid in generating an educated genetic counseling work force that can meet patient demands for genetic services. We will attempt to compare the curriculum offered by the 2 institutions, including their target student admission criteria and hope that these programs can serve as a potential model for developing a comprehensive 2 year Masters' program in genetic counseling in India, similar to the ones offered in the United States.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1864F

A strategy of applying comprehensive therapeutic approaches for training autism children. *H. Xu[1], N. Zhong[2].* 1) Dept. Children Health Care, Hubei Provincial Maternal and Children's Hospital, Wuhan, China; 2) New York State Institute for Basic Research in Developmental Disabilities, Staten Island, NY.

Objective: To explore a comprehensive training strategy for autism. Methods: 185 Children with core autism aged from 2.5 to 6.2 years were given comprehensive training between October 2009 and December 2010 in Hubei maternal and children's hospital. The training time was 40 hours per week, and lasted 6 months. Before and after treatment, autistic children were evaluated with both Autism Treatment Evaluation Checklist (ATEC) and Revision of Chinese Version of Psycho-Educational Profile (C-PEP). Results: The average ATEC total score of all children with autism was significantly decreased from (99. 12±15. 95) to (53. 91 ±11. 24) (P <0. 001). The subscale scores including language, social interaction, sensory and behavior were all decreased (P < 0. 05, P < 0. 001) as well. The average C-PEP score was increased from (30. 17 ±9. 58) to (49. 84 ±10. 17). The equivalent of age increased (9. 67±0. 59) months. When 6 months physiological development subtracted, it actually increased (3. 67 ±5. 32) months. Conclusions: Our results support the fact that for children with autism, the comprehensive training is a rather effective approach.

1865F

Genetics and Public Health: The Experience of a Referral Center for Diagnosis of 22q11.2 Deletion in Brazil and Suggestions for Implementing Genetic testing. *T. Vieira, I. Sgardioli, V. Gil-da-Silva-Lopes.* Medical Genetics, University of Campinas, Campinas, São Paulo, Brazil.

The introduction of new technologies of molecular diagnosis for health care has been a challenge in the last years, especially in Brazil, where the majority of the population is served by the public health system. The 22q11.1 deletion syndrome is the most common syndrome that has palatal anomalies as a major feature, with a prevalence of 1/4000 births. Considering this prevalence, the Brazilian health system characteristics and the current situation of medical and clinical genetic services in the country, the main aim of this study was to conduct a multicenter study for 22q11.2 deletion diagnosis as a model for the optimization of diagnostic strategies in medical genetics. We investigated the access to laboratorial diagnosis of 22q11.2 deletion at 11 genetic services and centralized this diagnosis for 100 patients with palatal abnormalities and suspicion of 22q11.2 deletion syndrome, referred from these centers during 30 months, at the Cytogenetics and Molecular Biology laboratories of the FCM/UNICAMP. To detect 22q11 deletions FISH (Fluorescence in situ Hibridization) and MLPA (Multiplex Ligation-dependent Probe Amplification) techniques were used. Previous and temporary availability for the diagnosis of 22q11 deletion, associated with research projects, was mentioned by seven centers, with remarkable geographic disparities. We detected 22q11 deletion in 35% of the patients, and chromosome abnormalities not related to 22q11 region in three patients; thus we reached diagnostic conclusion in 38% of the cases. There was significant difference between some clinical signs found in patients with or without 22q11 deletion. There was 100% of sensibility and specificity for both MLPA and FISH techniques. Considering the required infrastructure and the modifications in the FISH (allowing to reduce probe quantity), this technique was efficient, more economical and faster than MLPA. Centralizing the laboratorial diagnosis was considered advantageous, pointing to this model as an important and feasible strategy for genetic diagnosis in Brazil. These results allowed to suggestions for the improvement of laboratorial diagnosis of this and other genetic conditions in our country. Financial support: FAPESP, CAPES and CNPq.

1866F

Outreach, genetic counseling and testing of patients at risk for hereditary breast and ovarian cancer (HBOC) ascertained through the use of the electronic medical record. *H. Kershberg[1], E. Parkhurst[1], M. Alvarado[1], H. Zhou[2], J. Natoli[3], G. Tiller[1].* 1) Regional Genetics, Kaiser Permanente Southern California, Pasadena, CA; 2) Dept. of Research and Evaluation, Kaiser Permanente Southern California, Pasadena, CA; 3) Dept. of Clinical Analysis, Kaiser Permanente Southern California, Pasadena, CA.

Knowledge of mutation status has health management implications for patients and their family members, and the National Comprehensive Cancer Network (NCCN) recommends that women diagnosed with breast cancer under age 46 years be referred for genetic counseling. The genetics department in Kaiser Permanente Southern California (KPSC) provides clinical services in an integrated health care system serving nearly 3.5 million members. Using data from multiple sources including: KPSC tumor registry, a departmental cancer test results database, and the KPSC electronic medical record system, we noted that less than half of women diagnosed with early breast cancer had received genetic counseling and the offer of genetic testing. Using the same data sources, we identified 455 early breast cancer patients who were diagnosed between September, 2005 and September, 2010 and had not received genetic counseling. We sent these women a letter and followed up with a phone call offering a genetics consultation, and we offered BRCA testing to all those who came for counseling. **Results**: 32% (n=145) of women came for a genetic counseling appointment and were offered genetic testing. 83% (n=120) of those offered genetic testing underwent BRCA mutation analysis. 5 % (n=6) were positive for a deleterious mutation and 5% (n=6) had a variant of uncertain significance. The rate of acceptance of the genetic consultation and subsequent testing showed some significant differences. Hispanics were almost twice as likely to respond to the outreach and schedule an appointment compared to Caucasians (OR =1.87, 95% CI, 1.02, 3.43). Women diagnosed with cancer within a year of contact were 4 times as likely to schedule the genetic consultation than those diagnosed more than 5 years before (OR= 4.06, 95%CI, 1.03, 16.02). Of those who came in for counseling, African Americans were significantly less likely to accept genetic testing than Caucasians (OR = 0.25, 95%CI, (0.09, 0.73). **Conclusions:** This project demonstrates how an integrated care approach and electronic medical record system provide a unique opportunity to identify and contact women who are at increased risk. In view of the varied response rates among different racial/ethnic groups, multiple approaches may be required to reach diverse populations. Similar methods could be used to ascertain patients at risk for other hereditary cancers and genetic disorders.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1867F

Prevalence of Down syndrome in the United States. *A.P. Presson[1, 2], G. Partyka[3], K.M. Jensen[3, 4], S.A. Rasmussen[5], O.J. Devine[5], L. McCabe[3], E.R.B. McCabe[3].* 1) Division of Epidemiology, University of Utah, School of Medicine, Salt Lake City, UT; 2) Department of Biostatistics, UCLA, Los Angeles, CA; 3) Department of Pediatrics, University of Colorado School of Medicine, Aurora, CO; 4) Department of Internal Medicine, University of Colorado School of Medicine, Aurora, CO; 5) Centers for Disease Control and Prevention, Atlanta, GA.

Background: Life expectancy of people with Down syndrome (DS) in the United States (US) has increased remarkably over the past four decades. While births and deaths among people with DS have been studied in detail, we currently lack a reliable estimate of the population prevalence of DS. Using data available from the past century and projections of US population size, we estimated the prevalence of individuals with DS in 2008 and project these estimates to 2100. Objective: To estimate prevalence of DS in the US on January 1, 2008 and to project the prevalence of DS in 2100. Methods: Utilizing CDC data for annual US total births (1909–1939), births by maternal age (1940–2008) and published rates of DS by maternal age, we estimated annual births with DS (1909–2008). We estimated annual deaths by age for people with DS for 1900–1967 through extrapolation from available death certificate data (1968–2007). Using a time-specific life table approach, we estimated the US prevalence for DS on January 1, 2008 and applied Monte Carlo sampling methods to form 90% uncertainty intervals (UI) for our estimates. We then estimated the prevalence of DS in 2100 using the January 1, 2008 prevalence estimate and the middle and high projections of the US population size from the US Census Bureau, with the assumptions that DS prevalence would increase uniformly with the US population and there would be no change in the current rate of pregnancy terminations. Results: The estimated US prevalence of people with DS on January 1, 2008 was 240,500 (90% UI: 177,700 - 308,800) based on proportions of deaths by age from the most recent two years (2006–2007) of death certificate data. This estimate corresponded to a prevalence of 7.94 (90% UI: 5.86–10.19) people with DS per 10,000 in the US population. We projected an estimated 465,700–936,300 people with DS in 2100. Conclusion: Our estimate of DS prevalence is 30–40% lower than estimates based solely on current birth prevalence, but is similar to estimates in European countries. Survival among people with DS has increased dramatically since 1970, which suggests continued growth in DS prevalence over the next century. Thus, our projection of 465,700–936,300 people with DS in 2100 could be an underestimation. A more realistic projection based on expert opinion is warranted and planned. More accurate prevalence estimates are critical to facilitate planning for services for people with DS.

1868F

Rare Disease Initiatives in Europe: Centres of Expertise in Germany. *C. Huebner[1], M. Stuhrmann-Spangenberg[1], C. Zeidler[2], J. Schmidtke[1].* 1) Institue of Human Genetics, Hannover Medical School, Hannover, Niedersachsen, Germany; 2) Molecular Hematopoiesis and Pediatric Hematology and Oncology, Hannover Medical School, Hannover, Germany.

A disease which affects not more than 5 in 10,000 people is considered to be rare. Therefore 7,000–8,000 of 30,000 known diseases are categorized as rare diseases, with a great majority being of genetic origin. According to WHO estimates some 30 million people suffer from a rare disease in the European Union. After the European Council recommended an action in the field of rare diseases in 2009, national strategies were set up in many European countries, including the formation of centres of expertise. The Rare Disease Center at Hannover Medical School was initiated by the Department of Human Genetics and opened in November 2011. An interdisciplinary facility where research and care for patients with known diagnosis and patients with chronic conditions not yet having received a definite diagnosis was set up. All patient referrals including self-referrals have been documented in detail. A patient registry is being developed. Between 30 and 50 paediatric and adult patients were referred per month with a rising tendency over time. Preliminary data suggest that the majority of requests are self referrals (88% after 6 month) and only about 10% are referrals from primary care physicians. More than half of the patients did not have a reliable diagnosis at referral. As expected most of the undiagnosed patients were asking for a diagnosis and patients with known diagnosis were asking for experts. Interestingly more than 30% of patients were seeking for information, indicating a lack of information available on rare diseases. As a unique feature the German Orphanet team is part of the Hannover Rare Disease Centre. Orphanet is the reference portal for information on rare diseases and orphan drugs particularly in Europe. Here, we will report our experience of the first year, the organizational structure and financing of the Rare Disease Centre at Hannover Medical School including its 11 subcentres and specialized clinics, describe internal patient paths, diagnostic successes and failures, enrolment of patients in research projects, and compare our structure and experience with the few other centres in Germany and elsewhere, from which data are available. We will present an outlook into future regional, national and international networking of rare disease centres.

1869F

Implementation of Universal Lynch Syndrome Screening in an Integrated Health-Care Delivery System. *K.A.B. Goddard, T. Kauffman, M. Morse, J. Davis, C. Young, E. Esterberg, C. McGinley, J. Reiss.* Kaiser Permanente Northwest, Portland, OR.

Screening tests for Lynch syndrome (LS) have been recommended as an evidence-based practice which decreases disease burden and save lives. Nationally, more than half of patients meeting criteria do not receive LS screening. In a recent study of seven integrated health-care delivery systems, fewer than 5% of patients diagnosed with colorectal cancer (CRC) received screening for LS. How do we effectively improve access to LS screening? We hypothesize that universal LS laboratory-based screening will improve appropriate referrals for subsequent genetic counseling and consideration for additional testing, compared with the current practice of self-referral or provider-based referrals to medical genetics. Through a randomized-controlled trial, this project evaluates a strategy to maximize screening of CRC patients for LS. We will invite all CRC patients undergoing colon surgery to join the trial and compare universal LS screening in the intervention arm to usual care. We report on the development of processes and resources for the implementation of this screening program. We developed an algorithm to identify eligible patients undergoing bowel surgery daily from the electronic medical record. We implemented a procedure to order LS screening for patients randomized to the intervention arm, and to integrate a procedure to prepare tissue for microsatellite instability (MSI) testing within the pathology department procedures for managing the surgical specimens for the case. MSI test results are returned from an outside laboratory directly to medical genetics, and the laboratory report is integrated within the electronic medical record. We created genetic counseling educational talking points and result disclosure processes. Program implementation is being measured through a variety of mechanisms including interviews with clinical staff, patient surveys, and monitoring of downstream utilization through the electronic medical record. Evaluation of the steps required for implementation will aide in the determination of best strategies for successful universal LS screening, and will inform future implementation efforts at new sites considering this approach for LS screening.

1870F

Beyond NGS technology - overcoming key challenges for nation-wide implementation of next-generation sequencing into routine diagnostics. *T. Vrijenhoek on behalf of the Centre for Genome Diagnostics.* Centre for Genome Diagnostics, Utrecht, Netherlands.

We present an approach to implement next-generation sequencing (NGS) in routine diagnostics across clinical genetic centers. As diagnostic application of NGS has become **technically feasible**, many institutes have individually set up experimental and data analysis procedures accordingly. However, if we want to develop NGS into a **high-quality clinical decision tool** in national healthcare, we need to think beyond individual centers. At that point, issues like **standards** for diagnostic procedures, (legal) **agreement** on the use of personal genomic data, and **harmonization** of consent procedures are to be solved. The clinical genetics community in the Netherlands - a frontier country in self-regulation of technology-driven diagnostics - has overcome such cross-center challenges. All of the nation's clinical genomic research groups and diagnostic centers, as well as technology and service providers participate in the Centre for Genome Diagnostics, aiming to **identify and tackle** multi-center bottlenecks and challenges. Collectively we have developed an implementation approach based on e.g. experimental guidelines from the Board of Dutch DNA diagnostics laboratories (LOD), and circulating consent procedures within the Netherlands' Association of Clinical Geneticsts (VKGN). We piloted the approach in a joint project (CARDIO), thereby generating concrete data on how NGS-based diagnostics works in practice. CARDIO comprised 10 patients with cardiomyopathy, of whom DNA samples and medical records were sent to all clinical genetic centers. These performed comprehensive NGS-based diagnostics on all patients, from clinical assessment to diagnostic counseling. The centers were free in using their preferred approach of enrichment, sequencing, data analysis, and diagnostic interpretation. They all carefully reported on their internal diagnostic process. Based on 32 reports and detailed follow-up we were able (1) to assess bottlenecks, challenges and best practices based on concrete and concordant results, not on hypothetical situations; (2) to identify the key people in the diagnostic process and especially their position in the organization; and (3) to develop a framework for optimal coordination and communication among research groups and diagnostics centers. With the results of project CARDIO at hand, we present a **model for optimal transition** of NGS technology from research to diagnostics, regardless of technology.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1871F

Inherited Neuropathies Consortium: A Collaborative Approach to Genetic Research in Charcot-Marie-Tooth (CMT) disease. *F. Speziani[1], S.M.E. Feely[2], C. Siskind[3], S.H. Blanton[1], G. Beecham[1], J.M. Vance[1], M. Shy[2], S. Zuchner[1], Inherited Neuropathies Consortium.* 1) Hussman Institute for Human Genomics, University of Miami, Miami, FL; 2) University of Iowa, Iowa City, IA; 3) Stanford University, Palo Alto, CA.

Charcot-Marie-Tooth (CMT) disease encompasses a group of clinically and genetically heterogeneous hereditary peripheral neuropathies. Despite the identification of >50 CMT genes, only 30% of axonal CMT2 patients have mutations in known genes. Further, variable phenotypic expressions have long been noted within and across CMT families and patients, suggesting unknown genetic modifying factors. To efficiently approach these questions in a rare disease such as CMT, large-scale collaborations are necessary. With the support of NINDS we have formed the Inherited Neuropathies Consortium (INC), which is a member of the Rare Disease Clinical Research Network (RDCRN). The INC has currently 13 national and international sites led by prominent CMT investigators. The clinical aims of the INC include the development of standardized clinical measures for adult and pediatric CMT and natural history studies. Genetic studies are focused on the goal of exome sequencing 1,000 CMT2 samples and a genome wide association study of genomic modifiers of clinical outcome in CMT1A. INC members have organized a working structure that includes standardized IRB protocols, centralized DNA isolation/banking, and web-based data management portals. During the first three years of the INC, 2,120 patients have been registered; with 985 patients recruited in the past year. For genetic modifiers in CMT1A we have completed genotyping for the first GWAS analysis in CMT, which is currently under analysis. To date, 49 axonal CMT families from ten INC sites have had whole exome sequencing. Several novel candidate CMT genes are currently being followed up. We will present the unique and growing structure of the INC as a promising model for research of rare diseases.

1872F

Genomic Technologies: Effective Integration into Pediatric Clinical Care. *R. Hayeems[1,3], N. Hoang[1], S. Chenier[1], J. Stavropoulos[1], S. Pu[1], S. Wodak[1], R. Babul-Hirji[1], J. Davies[2], L. Velsher[2], J. Aw[2], R. Weksberg[1], C. Shuman[1].* 1) Clinical and Metabolic Genetics, Hospital for Sick Children, Toronto, Ontario, Canada; 2) Genetics, Medcan Clinic, Toronto, Ontario, Canada; 3) Institute of Health Policy Management and Evaluation, University of Toronto, Ontario, Canada.

Clinical interpretation, genetic counseling, and medical management for pediatric chromosome microarray (CMA) present ongoing challenges. To guide current practice and to inform future strategies for genome-wide analyses, the impact of CMA technology on clinical care warrants evaluation. Our study aims to understand the impact of CMA on pediatric medical management. A retrospective chart review of children who underwent CMA in 2010–2011 is underway. 900 charts were selected for review, 300 from each CMA result type: pathogenic, uncertain, benign. Data elements include: age, gender, ethnicity, phenotype, family history, microarray results, and medical investigations immediately prior to and following microarray testing. Descriptive and multivariate analyses are being used to understand the association between microarray results and post-test medical investigations. We hypothesize that an increase in medical investigations will be observed in children with pathogenic copy number variants (CNVs) compared to those with uncertain or benign CNVs. 129 charts have been reviewed to date (146 CMA results: 30.1% pathogenic, 37.7% uncertain, 32.2% benign). Preliminary results show that 44.9% of post-array investigations occurred in the pathogenic variant group, 34.1% occurred in the uncertain variant group, and 21.0% occurred in the benign variant group. For pathogenic CNVs, 53.4% of investigations were prompted by the CNV itself and 46.6% were prompted by phenotype. For non-pathogenic CNVs, 25.4% of investigations were prompted by the CNV and 74.6% of investigations were prompted by phenotype and diagnostic quest (p=0.0001). Our findings suggest that for pathogenic CNVs, medical investigations are prompted equally by the CNV finding itself and phenotype. For non-pathogenic CNVs, the majority of medical investigations are prompted by phenotype and diagnostic quest. These findings will be instructive in anticipating the impact of next generation genomic analyses on medical management and downstream use of health services.

1873F

Genomic Technologies: Effective Integration into Adult Preventive Medicine. *N. Hoang[1], R. Hayeems[1,3], J. Davies[2], L. Velsher[2], J. Aw[2], S. Chenier[1], J. Stavropoulos[1], S. Pu[1], S. Wodak[1], R. Babul-Hirji[1], R. Weksberg[1], C. Shuman[1].* 1) Clinical and Metabolic Genetics, Hospital for Sick Children, Toronto, Ontario, Canada; 2) Genetics, Medcan Clinic, Toronto, Ontario, Canada; 3) Institute of Health Policy Management and Evaluation, University of Toronto, Toronto, Canada.

Genome-wide single nucleotide polymorphism (SNP) arrays that assess individual genetic risk for common polygenic diseases can be useful tools to guide proactive health screening and overall preventive health care. While evidence is emerging related to why consumers are motivated to purchase genome-wide testing and how they are integrating results into health-related decision-making, little is known about how array information impacts clinical recommendations and health services used. Navigenics Health Compass is one example of such tests; it is currently available at Medcan, an executive health clinic in Toronto. Our study aims to understand how Navigenics risk results relate to medical recommendations and subsequent health service utilization. We are conducting a retrospective medical record review for two groups of Medcan clients, dating back to 2009: those who have pursued Navigenics testing (cases) and those who have not (controls). Navigenics risk estimates for nine conditions will be examined, including: abdominal aneurysm, atrial fibrillation, celiac disease, colon cancer, type 2 diabetes, glaucoma, heart attack, melanoma, and prostate cancer. Client data are being collected on: age, gender, ethnicity, clinical phenotype, environmental modifiers (exercise level, alcohol intake, and smoking behaviours), family history, medical recommendations, and services used within one year of receiving Navigenics results. We will analyze the association between Navigenics risk estimates and (i) medical recommendations and (ii) services used for each condition, controlling for the other variables. We hypothesize that for each condition, medical recommendations and services used will be increased for patients with elevated Navigenics risk estimates compared to those with average/reduced risk estimates and compared to controls. Analyses to date will be presented. Understanding the impact of genome-wide SNP arrays (i.e. used for individual risk assessment of common polygenic conditions) on downstream clinical management and health service utilization will inform strategies for effective delivery of these new technologies.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1874F

Evaluating laboratories' interpretation of genetic test results through a EuroGentest Proficiency Testing survey. *E. Girodon[1,11], S. Berwouts[2,11], M. Rodriguez de Alba Freiria[3,12], AF. Roux[4,13], D. Barton[5,13], L. Hoefsloot[6,13], S. Gardner[7,14], O. Kamarainen[7,13], B. Fowler[8,14], M.A. Morris[9,11], E. Dequeker[2,11], R. Hastings[10,12].* 1) Genetics, GH Henri Mondor APHP and INSERM U955, Creteil, France; 2) University of Leuven, Department of Public Health, Biomedical Quality Assurance Research Unit, Leuven, Belgium; 3) Fundacion Jimenez Diaz, Genetics Department, Madrid, Spain; 4) CHU de Montpellier et IURC, Laboratory of Molecular genetics, Montpellier, France; 5) National Centre for Medical Genetics and School of Medicine & Medical Sciences, University College Dublin, Dublin, Ireland; 6) Radpoud University Nijmegen Medical Centre, Department of Human Genetics, Nijmegen, The Netherlands; 7) Genetic Medicine, St Mary's Hospital, Manchester, United Kingdom; 8) Paediatrics, University Children's Hospital, Basel, Switzerland; 9) Hôpitaux Universitaires de Genève, Département de Médecine Génétique et de Laboratoire, Genève, Switzerland; 10) John Radcliffe Hospital, Oxford University Hospitals NHS Trust, Women's Centre, Oxford, United Kingdom; 11) Cystic Fibrosis European Network (CF Network); 12) Cytogenetic European Quality Assessment (CEQA); 13) European Molecular Quality Network (EMQN); 14) European Research Network for evaluation and improvement of screening, Diagnosis and treatment of Inherited Disorders of Metabolism (ERNDIM).

Interpretation of test results is a key component of a genetic laboratory report, ensuring full understanding of the results by the referring physician, including potential implications for the patient and the family. Proficiency Testing (PT) or External Quality Assessment (EQA), in which many laboratories report on the same clinical scenario, provides an ideal tool to study reporting practices. In order to evaluate laboratory performance in this area, a EuroGentest survey among PT programs from four European providers (CF Network, CEQA, EMQN & ERNDIM) was set up, with the aim of determining the interpretation elements that were missing in laboratory reports, in regard to existing guidelines and recommendations (*e.g.* OECD). PT reports from 2011 representing 422 laboratory participants and relating to six different PT programs: cystic fibrosis, Friedreich ataxia, hereditary deafness at the DFNB1 locus, postnatal microarray (CEQA), prenatal diagnosis of chromosomal abnormalities on amniotic fluid and metabolic diseases, were evaluated using criteria derived from the OECD Guidelines for Quality Assurance in Molecular Genetic Testing. Mean interpretation scores varied between cases, ranging from 1.61/2 to 1.98/2. Despite heterogeneity in the scoring systems and in assessment for the different schemes, the data analysis revealed the following: 1) although the answer to the question was provided in the majority of reports, the survey emphasized the need to clearly restate the indication of testing in reports; 2) items related to follow-up testing were the most critical: search for a 2nd mutation was missing in 36% of DFNB1 reports, testing the parents to confirm compound heterozygosity was missing in 29% of CF reports and need for parental bloods was not mentioned in up to 27% of prenatal cytogenetics reports (where applicable). Moreover, confirmation of unexpected results was not recommended in 75% of CF reports. A notable proportion of molecular genetics (~36% in FRDA, 28% in DFNB1 and 26% in CF) and cytogenetics (~29%) reports did not mention the recurrence risk of disease and/or the possibility of prenatal diagnosis for future pregnancies. This first and preliminary horizontal survey among different PT programs providers emphasizes the need to find effective training and communication processes, including harmonization of assessment between PT programs, so that laboratories can improve the interpretation content of their reports.

1875F

Interstate Compatibility and Emergency Preparedness of Newborn Screening Laboratories at the Interface of Clinical Medicine and Public Health. *J. Mulvihill[1], P. Hopkins[2], S. Berberich[3].* 1) Dept Pediatrics, University of Oklahoma, Oklahoma City, OK; 2) Missouri Public Health Laboratory, Jefferson City, MO; 3) State Hygienic Laboratory, University of Iowa, Iowa City, IA.

Clinical decisions about laboratory results for newborn screening assays should be identical across the US. Emergency preparedness is essential to maintaining the integrity and capacity of state-run newborn screening programs and laboratories. We assessed results of parallel assays by different state laboratories on the same dried blood spot—thereby, truly validating the clinical equivalence of results. To do so, we executed drills between such laboratories to rehearse continuity in the event of an unanticipated collapse of one state's program. After numerous in-person and electronic meetings with all stakeholders to dissect all points of each state's policies, procedures, and protocols, 14 complete drills were held, of one and three days' duration, first between Missouri and Iowa, then seven other states, involving both the laboratory staff and supervisors, as well as professionals in information technology, state emergency preparedness, and short-term reporting and follow-up. A total of 5,864 dried bloodspot samples were involved and 351,840 laboratory results on identical specimen cards were compared. Detailed minutes of meetings were examined for issues anticipated and resolved, and logs of the actual drills revealed additional lessons learned. The quantitative laboratory results and clinical decisions between laboratories were compared with usual statistics. The 14 exchanges among nine states included 18 days of newborn screening. A 1996 Federal Law, the Emergency Management Assistance Compact (EMAC), successfully resolved legal, states' rights, and sovereignty issues and permitted the demonstration of emergency preparedness. Respective staffs achieved sufficient harmonization of computer and laboratory systems to assure comparability of results. Interstate laboratory results were highly correlated with only minimal differences in a few borderline abnormal results. No new positive cases were uncovered that would have required clinical intervention. The laboratories of nine participating states produced findings that led to exactly the same rate of positive results that would require diagnostic evaluation of the newborn, documenting the great reliability the laboratories have achieved. Emergency preparedness that includes fully functional drills should be a routine element of all states' newborn screening programs. They are feasible and economical. Our regional approach has been a successful model. HRSA Grant U22MC03962.

1876F

Evaluation of the Bogotá Congenital Malformations Surveillance Program. *I. Zarante[1,2], C. Mallarino[1], G. Gracia[2].* 1) Inst de Gen Humana, Pontificia Univ Javeriana, Bogota, Bogota, Colombia; 2) Secretaria de Salud de Bogota.

Introduction: The Bogotá Congenital Malformations Surveillance Program was designed in the year 2001, based on the Latin American collaborative study of congenital malformations (ECLAMC). The program is hospital-based and currently covers almost 90% of the total births of the city. CDC guidelines for evaluating public health surveillance systems clearly state that these programs should be evaluated. Methods: We evaluated the attributes of the Bogotá Congenital Malformations Surveillance System based on the CDC guidelines for evaluating public health surveillance systems. Results: 1. Simplicity: the general structure of the program is relatively simple, compared to the majority of surveillance systems functioning in Colombia. Additionally, only 14 people are needed for the program to work properly, and it functions with a single centralized system for capturing, managing and analyzing data, which also adds simplicity. 2. Flexibility: In retrospective, the program has been able to adapt to changes without great inconvenient. 3. Data quality: The mean percentage of blank responses to items on the surveillance forms was 5% for the case-control modality and 17% for the monitor modality in the year 2011. 4. Sensitivity: Compared to the Metropolitan Atlanta Congenital Defects Program, our program has a 48% sensitivity for the detection of all defects, the lowest being congenital cataract and interatrial communication, and the highest being Down syndrome. 5. Predictive value positive: in the year 2011 we followed a total of 513 patients with 510 having a confirmed diagnosis, which gives us a predictive value positive of 99.4%. 6. Representativeness: The program covers almost 90% of the births that occur in the city of Bogotá and the population is heterogeneous, which means that it is fairly representative of the general population of newborns in the city. 7. Timeliness: Information gathered during one month is entered into the central electronic server during the first ten days of the following month for the case-control modality and during the first 20 days for the monitor modality. Follow-up is initiated at two months of age. 8. Stability: Before the year 2010, the program had some periods of instability because resources were very limited making it very difficult to hire essential personnel, which resulted on inactivity of the surveillance in some hospitals.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1877F

Drivers, barriers and opportunities for genetic testing services in emerging economies: the GenTEE (Genetic Testing in Emerging Economies) project. *I. Nippert[1], A. Christianson[2], D.D.G. Horovitz[3], R. Raouf[4], C.D. Padilla[5], V. Penchaszadeh[6], A. Rajab[7], I.C. Verma[8], N. Zhong[9], L. Gribaldo[10], U. Kristoffersson[11], J. Schmidtke[12].* 1) Women's Health Research, Muenster Medical School, Westfaelische Wilhelms-Universitaet, Muenster, Germany; 2) Division of Human Genetics, National Health Laboratory Service & University of the Witwatersrand, South Africa; 3) Centro deGenética Médica - Instituto Nacional de Saúde da Mulher, da Criança e do Adolescente Fernandes Figueira (IFF/Fiocruz), Rio deJaneiro, Brazil; 4) Institute of Post Graduate Childhood Studies/Ain-Shams University, Cairo, Egypt; 5) Department of Pediatrics, College of Medicine and Philippine General Hospital; Institute of Human Genetics, National Institutes of Health, University of the Philippines Manila, Philippines; 6) Centre of Genetics and Public Health. Department of Health Sciences, Universidad Nacional de La Matanza, Buenos Aires, Argentina; 7) Genetic Unit, Directorate General of Health Affairs, Ministry of Health, Muscat, Sultanate of Oman; 8) Centre of Medical Genetics, Sir Ganga Ram Hospital, Rajender Nagar, New Delhi, India; 9) Peking University Center of Medical Genetics, Beijing, People's Republic of China; 10) Institute for Health and Consumer Protection, European Commission Joint Research Centre, Ispra, Italy; 11) Department of Clinical Genetics, University Hospital Lund, Lund, Sweden; 12) Institute of Human Genetics, Hannover Medical School, Hannover, Germany.

Background: Due to the epidemiological transition in the emerging economies of China, East Asia, India, Latin America, the Middle East and South Africa, these economies are facing (i) an increasing proportion of morbidity and mortality due to congenital and genetic conditions, (ii) a rising need for genetic services to improve patient outcomes and overall population health. These economies are facing the challenge how: (i) to ensure the successful translation of genetic/genomics laboratory and academic research into quality assured pathways, (ii) to develop a service delivery infrastructure that leads to equitable and affordable access to high quality genetic/genomic testing services. *Objectives:* (i) to document and compare current practices and the state of genetic service provision in eight emerging economies: Argentina, Brazil, China, Egypt, India, Oman, Philippines and South Africa, (ii) to identify current knowledge gaps and unmet service needs. The GenTEE international project is intended to inform policy decisions for the challenges of delivering equitable high quality genetic services and to promote international collaboration for capacity building. *Methods:* (i) a standardized survey that is the first of its worldwide that allows comparison of services internationally across a number of key dimensions by using a core set of indicators, selected by the GenTEE consortium for their relevance and comparability, (ii) capacity building demonstration projects. To date, the GenTEE project has completed its survey that maps the current state of genetic services in the participating countries and identifies current drivers, barriers and opportunities for genetic services development. *Results:* There is no equitable access to genetic services in all countries mainly due to financial barriers (underfunded fragmented public services, out-of-pocket expenses tend to be the norm for genetic testing services), geographical barriers (concentration of services in main cities) and skill gaps, resulting in inequitable services or delayed access. The development of services in the private sector is opportunistic and mostly technology and market driven. There is a marked lack of standard operating procedures and agreed quality assessment processes for new technologies. *Discussion:* International collaborative networks can provide support for capacity building and help to strengthen the provision of quality genetic/genomic services in emerging economies.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1878W

Genetic variation in susceptibility to hypothyroidism induced hearing impairment. *Q. Fang[1], T.J. Jones[1], T.-W. Gong[1], A.H. Mortensen[1], M.T. Fleming[1], D.F. Dolan[1], M.I. Lomax[1], K.R. Johnson[2], M. Mustapha[3], S.A. Camper[1].* 1) University of Michigan, Ann Arbor, MI; 2) The Jackson Laboratory, Bar Harbor, ME; 3) Stanford University, Stanford, CA.

Hypothyroidism during pregnancy or in newborns can impair hearing and cognition, although individual children vary in their sensitivity. Our studies in genetically hypothyroid mice demonstrated that the genetic background of the fetus confers susceptibility or resistance to hypothyroidism-induced hearing impairment (1, 2). Thyroid hormone deficiency has pleiotropic effects on development of neurosensory and conductive aspects of hearing in susceptible mouse strains, including effects on innervation and potassium flux in the cochlea and maturation of the middle ear (2, 3). The DW/J-*Pou1f1*$^{dw/dw}$ mutants produce little or no thyroid hormone and have severe dwarfism and profound, congenital deafness. Mutant F2 progeny from a cross of this strain with CAST/Ei had equal growth insufficiency but exhibited a range of hearing abilities, from normal hearing to profound deafness. Initial microsatellite mapping studies suggested that a locus on chromosome 2 could explain about 20% of the variation in hearing, and we named it Modifier of *dw* hearing, *Mdwh* (4). Both AKR and C57BL/6 are susceptible, and C3H/HeJ and 129P2/Ola are resistant. The strain distribution pattern analysis of the DW/J stock revealed that it is similar to C57BL/6 in the *Mdwh* region. Here we report the results of genome-wide SNP typing for 1500 loci on DW/J × CAST/Ei F2 *Pou1f1*$^{dw/dw}$ mice and DW/J × 129P2/Ola F2 mutants whose hearing was quantified by auditory brain stem response testing. The data from both crosses implicate the same region of chromosome 2 and refine the *Mdwh* to 20 Mb with over 500 genes. We carried out gene expression profiling with cochlear RNA from DW/J-*Pou1f1*$^{dw/dw}$ and wild type mice and identified over 500 genes that are regulated by thyroid hormone in the cochlea, and 10 map to the *Mdwh* region. We are comparing expression of these genes in susceptible and resistant strains to evaluate them as candidates for *Mdwh*. Identification of a gene that can protect against the pleiotropic effects of thyroid hormone deficiency on the development of hearing could lay the foundation for designing therapeutics for children who are not responsive to thyroid hormone supplementation. Funded by Endocrine Society Summer Fellowship (TJJ), March of Dimes (SAC) and NIH (DFD and MM). References: 1. Karolyi et al. Mamm Genome 19:596, 2007 2. Fang et al. J Assoc Res Otolaryngol 13:173, 2012 3. Mustapha et al. J Neurosci 28:1212, 2009 4. Fang et al., Genetics 189:665, 2011.

1879T

Transethnic Mapping of Putative Diabetic Nephropathy Loci in African Americans. *N.D. Palmer[1,2,3], M.C.Y. Ng[2,3], J.N. Cooke[5], M. Petrulis[1], B.I. Freedman[4], D.W. Bowden[1,2,3].* 1) Biochemistry Department, Wake Forest School of Medicine, Winston-Salem, NC; 2) Center for Diabetes Research, Wake Forest School of Medicine, Winston-Salem, NC; 3) Center for Genomics and Personalized Medicine Research, Wake Forest School of Medicine, Winston-Salem, NC; 4) Molecular Medicine and Translational Sciences, Wake Forest School of Medicine, Winston-Salem, NC; 5) Internal Medicine, Section on Nephrology, Wake Forest School of Medicine, Winston-Salem, NC.

Diabetes-associated nephropathy (DN) is the most common cause of end-stage renal disease (ESRD) in the United States, with disproportionately higher rates in African Americans. The purpose of this study was to perform fine mapping of 23 reported DN loci in the African American population. Our African American Genome-Wide Association Study (GWAS) dataset was composed of 965 DN cases and 1029 population-based controls. In total, 3,122,014 SNPs were evaluated including those from the Affymetrix 6.0 array with imputation to HapMap Phase II. All analyses were performed with adjustment for admixture using a principle components approach and adjusted for multiple comparisons by accounting for the effective number of independent SNPs (n_{eff}) at each locus. Examination of 23 reported DN genes revealed three loci statistically associated (P<0.05) with DN in African Americans after adjustment for multiple comparisons: myosin heavy chain 9 non muscle (*MYH9*; n_{eff}=56, rs5750250, OR=0.72, P=1.6E-05), carnosine dipeptidase 1 (*CNDP1*; n_{eff}=45, rs65666815, OR=1.31, P=7.6E-03 and chimerin 2 (*CHN2*; n_{eff}=194, rs3729621, OR=1.57, P=0.037). Given the high frequency and strong effect of variants in apolipoprotein L1 (*APOL1*) on DN-susceptibility, G1 and G2 haplotypes were included as covariates in the analysis. As a result, association was eliminated at *MYH9* while the significance at *CNDP1* (rs65666815, P=0.011) was attenuated and a new variant in *CHN2* rose to significance (rs17157914, OR=0.65, P=0.025). These results suggest that causal genes may be shared across populations. The reduced, differential linkage disequilibrium pattern observed in African Americans will facilitate fine mapping of causal variants associated with DN.

1880F

The Exon 3 deleted/full-lenght polymorphism of the Growth Hormone Receptor in the Brazilian Population: Association with idiopathic short stature and body composition measurements in osteoporosis. *F.A. Marques[1], T.C.L. Lins[2], C.T. Neves[3], M.T.O. Cardoso[3], R.M. Lima[4], R.J. Oliveira[4], R.W. Pereira[1,2], R. Pogue[1].* 1) Programa de Pós-Graduação em Ciências Genômicas e Biotecnologia, Universidade Católica de Brasília, Brasília, DF, Brazil; 2) Programa de Pós-Graduação em Patologia Molecular, Universidade de Brasília, Brasília, DF, Brazil; 3) Núcleo de Genética da Secretária de Saúde do Distrito Federal, DF, Brazil; 4) Faculdade de Educação Física, Universidade de Brasília, Brasília, DF, Brazil.

The influence of the growth hormone receptor (GHR) exon 3 full-length/deleted (fl/d3) polymorphism on the response to growth hormone is undetermined, but has been suggested to be associated with linear growth, and also with phenotypes such as obesity and diabetes. The work described here involves genotyping of this polymorphism in different Brazilian population groups. Instead of direct genotyping, the use of the single nucleotide polymorphism rs6873545 as anindirect/surrogate marker was validated. Genotyping by two methods (multiplex PCR/direct and TaqMan-SNP/indirect) of 105 admixed individuals from the general Brazilian population showed 100% concordance between the GHR polymorphism and the SNP alleles. The distribution of genotypes was in Hardy-Weinberg equilibrium (HWE) and no statistical difference was observed according to genetic ancestry among genotypes. Subsequent SNP genotyping of 32 patients with idiopathic short stature showed a significant association (p=0.013) between the full-length (fl) allele and short stature, and the polymorphism was not in HWE in this group. Furthermore, genotyping in 175 women between the ages of 56 and 82 with and without osteoporosis showed that while there was no association with the development of osteoporosis, there was an association between the fl/fl genotype and lower bone mineral density (p=0.0004), and between this genotype and lower lean mass (p=0.018) in women with the condition. These results suggest that: a) the SNP rs6873545 can be used as a surrogate for the GHR fl/d3 polymorphism due to linkage disequilibrium in the Brazilian population; b) the polymorphism may play a role in the multifactorial determination of normal and pathological height; and c) the interference of the fl allele on the response to growth hormone may affect the severity of osteoporosis in post-menopausal women.

1881W

Variants in PTPN22 and SMOC2 genes and the risk of thyroid disease in the Jordanian population. *A. Alkhateeb[1], N. Marzouka[1], R. Tashtosh[2].* 1) Biotechnology and Genetics Department, Jordan Univ Science & Tech, Irbid, Jordan; 2) Endocrinology Department, Ibn Alnafees hospital, Irbid, Jordan.

Autoimmune thyroid diseases (AITDs) are one of the most prevalent autoimmune diseases, especially among females. AITDs contain two main categories: Graves Disease (GD) and Hashimoto Thyroiditis (HT). The causes of AITDs are not clear, though twins and familial studies give many evidences on a genetic susceptibility. Our study focused on two polymorphisms: rs2476601 (C1858T) polymorphism in PTPN22 gene which is an autoimmune risk locus and rs13208776 polymorphism in intron 4 in SMOC2 gene which was found to be associated with generalized autoimmune vitiligo. In our case-control study, we collected blood samples from 216 normal Jordanians and 204 Jordanian patients with clinical symptoms of autoimmune thyroid diseases, these samples were used for DNA extraction. Genotyping of polymorphisms in PTPN22 and SMOC2 genes were done by Polymerase chain reaction (PCR) followed by Restriction Fragment Length Polymorphism (RFLP) technique. We statistically analyzed the genotyping data by Open-Epi software to calculate P-value, our results show that there is no difference between patients and controls genotypes and allele frequencies on C1858T locus (P=1), thus no association was found between this locus and AITDs as a whole or GD and HT separately. Rs13208776 variant in SMOC2 gene showed more variant distribution between controls and patients genotypes (P=0.17) and allele (P=0.12) frequencies, especially in case of females subgroup when the genotypes statistically tested on combined dominant model (P=0.052). Even though P-values were above 0.05 we suggest that SMOC2 gene may play role in AITD susceptibility as a dominant variant. Additional samples might be needed to confirm or exclude association of SMOC2 and AITD.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1882T

Genetic associations with neonatal thyroid stimulating hormone levels. *F. Alul[1], O. Shchelochkov[1], S. Berberich[2], J. Murray[1], K. Ryckman[1].* 1) Department of Pediatrics, University of Iowa, Iowa City, IA; 2) State Hygienic Laboratory, University of Iowa, Iowa City, IA.

Background: Thyroid hormones are critical regulators for protein synthesis, carbohydrate metabolism and cellular differentiation. Elevations or deficits in thyroid hormone levels are responsible for a wide range of neonatal and adult phenotypes. Even small differences in thyroid function are correlated with body mass index, blood pressure, lipid levels, atrial fibrillation and cardiovascular disease. Newborns are commonly screened for biochemical evidence associated with hypothyroidism which, if untreated, can result in severe cognitive disability. Several genome-wide, candidate gene and meta-analysis studies have been performed with thyroid hormones in adults; however, to our knowledge no studies have examined genetic associations with neonatal thyroid levels. Objective: To identify genetic factors contributing to thyroid stimulating hormone (TSH) variation in newborns. Design: A population of Iowa neonates; term (n=827) and preterm (n=815), were genotyped at 45 single nucleotide polymorphisms encompassing 24 candidate genes. TSH measurements were obtained from the Iowa Neonatal Metabolic Screening Program (INMSP). Analysis of variance (ANOVA) was performed to identify genetic associations with TSH concentrations. Results: The strongest association was rs4704397 in the *PDE8B* gene (p=1.3×10-4), followed by rs965513 (p=6.4×10-4) on chromosome 9 upstream of the *FOXE1* gene. Both of these SNPs met statistical significance after correction for multiple testing; Bonferonni correction significance threshold set at p<0.001 (0.05/45 markers). Six other SNPs show marginal significance (p<0.05) after correction for multiple testing. Conclusions: Several genetic polymorphisms previously associated with adult TSH levels are also associated with neonatal levels. This may be useful in the early prediction of risk for adult diseases and conditions associated with thyroid hormone levels. Furthermore, this provides a better understanding of the thyroid profile and potential risk for thyroid disorders in newborns.

1883F

TNF-β Nco1 polymorphism in relation to postoperative sepsis outcome in joint care surgery. *K. Baghel, R. Srivastava, S. Raj, D. Sanghi.* Orthopaedic Surgery, KG Medical College, CSM Medical University, Lucknow, UP, India.

Background: Tumor Necrosis Factor (TNF) is believed to be a cytokine central to pathogenesis of sepsis and the TNF-β Nco1 polymorphism has been found to be associated with increased mortality rate in severe sepsis following major surgeries. We postulated that Nco1 polymorphism may be associated with post operative infection in some patients but not in others Methods: The study group consisted of 153 patients undergoing major elective surgery (surgical time more than 1 hour and requiring general anesthesia and or respiratory assistance). Blood samples were obtained for genomic DNA isolation. Genotyping of each patient for TNF-β polymorphism was performed. All the patients were followed for 1 month following surgery for any evidence of sepsis as determined by guidelines from Bone et al. Subjects with and without post-operative infection were compared on presence or absence of polymorphism and of confounder status. Results: The overall allele frequency for TNF-β genotype was 0.32 for TNFB1 and 0.68 for TNFB2. In TNF-β genotype, homozygous recessive TNFB1 were 17 (11.1%), heterozygous TNFB1/ TNFB2 were 63 (41.2%) and homozygous dominant TNFB2 were 73 (47.7%). 125 patients showed an uncomplicated postoperative recovery, while 25 developed mild sepsis and 3 developed severe sepsis. Genotype distribution in patients with an uncomplicated clinical course was significantly different from that in patients with postoperative sepsis. Development of postoperative sepsis was significantly higher in patients homozygous for the allele TNFB2. When compared with patients carrying at least one TNFB1 allele (TNFB1 homozygous and heterozygous genotype), the TNFB2 homozygous genotype was associated with an OR of 3.39 (p=0.005) for the development of severe sepsis. Compared with the heterozygous genotype, the OR for the homozygous TNFB2 genotype was 5.5 (p=0.001) Conclusions: The Nco1 polymorphism within the TNF-β gene influences the development of postoperative sepsis suggesting a genetic determination in susceptibility to postoperative infection in surgerical interventions.

1884W

Preeclampsia is associated with variation in endoglin pathway genes. *M. J. Bell[1], J. M. Roberts[1, 2], S. A. Founds[1, 2], A. Jeyabalan[1,2], L. Terhorst[1], Y. P. Conley[1].* 1) University of Pittsburgh, Pittsburgh, PA; 2) Magee-Womens Research Institute and Foundation, Pittsburgh, PA.

Preeclampsia is a pregnancy disorder that greatly impacts maternal and fetal/neonatal health and wellbeing. Preeclampsia's pathophysiology remains unclear, but gene expression studies and the presence of increased circulating soluble endoglin before clinically evident disease have implicated endoglin as a potential factor involved in its development. The purpose of this case-control candidate gene association study was to investigate the association between endoglin pathway genetic variation and preeclampsia. Data collected from 355 white (181 cases/174 controls) and 60 black (30 cases/30 controls) women matched on ancestry, age, and parity were analyzed. Tagging single nucleotide polymorphisms (tSNPs) in ENG, TGFβ1, TGFβR1, ALK1, and TGFβR2 were evaluated with iPLEX® and TaqMan® technologies. Allele/genotype/haplotype tests were conducted separately in white/black subgroups with a $\chi 2$ or Fisher's exact test. Odds ratios were computed with binary logistic regression for tSNPs with significant genotype tests. In white women, variation in ENG (rs11792480, rs10121110) and TGFβR2 (rs6550005) was associated with preeclampsia. Allelic frequency distributions in rs11792480, rs10121110, and rs6550005 were significantly different among cases and controls while genotype distributions of rs10121110 and rs6550005 were further associated with preeclampsia (*p*-values < .05). For rs10121110, women with the AA genotype were 2.290 times more likely to develop preeclampsia compared to the GG genotype (99% CI [1.022, 5.133], p = .008). The ENG haplotype TACGA, which contains the risk alleles for rs11792480 and rs10121110, was also over-represented in cases (p = .022). In black women, variation in TGFβ1 (rs4803455, rs4803457), TGFβR1 (rs10739778), and TGFβR2 (rs6550005, rs1346907, rs877572) was associated with preeclampsia. Allelic frequency distributions in rs10739778, rs6550005, rs1346907, and rs877572 were significantly different among cases and controls while genotype distributions of rs10739778, rs4803455, and rs4803457 were further associated with preeclampsia (*p*-values < .05). For rs4803457, women with the CT genotype were 7.437 more times likely to develop preeclampsia compared to the CC genotype (99% CI [1.192, 46.408], p = .005). These results demonstrate that variation in ENG pathway genes is associated with preeclampsia in white and black women, with different genes from the same pathway being involved in white women compared to black women.

1885T

Genetic Polymorphisms in ESR1 and ESR2 Genes and Risk of Hypospadias in a Multi-ethnic Study Population. *S. Choudhry[1], E.J. Lammer[2], J.S. Witte[3], S. Dasgupta[1], C. Ma[4], G.M. Shaw[4], L.S. Baskin[1], S.L. Carmichael[4].* 1) Dept Urology, Univ California, San Francisco, CA; 2) Children's Hospital Oakland Research Institute, Oakland, CA; 3) Dept of Epidemiology & Biostatistics, Institute of Human Genetics, Univ California, San Francisco, CA; 4) Dept of Pediatrics, Division of Neonatology, Stanford University, Stanford, CA.

Hypospadias is a common congenital malformation of the male external genitalia with a multifactorial etiology. Little is known about the genes involved in hypospadias. Estrogenic endocrine disruptors acting via estrogen receptors α (ESR1) and β (ESR2) have been implicated in the etiology of hypospadias. Our objective was to examine genetic polymorphisms in ESR1 and ESR2 genes and their association with hypospadias in a racially-ethnically diverse study population of California births. We investigated the relationship between 108 ESR1 and 37 ESR2 SNPs and hypospadias in 647 cases and 877 population-based non-malformed controls among infants born in selected California counties from 1990–2003. For risk of hypospadias, odds ratios (OR) for 16 of the 108 ESR1 SNPs and 2 of the 37 ESR2 SNPs had p-values <0.05 (p-value range: 0.05–0.002). Due to high LD across the SNPs, haplotype analyses were conducted, and 18 and 5 two-SNP or larger haplotype blocks were examined for ESR1 and ESR2 respectively. These analyses identified two haplotype blocks in ESR1 that were associated with increased risk of hypospadias (OR 2.67, p<0.002 and OR 2.5, p<0.006). No significant haplotype association was observed for ESR2. These findings suggest that SNPs in ESR1 may influence the risk of hypospadias.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1886F

A missense mutation in Exon 2 of Adiponectin receptor 2 (ADIPOR2) is associated with serum insulin in overweight and obese African-American subjects. *A. Doumatey, G. Chen, J. Zhou, H. Huang, A. Adeyemo, C. Rotimi.* NHGRI/CRGGH, National Institutes of Health, Bethesda, MD.

Introduction: Obesity (Ob) and overweight (OW) are associated with a number of metabolic disorders including insulin resistance, type 2 diabetes, hypertension, and dyslipidemia. Insulin is one of the most perturbed biochemical markers in OW and Ob, and it has been shown that adiponectin has an insulin-sensitizing action. Thus, we hypothesized that variants in genes encoding adiponectin and/or adiponectin receptors may influence insulin levels and action in Ob and OW persons. Methods: The subjects comprised 490 Ob and OW (BMI≥ 25 kg/m2) African Americans from the Howard University Family study. Genotypes for SNPs in ADIPOQ, ADIPOR1 and ADIPOR2 were obtained from the Affymetrix genome- wide Human SNP array 6.0. Imputation was done using MACH with HapMap CEU and YRI as the reference population panels. A total of 169 genotyped and imputed SNPs ±50 kb around adiponectin (ADIPOQ) and adiponectin receptors (ADIPOR1, ADIPOR2) were tested for association with fasting serum insulin levels. Association models assumed an additive genetic model with adjustment for age, sex and the first two principal components of the genotypes. Results: Nine variants in AdipoR2 gene including a missense mutation (rs12298275) in exon 2 were associated with fasting insulin levels. In addition, one variant in CYB5R1 gene and four others in CACNA2D4 gene were associated with fasting insulin (p-values ranging from 4.1×10-2 to 5.4×10-3). CYB5R1 and CACNA2D4 are the closest genes to ADIPOR1 and ADIPOR2 respectively. Conclusion: These findings provide evidence for the role of adiponectin receptors in insulin metabolism in OW and Ob African Americans.

1887W

Common variants associated with normal tension glaucoma and optic nerve degeneration are also associated with glaucoma in exfoliation syndrome. *B. Fan[1], S. Loomis[1], J.H. Kang[2], D.Y. Wang[1], B. Yaspan[3], M.A. Hauser[4], L.R. Pasquale[1], J.L. Haines[3], J.L. Wiggs[1].* 1) Dept of Ophthalmology, Harvard Med Sch, Massachusetts Eye & Ear Infirmary, Boston, MA; 2) Dept of Medicine, Channing Lab, Brigham and Women's Hospital, Boston, MA; 3) Center for Human Genetics Research, Vanderbilt University, Nashville, TN; 4) Dept of Ophthalmology & Medicine, Duke Univ Medical Center, Durham, NC.

The *CDKN2BAS* region at 9p21and the *SIX1/SIX6* region at 14q22 has been previously associated with primary open angle glaucoma (POAG), a genetically and phenotypically complex trait that is a leading cause of blindness worldwide. Recently we showed that the *CDKN2BAS* region and a novel regulatory region on 8q22 are associated with normal-tension glaucoma (NTG), suggesting that these SNPs primarily influence the development of optic nerve disease in glaucoma. In this study we evaluated the association of SNPs located in the 9p21, 14q22 and 8q22 genomic regions with another common form of glaucoma, exfoliation glaucoma (EG), to determine if these SNPs also contribute to optic nerve disease in other types of glaucoma. The lead SNP from each region was analyzed in 104 exfoliation glaucoma patients and 344 controls. Single-SNP associations were initially analyzed using the chi-square test and further evaluated using logistic regression models adjusted for age. The lead SNP in *CDKN2BAS* was significantly associated with EG (rs2157719, p=0.004, OR=0.59, 95%CI: 0.41-0.84), while the 8q22 SNP was nominally associated with EG (rs284489, p=0.02; OR=0.68, 95%CI: 0.48-0.95) and the *SIX1/SIX6* SNP did not show evidence of association (rs10483727, p=0.14, OR=1.28, 95%CI: 0.92–1.77). These results suggest that common SNPs at 9p21 and 8q22 contribute to optic nerve disease in exfoliation glaucoma, and thus may be risk factors for optic nerve disease in other common forms of glaucoma. **Grant support:** NEI Grants R01EY020928, R01EY015872 and P30EY014104-06, Research to Prevent Blindness and The Massachusetts Lions Eye Research Fund.

1888T

Genetic variants in Selenoprotein P plasma 1 gene (*SEPP1*) are associated with fasting insulin and first phase insulin response in Hispanics. *J.N. Hellwege[1,2,3], N.D. Palmer[2,3,4], J.T. Ziegler[5], C.D. Langefeld[5], C. Lorenzo[6], J.M. Norris[7], T. Takamura[8], D.W. Bowden[2,3,4,9].* 1) Molecular Genetics & Genomics, Wake Forest School of Medicine, Winston Salem, NC; 2) Center for Genomics and Personalized Medicine Research, Wake Forest School of Medicine, Winston-Salem, NC; 3) Center for Diabetes Research, Wake Forest School of Medicine, Winston-Salem, NC; 4) Department of Biochemistry, Wake Forest School of Medicine, Winston-Salem, NC; 5) Division of Public Health Sciences, Wake Forest School of Medicine, Winston-Salem, NC; 6) Department of Medicine, University of Texas Health Science Center, San Antonio, TX; 7) Department of Epidemiology, Colorado School of Public Health, University of Colorado Denver, Denver, CO; 8) Department of Disease Control and Homeostasis, Kanazawa University Graduate-School of Medical Science, Kanazawa, Ishikawa, Japan; 9) Internal Medicine-Endocrinology, Wake Forest School of Medicine, Winston-Salem, NC.

Type 2 diabetes (T2D) is a major health problem worldwide, with over 25 million people in the US affected. T2D is characterized by insulin resistance, as well as the body's inability to produce sufficient amounts of insulin to maintain homeostasis. Insulin resistance is not yet fully explained on a molecular level, though several genes and proteins have been tied to this defect. Knockdowns of the *SEPP1* gene, which encodes the Selenoprotein P (SeP) protein, increase insulin sensitivity in mice. SeP is a liver-derived plasma protein and a major supplier of selenium. Selenium is a proposed insulin mimic and antidiabetic agent. Due to a suggestive role in insulin resistance, 21 *SEPP1* polymorphisms were selected for analysis with glucometabolic measures in 1424 Hispanics from families in the Insulin Resistance Atherosclerosis Family Study (IRAS-FS). A frequently sampled intravenous glucose tolerance test was used to obtain precise measures of first phase insulin release (acute insulin response, AIR) and insulin sensitivity (S_I). Tagging SNPs (12) from HapMap (including 2 SNPs previously associated with SeP levels), 4 coding variants from the ESP_GO database and 5 SNPs in the promoter region were genotyped and analyzed for association using SOLAR. Two highly correlated (r^2=0.88) coding SNPs showed significant association with AIR (rs28919926; Cys368Arg; p=0.0028 and rs146125471; Ile293Met; p=0.0026) while rs16872779 (intronic) was associated with fasting insulin levels (p=0.0096). Replication genotyping was performed in the multi-ethnic non-family based IRAS cohort, which included 480 Caucasian, 382 Hispanic and 288 African American samples. Meta-analysis of the IRAS cohorts identified two markers (rs6882786, promoter and rs7579, 3' UTR) associated with S_I (p=0.0079 and 0.010, respectively), while rs6882786 and rs16827262 (3' UTR) were found to be associated with AIR. In the IRAS Hispanic cohort, the associations seen in the IRAS-FS were not replicated (p≥0.05) but meta-analysis of IRAS-FS and all 3 IRAS cohorts (N= 2446) supported association of rs7579 with S_I (p=0.047) and rs28919926 with AIR (p=0.013). Overall, these results are consistent with the literature suggesting an important role for *SEPP1* in insulin resistance and the development of T2D.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1889F

Replication of GWAS meta-analysis associations in a large Amish pedigree support *VWF*, *ABO*, *STXBP5*, *STAB2*, and *SCARA5* as modifiers of VWF and FVIII levels. *J. Hinckley[1], K. Wang[2], T. Burns[2], L. Law[1], A. Shapiro[3], J. Di Paola[1].* 1) Univ Colorado, Denver, Aurora, CO; 2) Univ Iowa, Iowa City, IA; 3) Indiana Hemophilia and Thrombosis Center, Indianapolis, IN.

Closed populations have proven valuable in elucidating genes of Mendelian traits and are becoming more recognized in the study of complex and quantitative traits due in part to a more homogenous environmental component and a similar distribution of common variants compared to outbred populations. We characterized a 2171-member Amish family with von Willebrand disease caused by a single autosomal-dominant C4120T mutation and demonstrated the variability between individuals and the reproducibility within individuals of coagulation factor VIII (FVIII:C) and von Willebrand factor (VWF) levels. Analysis of VWF levels and FVIII:C in SOLAR (covariates: age, sex, C4120T) estimated a heritability (h^2) of 0.301 for VWF antigen levels (VWF:Ag; $p=2.7\times10^{-6}$), 0.386 for VWF ristocetin cofactor activity (VWF:RCo, functional assay; $p=1.16\times10^{-8}$), and 0.280 for FVIII:C ($p=1.52\times10^{-5}$). The recent CHARGE Consortium meta-analysis reported 8 chromosomal regions associated with VWF levels, 5 of which are also associated with FVIII levels, and the most significant intragenic SNP for each region. We genotyped 10 of 13 SNPs covering all 8 genes in 448 members of our Amish cohort and replicated significant associations for 6 SNPs (SOLAR, covariates: age, C4120T): rs1063857 (VWF; VWF:Ag, p= 0.01), rs1063857 (ABO; FVIII:C, $p=8.28\times10^{-6}$; VWF:Ag, $p=1.34\times10^{-5}$), rs9390459 (STXBP5; VWF:Ag, p=0.03), rs9644133 (SCARA5; FVIII:C, p= 0.01), and rs12229292 (STAB2; FVIII:C, p=0.021). Of 94 HapMap tagging SNPs genotyped through VWF, rs7964554 (intron 4), demonstrated the most significant association with VWF:Ag and VWF:RCo ($p=6.46\times10^{-4}$ and 4.17×10^{-7}, respectively; SOLAR, covariates: age, C4120T, ABO). SNPs near the functional domains of VWF were associated with VWF:RCo but not VWF:Ag. ABO, the best-characterized modifier of VWF levels, explained an additional 8, 12, and 7% of the variability of VWF:Ag, VWF:RCo, and FVIII:C, respectively, in the Amish cohort (covariates: age, sex, C4120T). Inclusion of VWF:Ag as a covariate of FVIII:C suggests the effect of ABO on FVIII:C is mediated through VWF. Exonic sequencing of STXBP5 and STAB2 in 48 samples identified only common variants in dbSNP. In summary, we report replication of 6 CHARGE study associations and evaluation of VWF, ABO, STXBP5, and STAB2 as candidate modifier genes of FVIII and VWF levels, demonstrating the usefulness of closed populations in the study of complex traits.

1890W

Mutation Analysis in Zinc Finger Protein Isoform 1 (*ZNF644*) in Cases with Moderate and High Myopia. *N. Kaur, S. Saini, V. Vanita.* Department of Human Genetics, Guru Nanak Dev University, Amritsar, Punjab, India.

Purpose: To perform mutation screening in *ZNF644* in patients with moderate and high myopia from North India. Materials and Methods: The present study included 58 families having patients with moderate and high myopia. All the cases were recruited from Dr. Daljit Singh Eye Hospital, Amritsar. All these patients went through complete ophthalmologic examination that included visual acuity, refractive error, corneal thickness and axial length. Of these 40 patients were high myopic (-6.00 D and above) and 18 were with moderate myopia (-2.00 D to -5.00 D). The mutation screening in *ZNF644* was performed by bi-directional sequence analysis of the amplified products. Results: We observed few previously unreported polymorphisms and substitutions in the coding and 3' UTR region of *ZNF644* in the analyzed patients. Also a polymorphism i.e. Arg455Arg was observed in one of the analyzed patient. Conclusions: We report identification of few previously unreported nucleotide substitutions in *ZNF644* in the analyzed moderate and high myopia cases in present study.

1891T

Title: To study the association of Peroxisome Proliferative Activated Receptor Gamma (*PPARG*) gene polymorphism (p.Pro12Ala) with diabetic retinopathy. *R. Kaur, S. Goyal, V. Vanita.* Department of Human Genetics, Guru Nanak Dev University, Amritsar, Punjab, India.

Purpose: To study the association of Peroxisome Proliferative Activated Receptor Gamma (*PPARG*) gene polymorphism (p.Pro12Ala) with diabetic retinopathy (DR) in type 2 diabetic patients from North India. Material and Methods: The present case control association study included 148 type 2 diabetic patients. All the subjects were unrelated and were recruited from Dr. Daljit Singh Eye Hospital, Amritsar. Out of these, 65 patients were suffering from different types of retinopathies. Of these 49 were having Proliferative Diabetic Retinopathy (PDR), eleven had Background Diabetic Retinopathy (BDR) and five were with Non Proliferative Diabetic Retinopathy (NPDR). The 83 diabetic patients without any sign of retinopathy were also collected as controls. The genotyping was carried out by bi-directional sequencing analysis of the amplified products. The statistical analysis was done using SPSS (version 16.0) and one way ANOVA test was performed to find any association with clinical parameters (age, age of onset, duration of disease, basal metabolic index (BMI), systolic and diastolic blood pressure). Genotypic and allelic frequencies were calculated using statistical calculator (StatPac V. 3.0). Odds Ratio (OR) and 95% confidence intervals (CI) were determined by MedCalc (V 9.3.9.0). Results: The present study showed statistically significant differences between age of onset (p=0.009), duration of diabetes (0.005), and systolic blood pressure (0.001), when DR patients were compared with controls. Significant difference was found in allelic frequencies (p=0.0033) when DR patients were compared with controls and on comparison of different types of diabetic retinopathy (NPDR, BDR, PDR) patients among themselves as well as in comparison to controls. OR analyses revealed a significant relationship of (CG) heterozygote of *PPARG* with 1.8 times increased risk of diabetic retinopathy. Conclusion: The present finding suggest positive association of *PPARG* polymorphism (p.Pro12Ala) with DR. Present data also showed an increased risk of DR with CG heterozygote genotype. We conclude that *PPARG* polymorphism (p.Pro12Ala) seem to be a useful marker for diabetic retinopathy in the analyzed cases in present study.

1892F

Association of the IL17 gene polymorphisms with Lumbar Disc Disease. *A. Kelempisioti[1], I. Daavittila[1], N. Noponen[1], S. Barral[2], J. Ott[3], J. Niinimaki[4], T. Koskelainen[5], L. Ala-Koko[1,6], J. Karppinen[7,8], M. Männikkö[1,9].* 1) Department of Medical Biochemistry and Molecular B, University Of Oulu, Oulu, Finland; 2) Gertrude H. Sergievsky Center, College for Physicians and Surgeons, Columbia University, New York, USA; 3) Institute of Psychology, Chinese Academy of Sciences, Beijing, China; 4) Department of Diagnostic and Interventional Radiology, Oulu University Hospital, Oulu, Finland; 5) Department of Neurosurgery, Oulu University Hospital, Oulu, Finland; 6) Connective Tissue Gene Tests, Allentown, Pennsylvania, USA; 7) Institute of Clinical Medicine, Department of Physical and Rehabilitation Medicine, University of Oulu, Oulu, Finland; 8) Finnish Institute of Occupational Health, Oulu, Finland; 9) Institute of Health Sciences, Public Health and General Practice, University of Oulu, Oulu, Finland.

Lumbar disc disease (LDD) is a common condition influenced by genetic and environmental factors. LDD causes symptoms such as sciatica and low back pain with negative effect on patient's quality of life. Although various risk factors have been reported, the etiology and pathogenesis remain unclear. Genome wide analysis in a total of 186 members of 14 Finnish families with sciatica identified potential loci in 10 chromosomal regions with suggestive logarithm of odds (LOD)greater than 1. Fine mapping confirmed the susceptibility locus on chromosome 6 with a maximum two point LOD score of 2.36 suggesting evidence for linkage around marker D6S294. This marker is located near IL17A and IL17F genes. The exons, as well as the intron-exon boundaries of these two genes were screened in 30 symptomatic individuals using direct sequencing. Out of 17 single nucleotide polymorphisms (SNPs) (nine in IL17A, and eight in IL17F), we proceeded with three SNPs that showed most variation (rs7747909 in IL17A, rs763780 and rs11465553 in IL17F). Association studies of the candidate SNPs were performed in a total of 635 individuals from a Finnish case control set characterized by sciatica symptoms. The SNPs were genotyped using the SNaPshot multiplex method and were in Hardy Weinberg Equilibrium. Association analysis under an additive model of inheritance showed association ofSNP rs763780 [OR 1.61 (95percent CI 1.01–2.57)] but not with the other 2 SNPs. G-C-G haplotype was shown to be the risk haplotype [OR 1.75 (95percent CI 1.12–2.74)]. Both IL17A and IL17F are proinflammatory cytokines that induce expression of other cytokines and chemokines. Inflammatory etiology has been suggested to contribute to the generation of back pain. Our findings indicate that IL17 is involved in LDD.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1893W

Association and interaction of Polymorphisms in Angiogenic genes with Endometriosis. *M.M. Latha[1,2], Q. Hasan[3], V.L Kodati[1].* 1) Dept of Genetics and Molecular Medicine, Vasavi Medical and Research Centre, Hyderabad, India; 2) Nizam Institute of Medical Sciences,Hyderabad India; 3) Kamineni Hospital,, India.

Endometriosis is a multifactorial disorder where angiogenesis (neo vascularisation) plays an important role in the etiopathogenesis of the disease associated infertility and pelvic pain with a prevalence rate of 10–15% in women of reproductive age. It is know that the endometrium of women with endometriosis has an increased capacity to proliferate, implant and grow in the peritoneal cavity. Polymorphisms in functionally relevant angiogenic candidate genes are considered to be important for identifying individuals at high risk especially in case of polygenic disorders like endometriosis. Hence three single nucleotide polymorphisms in MMP-2 (rs243865), MMP-9 (rs3918242) and ICAM-1 (rs1799969) genes were evaluated in individual with endometriosis. The present study indicates significant difference in allele distributions of three SNPs between endometriosis patients and control women (P>.05;). Compared with the C allele, T allele frequency of MMP-2 is showing risk for endometriosis (OR=1.5556 95%CI=1.046–2-3133, P<0.05). Results of MMP-9 indicate that `TT' genotype is present with a higher frequency in women with endometriosis in our population (OR OR= 1.7147 95%CI=1.1029–2.6659 P<0.05). Results of ICAM-1 indicate a strong association of G allele association with the disease (OR=1.7983,95%CI 1.3101–2.4685,P<0.05). Clinical features like dysmenorrhea and menstrual irregularities were associated with MMP2 and MMP9 but no association with ICAM-1 genotypes. According to MDR analysis MMP-2, MMP-9 and ICAM-1 genes were individually contributing their role in causing the disease but they are not interactive with each other. Hence these three SNPs can be used for risk assessment of endometriosis in South Indian population. Interaction among MMP-2 (rs243865), MMP-9 (rs3918242) and ICAM-11799969) SNPs with the Risk of Endometriosis As shown in radial graph (fig-1) there was an additive influence of MMP-2 and MMP-9 polymorphisms in the risk of developing endometriosis. Where has ICAM1 and MMP9 are showing synergistic action and more significantly modify the risk of developing endometriosis and MMP2, ICAM1 are counteracting in the development of the disease and showing a less significant risk and less interaction compared with the other two genes in causing the endometriosis.

1894T

The HLA-DPB1 and DRB1 genes are synergistically associated with Graves disease in Han Chinese children. *Y. Lee[1,2,3], C. Huang[1], W. Ting[1], F. Lo[4], T. Chang[1], C. Chu[1], Y. Wu[1], S. Chang[1], W. Lin[1], M. Lin[1].* 1) Dept Pediatrics & Med Res, Mackay Memorial Hosp, New Taipei City, Taiwan; 2) 8Institute of Biomedical Sciences and Department of Medicine, Mackay Medical College, New Taipei City, Taiwan; 3) Department of Pediatrics, Taipei Medical University, Taipei, Taiwan; 4) 2Department of Pediatrics, Chang Gung Memorial Hospital; College of Medicine, Chang Gung University, Taoyuan, Taiwan.

Associations of HLA with Graves disease (GD) vary between races. Our study showed DRB1*09:01 is associated with and linked to GD. Recently the DPB1 gene was found to be independently associated with GD. We investigated which was the strongest factor associated with GD. **Subjects and Methods** The patients were 181 unrelated children with GD. Their age at the diagnosis was 10.5 ± 3.4 years. The controls consisted of 377 subjects. **Statistical analysis** Agreement with Hardy-Weinberg equilibrium, linkage disequilibrium (LD), and estimation of HLA DPB1-DRB1 haplotypes were performed on both patients and controls using PyPop 0.7.0. Differences in the frequencies of alleles, carriers, and haplotypes between patients and controls were evaluated using the Chi-square test. To evaluate the influence of linkage disequilibrium (LD) between HLA-DPB1 and DRB1 loci and find which of them conferring the stronger risk, we tested association with one locus in the presence or absence of the associated allele at the second locus. ORs and 95% CIs were calculated. The Bonferroni correction was used for multiple comparisons. A value of Pc < 0.05 was considered statistically significant. **Results** The distribution of genotypes of the HLA-DPB1 and -DRB1 genes of controls were in Hardy-Weinberg equilibrium. The frequency of the DPB1*05:01 allele was significantly higher in patients than in controls, OR (95%CI) = 1.61 (1.25–2.07), Pc = 0.0024. The HLA-DPB1-DRB1 haplotype 05:01-09:01 was significantly more frequent in patients than in controls, 75 (20.7%) vs. 75 (9.9%), OR = 2.37, Pc = 1.49E-05. In DRB1*09:01 carriers, DPB1*05:01 was significantly more frequent in patients than in controls, 88 (93.6%) vs. 86 (76.8%), OR = 4.43, Pc = 0.0054. But in those carrying no DRB1*09:01, the frequency of DPB1*05:01 was not significantly different between patients and controls. In DPB1*05:01 carriers, DRB1*09:01 was significantly more frequent in patients than in controls, 88 (54.3%) vs. 86 (30.9%), OR = 2.65, Pc = 7.84E-06. But its frequency was not significantly different between patients and controls carrying no DPB1*05:01. The frequency of carriers of both DPB1*05:01 and DRB1*09:01 was greater in patients than in controls, 88 (87.1%) vs. 86 (54.1%), OR = 5.75 (2.97–11.12), Pc = 3.07E-07. **Conclusion** HLA-DPB1*05:01 and DRB1*09:01 were synergistically associated with GD in Han Chinese children. Both of them must be simultaneously present to confer a risk of GD.

1895F

Fine-Mapping within Regions of Admixture Linkage to Sarcoidosis Risk in African Americans. *A.M. Levin[1], P. McKeigue[2], I. Datta[1], C.G. Montgomery[3], I. Adrianto[3], M. Colombo[2], M.C. Iannuzzi[4], B.A. Rybicki[1].* 1) Public Health Sciences, Henry Ford HIth System, Detroit, MI; 2) Public Health Sciences Section, University of Edinburgh Medical School, Edinburgh, Scotland; 3) Arthritis and Immunology Research Program, Oklahoma Medical Research Foundation, Oklahoma City, OK; 4) Department of Medicine, Upstate Medical University, Syracuse, NY.

Sarcoidosis is a multi-organ granulomatous disease of unknown etiology. Populations of primarily West African descent have higher sarcoidosis incidence than European populations, which compelled us to search for sarcoidosis genes using admixture mapping. In previously identified admixture linkage regions, we conducted fine mapping using genome-wide single nucleotide polymorphism (SNP) data genotyped on a sample of 2,727 self-identified African Americans and additional SNP imputation using the 1000 Genomes project data. Local ancestry was estimated genome-wide using the Local Ancestry in Admixed Populations method[1]. Employing a case-only technique we more finely mapped admixture linked regions using two approaches: 1) association testing adjusting for local ancestry and 2) the MIX score method[2], which tests the likelihood of whether a particular SNP association explains an admixture signal. We confirmed and more finely mapped sarcoidosis admixture linkage regions, two where increasing West African (6p24.3-p12.1 and 17p13.3-13) and two where increasing European (2p13.3-q12.1 and 6q23.3-25.2) ancestry were associated with susceptibility to sarcoidosis. The most significant admixture linkage locus was 6p24.3-p12.1 ($p=9*10^{-8}$), which encompasses the HLA class II region known to include sarcoidosis risk loci. Within this region, the most significant SNP after adjustment for local ancestry was rs74318745 (OR=0.69, 95%CI=0.62-0.78, $p=4.5*10^{-11}$), which resides within the HLA-DRA gene and explains the majority of the admixture signal at this locus (MIX $p=7.9*10^{-12}$). The 17p13.3-13.1 region ($p=2.0*10^{-4}$) was the most significant novel region of admixture linkage, and the association at SNP rs6502976 (OR=0.74 95%CI= 0.65–0.84, $p=1.2*10^{-4}$) within the X-linked inhibitor of apoptosis associated factor (XAF1) gene accounted for the majority of the admixture linkage (MIX $p=7.9*10^{-5}$). The SNP rs6502976 is in high linkage disequilibrium ($r^2=0.80$) with rs9891567, which has been identified as an expression quantitative trait locus for XAF1. As a negative inhibitor of apoptosis, XAF1 may play a role in granuloma formation and maintenance. In summary, our findings suggest that XAF1 is a novel sarcoidosis susceptibility locus in African Americans and demonstrate how admixture mapping can be used to prioritize regions of the genome for fine mapping susceptibility loci for sarcoidosis. **References:** [1]Baran et al. Bioinformatics (2012). [2]Pasaniuc et al. PLoS Genet (2011).

1896W

Fine-mapping of Central Adiposity Loci Using Association Results from Multi- ethnic Populations. *C. Liu, the CHARGE—GIANT—CARe—African American Adiposity Consortium.* Biostatistics, Boston University SPH, Boston, MA.

Background: Genome-wide association studies in European Ancestry (EA) samples have identified many loci associated with central obesity, measured by waist-to-hip ratio (WHR). Yet, the causal variants at these loci remain unknown. We used an integrated approach combining results from European and African American (AA) ancestry samples to aid in localizing causal variants at these loci identified in EA samples. **Methods:** Population-specific meta-analysis of association results were obtained from ~77,000 participants of European Ancestry and ~ 23,500 participants of African Ancestry. We interrogated SNPs within the ±250kb flanking region of 14 previously reported waist-related loci in European ancestry populations by performing trans-ethnic meta-analysis of association studies from these samples. We used a Bayesian approach (Morris 2011 - MANTRA) that leverages allele frequency variability across populations and allows heterogeneity in effect size. We applied this approach to meta analytic results from GIANT EA and separately from AA samples. **Results:** All fourteen loci identified in GIANT EA samples were confirmed to be associated with WHR-adjusting for BMI in this joint analysis of EA and AA samples. All fourteen but one (at *CPEB4*) index SNPs retain evidence of association with log-transformed bayes factor > 5. In interrogation of the genomic regions around these 14 loci with trans-ethnic meta-analysis, four index SNPs (rs984222 at *TBX15-WARS2*; rs6784615 at *NISCH-STAB1*; rs1443512 at *HOXC13*; rs4823006 at *ZNRF3-KREMEN1*) still serve as the most significant SNP within the region. Among the other ten loci, the surrogates for the functional variants, with stronger signals than the original index SNPs, were identified from trans-ethnic meta-analysis: rs9286854 at *DNM3-PIGC*, rs2820443 at *LYPLAL1*, rs1128249 at *GRB14*, rs4132228 at *ADAMTS9*, rs10516107 at *CPEB4*, rs1294410 at *LY86*, rs7766106 at *RSPO3*, rs1358980 at *VEGFA*, rs4141278 at *NFE2L3*, rs7302344 at *ITPR2-SSPN*. **Conclusion:** In leveraging varying linkage disequilibrium structures across different populations, the SNPs with strongest signals from our trans-ethnic meta-analysis provide a more precise localization of functional variants for future functional analysis.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1897T

Investigation of genetic variation in scavenger receptor class B, member 1 (*SCARB1*) and association with serum carotenoids. *G.J. McKay[1], E. Loane[2], J.M. Nolan[2], S. Beatty[2], G. Silvestri[3].* 1) Centre for Public Health, Queen's University Belfast, Belfast, Northern Ireland, United Kingdom; 2) Macular Pigment Research Group, Waterford Institute of Technology, Waterford, Ireland; 3) Centre for Vision and Vascular Science, Queen's University Belfast, Belfast, Northern Ireland, United Kingdom.

Objective: To investigate association of scavenger receptor class B, member 1 (*SCARB1*) genetic variants with serum carotenoid levels of lutein (L) and zeaxanthin (Z) and macular pigment optical density (MPOD). **Introduction:** Age-related macular degeneration (AMD) is the most common cause of blindness in older people in developed countries. Previous studies suggest carotenoids confer protective measures against oxidative stress ongoing in the macular region of the retina and variation in *SCARB1* has been implicated in this process. SCARB1 is a key lipoprotein receptor which participates in the transfer of neutral lipids. Common variants within the gene have been reported in association with serum lipid levels. **Methods:** This was a cross-sectional study of 302 healthy adult subjects. MPOD was measured by customised heterochromatic flicker photometry. Fasting blood samples were taken for serum L and Z measurement by HPLC and lipoprotein analysis by spectrophotometric assay. Forty-seven single nucleotide polymorphisms (SNPs) across *SCARB1* were genotyped using Sequenom technology. Association analyses were performed using PLINK to compare allele and haplotype frequencies, with adjustment for multiple testing by permutation. **Results:** Following logistic regression analysis adjusted for age, body mass index, gender, high-density lipoprotein (HDL), low-density lipoprotein (LDL), triglycerides and smoking status, five SNPs (rs11057841, P=3.4×10-4; rs10773109, P=0.002; rs11057830, P=0.003; rs12581963, P=0.004 and rs838873; P=0.005) were significantly associated with serum L levels, although only rs11057841 survived after correction for multiple testing by permutation (P<0.01). Haplotype analysis supported single SNP analysis but failed to identify additional novel association. **Discussion:** Our study has examined common variants in *SCARB1* and identified association between rs11057841 and serum L in healthy subjects. We have adjusted for potential confounding through levels of HDL and LDL cholesterol, triglycerides, BMI, age and gender with the association observed independent of these factors. The high level of linkage disequilibrium between rs11057841 and rs11057830 (r^2=0.93) further supports the significant association observed with serum L. Replication in independent cohorts is necessary to evaluate the role of SCARB1 in the transportation of macular pigments and possible modulation of AMD risk in combating the effects of oxidative stress within the retina.

1898F

Genetic Analysis of FUS/TLS gene in Essential Tremor. *K. Mirzozoda[1], N. Parmalee[1], S. Kisselev[2], N. Merner[4], P. Dion[4,5], G. Rouleau[4,6,7], E.D. Louis[1,8,9,10], L. Clark[1,2,3].* 1) Taub Institute for Research on Alzheimer's Disease and the Aging Brain, Columbia University, New York, NY; 2) Department of Pathology and Cell Biology, Columbia University, New York, NY; 3) Center for Human Genetics, Columbia University, New York, NY; 4) Centre of Excellence in Neurosciences, CHUM Research Center, and Department of Medicine, Université de Montréal, Montréal, Québec, Canada; 5) Department of Pathology and Cellular Biology, Université de Montréal, Montréal, Québec, Canada; 6) Research Center, CHU Sainte-Justine, Montreal, Quebec, Canada; 7) CHUM and the Department of Medicine, Université de Montréal, Montreal, Québec, Canada; 8) Gertrude H. Sergievsky Center, Columbia University, New York, NY; 9) Department of Neurology, Columbia University, New York, NY; 10) Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY.

Background: Although essential tremor (ET) has a genetic basis, specific genes have not been identified. Recently, in a large ET family (FET1) from Quebec, a nonsense mutation (p.Q290X) in the amyotrophic lateral sclerosis (ALS) gene fused in sarcoma/translated in liposarcoma (FUS/TLS) was identified by exome sequencing. No confirmatory studies have been published. Methods: 259 ET cases and 262 controls were enrolled in a study at Columbia University. We performed a comprehensive analysis of the FUS/TLS gene by sequencing all exons in a subsample of 116 ET cases with early onset (≤40 years) ET. We evaluated an association between ET and SNPs in the FUS/TLS gene by genotyping 4 haplotype tagging SNPs in all 259 ET cases and 262 controls. Additionally, 7 clinically associated variants, 2 variants of unknown pathogenicity detected in ALS cases, 8 missense variants predicted to be damaging, and 6 rare variants were genotyped in these 259 ET cases and 262 controls. Results: FUS/TLS mutations previously reported in ALS, the FET1 family, or novel mutations were not found in any of the 116 early onset ET cases. In the case-control analyses, no significant association between tagging SNPs in FUS/TLS and ET was observed, and none of the analyzed SNPs showed evidence of association with ET. Conclusion: Our study suggests that pathogenic mutations in FUS/TLS are rare in a sample of early onset ET cases in the U.S. Also, we did not find evidence that the FUS/TLS gene is a risk factor for ET. Funding for this study was provided by: NIH, NINDS #R01 NS073872 (Louis, Clark), the Parkinson's Disease Foundation (Louis, Clark), the Arlene Bronstein Essential Tremor Research Fund, and the Claire O'Neil Essential Tremor Research Fund.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1899W

Genetic and functional data implicates BCHE as a novel vitiligo candidate gene. *L. Nascimento[1], C. Castro[1, 2], V. Fava[1], R. Werneck[3], M. Mira[1, 3].* 1) Group for Advanced Molecular Investigation, Graduate Program in Health Sciences, School of Medicine, Pontifical Catholic University of Paraná, Curitiba, Brazil; 2) Department of Dermatology, Santa Casa de Misericórdia Hospital, Curitiba, Brazil; 3) School of Health and Biosciences, Pontifical Catholic University of Paraná, Curitiba, Brazil.

Vitiligo is an acquired, systemic, chronic disease of unpredictable evolution, characterized by achromic macules on skin resulting from the absence of functioning melanocytes in the affected areas. The disease affects 1% of the population, both genders equally, with onset at any age. The pathogenesis of vitiligo is still poorly understood; however, experimental data suggests that melanocyte damage may be mediated by impaired clearance of cytotoxic metabolites. In this context, oxidative stress and the accumulation of free oxygen radicals in the epidermal layer of skin may play a role in the pathophysiology of vitiligo. An immunohistochemistry study has shown that epidermal butyrylcholinesterase (BChE) enzyme expression and activity are altered in vitiligo skin, suggesting the BCHE gene as a candidate for genetic analysis. Here we present the results of a family-based association analysis between vitiligo and ten tag SNPs capturing the entire information of the BCHE locus. SEQUENOM MassARRAY and fluorescence-based TaqMan technology were used to genotype 189 family trios recruited from Southern Brazil. Two independent association signals were observed between the disease and BCHE: allele “C” of SNP rs1355538 was associated with an increased risk of vitiligo (P= 0.006; OR 1.90; 95% CI 1.20–3.02) and allele “A” of SNP rs1803274 was associated with protection (P = 0.03; OR 0.70; 95% 0.49–0.98); the latter signal was concentrated among trios presenting non-segmental vitiligo, in the presence of autoimmune co-morbidities. Both association signals were then tested for replication in an independent case-control population sample consisting of 134 non-related affected individuals and 134 controls matched according to age and gender. Again, statistically significant association was observed between allele “A” of SNP rs1803274 and non-segmental vitiligo, in the presence of autoimmune co-morbidities (P = 0.02; OR 0.56; 95% CI 0.32–0.97). Finally, we confirmed previous data showing an impact of rs1803274 alleles over BChE enzymatic activity: a genotype-phenotype correlation experiment in 49 vitiligo-affected subjects resulted in positive evidence for association between rs1803274 allele “A” and lower BChE plasma activity (P = 0.01). In conclusion, our study provides strong experimental evidence supporting BCHE as a novel vitiligo susceptibility gene, and implicates the oxidative stress pathway in vitiligo pathogenesis.

1900T

Targeted deep-resequencing of the CLDN1 gene in African Americans for asthma. *N.M. Rafaels[1], L. Huang[1], C. Vergara[1], R. Lewis[1], L. Gao[1], I. Ruczinski[2], T.H. Beaty[3], A. De Benedetto[4], L.A. Beck[4], R.A. Mathias[1], K.C. Barnes[1].* 1) Division of Allergy and Clinical Immunology, Johns Hopkins, Baltimore, MD; 2) Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins, Baltimore, MD; 3) Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins, Baltimore, MD; 4) Department of Dermatology, University of Rochester Medical Center, Rochester, NY.

Rationale: Asthma is a common complex disease with higher prevalence among African Americans which cannot be entirely explained by environmental, social, cultural, or economic risk factors. Increased airway smooth muscle (ASM) mass is essential to airway remodeling and asthma development. Tight junction proteins, specifically claudin-1 (*CLDN1*), expressed in endothelial and epithelial cells, were found to have increased expression in ASM of asthmatics. **Methods:** We sequenced the entire *CLDN1* gene and 2kb up and downstream using Agilent's targeted deep-resequencing platform on 200 asthmatics and 200 non-asthmatic controls of African ancestry from the Baltimore/Washington area from the *Genomic Research on Asthma in the African Diaspora Study* (GRAAD) previously genotyped on the Illumina HumanHap650Y BeadChip. Genetic associations for risk of asthma were determined by Fisher's exact test under a dominant model for rare variants (minor allele frequency or MAF< 0.05) and logistic regression using PLINK for common variants under the additive model (MAF>0.05) in a final dataset of 183 cases and 192 controls. **Results:** We identified 121 novel variants not previously documented in dbSNP or the Thousand Genomes Project: TGP). Tests of association revealed 13 variants associated with asthma at the α=0.05 level, the most significant being rs9825088 in the 3'-UTR region, and intronic SNPs rs9847792 and rs9834611 (*P*= 0.002). In addition, these 3 variants are common among TGP Yoruban (MAF=0.155), but rare or absent among Europeans and Asians (MAF= 0.003; MAF=0, respectively). None of the 13 variants were included in the GWAS marker panel previously genotyped in GRAAD. **Conclusions:** Our findings suggest variants in *CLDN1* may be associated with asthma in populations of African descent. Validation and replication genotyping in the full GRAAD population is underway.

1901F

TARGETED RESEQUENCING OF SLE SUSCEPTIBLE LOCI. *E. Rai[1], B. Wakeland[1], C. Liang[1], P. Raj[1], K. Viswanathan[1], D. Karp[1], N. Olsen[1], I. Dozmorov[1], L. Davis[1], P. Doshi[1], QZ. Li[1], G. Wiley[2], K. Kaufman[2], JA. Kelly[2], J. Harley[2], P. Gaffney[2], EK. Wakeland[1].* 1) Dept Immunology, UTSouthwestern Med Ctr, Dallas, TX,USA; 2) Oklahoma Medical Research Foundation, Oklahoma City, Oklahoma, USA.

Susceptibility to SLE is impacted by both genetic and environmental factors. More that 30 SLE susceptible loci have been identified, however, the causal variants responsible for these associations are largely unknown. We have developed a high throughput sequencing technology to perform targeted resequencing of 30 SLE susceptible LD segments (~4.3 Mb genomic sequence). 192 Caucasian ((107 SLE cases and 85 controls) were sequenced using combined method of SureSelect array capture (Agilent) and Paired End sequencing protocol from Illumina. High quality non redundant reads were assembled and variations were called using GATK and VCFtools. Of the targeted bases, an average of ~99% bases was captured by at least one non-redundant read and ~93.4% bases by at least 15 non-redundant reads, yielding average fold coverage of ~107X. The high quality sequencing calls were confirmed by having >99% concordance with the Immunochip array data. A total of ~21,000 variations (SNPs and Indels) were identified, of which ~38% were novel. Of the total variations, ~32% potentially impacted function, categorized as ~5% non-synonymous; ~4% synonymous; ~13% UTR; ~2% deleterious; ~1% splice; and 37% cis-eQTL. Most of the non-synonymous and deleterious variations were rare, suggesting that either they are newly evolved or have been subjected to purifying selection. Most of the non-synonymous and deleterious variations were rare, suggesting that either they are newly evolved or have been subjected to purifying selection. Interestingly, a high accumulation of rare deleterious variations restricted to cases were observed in the loci reported to have major effect on SLE susceptibility in humans or animal models (C1Q; TREX1; C2-CFB; MSH5; PTPN22 etc). To explore the allelic architecture of functional variations in tight LD with SLE tagging SNPs, phylogenic networks were drawn using the Median joining networks. Phylogenic networks were drawn using haplotypes formed by potentially functional SNPs in LD with SLE associated SNPs. This analysis identified a specific CLADE of alleles containing multiple SLE associated SNPs in several regions (IRF5, BLK, TNFAIP3, ITGAM etc.) that have strong cis-eQTL impact and identified extensive variations in functional SNPs among families of alleles that carried disease associated SNPs for many other genes. These results suggest that many disease-associated SNPs actually identify a cluster of functional variants, rather than a single disease allele.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1902W

Study of BMP4 gene in a Brazilian sample with Congenital Anomalies of Kidney and Urinary Tract. *G.S. Reis[1], A.C.S. Silva[1,2], T.R. Heilbuth[1], I.S. Freitas[1], L.A. de Marco[1,3], D.M. Miranda[1,2].* 1) Universidade Federal de Minas Gerais - Faculdade de Medicina - INCT MM. Avenida Alfredo Balena, 190, SL114. Bairro Santa Efigenia. Belo Horizonte - MG. CEP.: 30130-100 Brazil; 2) Universidade Federal de Minas Gerais - Faculdade de Medicina - Departamento de Pediatria. Avenida Alfredo Balena, 190. Bairro Santa Efigenia. Belo Horizonte - MG. CEP.: 30130-100 Brazil; 3) Universidade Federal de Minas Gerais - Faculdade de Medicina - Departamento de Cirurgia. Avenida Alfredo Balena, 190. Bairro Santa Efigenia. Belo Horizonte - MG. CEP.: 30130-100 Brazil.

Congenital anomalies of the kidney and urinary tract (CAKUT) occur in 0,5–6% of all births and include a wide spectrum of malformations such as dysplasia, vesicoureteral reflux, polycystic kidney, horseshoe kidney and others. CAKUT are an important cause of renal failure in children and renal end stage disease in adults. CAKUT conditions are polygenic and multifactorial. Genes modulating nephrogenesis are good candidate genes for CAKUT. BMP4, a TGF-β family member, regulates various processes during human development, especially in those organs in which epithelial-mesenchymal interactions are essential for development, as kidney. The BMP4 determine the right insertion position of the ureteric bud. For this reason, we established a case-control study composed by 457 individuals from several areas of Brazil to identify association between BMP4 gene and the CAKUT diagnosis. Samples were classified in two groups: control (n=246) and case (n=211). We investigate three tagSNPs (rs762642, rs207147 and rs17563) allelic and genotype frequencies and determinated its association with CAKUT. The study followed the guidelines of the Declaration of Helsinki and was approved by the Ethics Local Board. Statistical analysis were performed from the UNPHASED 3.1.4©. Differences between non-contiguous variables, genotype distribution and allele frequency were tested by calculating chi-square (X 2) and P <0.05 was considered statistically significant. As expected, the prevalence of CAKUT was higher in male population than in female population. By haplotype analysis we found two statistically significant results for P-value post 1000 permutations: between SNPs rs17563 and rs2071047 (A-G) P-value was 0.01856. Between the three markers, rs17563, rs2071047and rs762642, (A-G-A) we found P-value equal to 0.02117. We also found the following allelic frequency: rs17563, P value = 0.002552. Genotype frequencies were higher when rs17563 was present. Our data suggests a higher risk to develop CAKUT in individuals with polymorphisms in the BMP4 gene. The growing knowledge about the molecular origins of CAKUT might improve the early diagnoses and new targets for the pharmacological treatment of those disorders.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1903T

Detection of 4 new mutations related to Oral Clefts by direct sequencing. *M. Simioni, T.K. de Araujo, R.G Faria, C.V. Maurer-Morelli, V.L. Gilda-Silva-Lopes.* Medical Genetics, FCM/UNICAMP, Campinas, Brazil.

Oral clefts (OC) are the most common craniofacial defects with a worldwide prevalence ranging from 1:500 to 1:2500 newborns. This group includes cleft lip (CL), cleft palate (CP) and cleft lip and (or) palate (CLP). OC can occur as an isolated feature (non-syndromic) or with additional findings (syndromic). Among different etiologies, mutations in several genes have been documented. Considering this, a total of 23 patients with OC (15 syndromic and 8 non-syndromic) were simultaneously investigated by direct sequence for mutation at following genes: IRF6, FOXE1, GLI2, MSX2, SKI, SATB2, SPRY, MSX1, FGF8 and FGFR1. All individuals were previously evaluated by clinical geneticists from Crânio-Face Brazil Project. It was identified in unrelated individuals 4 novel sequence alterations. In syndromic cases, it was detected one mutation at FOXE1 gene (748C ->T; Ala28Ala) in a CLP individual and one mutation at MSX1 gene (329C->T; Ala32Val) in a patient with CL. In non-syndromic, a mutation at GLI2 gene (2341C->T; Leu761Fenil) was found in one case of CL and a mutation in FGF8 gene (765C->A; Glu236Lis) in a CLP child. Except by the MSX1 mutation, which is still under investigation, none of them were detected in a group of 100 controls individuals. Results of in silico analysis to nonsynonymous alterations based in four different algorithms were not concordant about the effect at protein level. Functional studies would be necessary to elucidate the effect of each one. These genes are expressed during embryonic development of lip or palate and had mutations previously described in patients with OC. The results herein described reflect the diversity of genetic factors involved in the etiology of OC and provide new insights for future analyses of FOXE1, MSX1, GLI2 and FGF8 genes. Financial support: Fapesp and CNPq.

1904F

Meta-analysis Demonstrates that an Interleukin-6 Polymorphism is Protective Against Preterm Birth in Women of European Descent. *W. Wu[1], E. Clark[2], G. Stoddard[3], S. Esplin[2], T. Manuck[2], J. Xing[1,4], M. Varner[2], L. Jorde[1].* 1) Department of Human Genetics, University of Utah, Salt Lake City, UT; 2) Dept of Obstetrics and Gynecology, University of Utah, Salt Lake City, UT; 3) Study Design and Biostatistics Center, University of Utah, Salt Lake City, UT; 4) Department of Genetics, Rutgers, the State University of New Jersey, Piscataway, NJ.

Preterm birth (PTB), defined as birth before 37 weeks gestation age, is the leading cause of neonatal mortality and long-term morbidity. The PTB rate is increasing worldwide despite available intervention and prevention strategies. Genetics is a known contributor to PTB and and the heritability of PTB is approximately 30%. PTB is associated with increased concentration of interleukin-6 (IL-6) in maternal serum, as well as cervical and amniotic fluid. Many IL6 polymorphisms have been studied, including -174 G/C, a single nucleotide polymorphism (SNP) located in the IL6 promoter region. Functional studies have shown that the derived C-allele is associated with decreased IL6 expression. Homozygous C/C individuals have lower plasma IL-6 concentration, while individuals with one or more G alleles have higher IL-6 production. However, previous reports associating the IL6–174 polymorphism with PTB have yielded conflicting results. In addition, HapMap data show this SNP has very different frequencies in different populations. Therefore, in addition to small sample size issue, we hypothesize that this conflicting result can in part be explained by differences in population structure and admixture. To address this issue, we conducted a meta-analysis with subgroup analysis by population. We searched PubMed for studies testing associations between the IL6 -174 polymorphism genotype, and the PTB phenotype. We identified eligible studies, extracted and pooled the data, and stratified them by their underlying population structure. Authors were contacted to supply missing information. Maternal and fetal genotypes were analyzed separately. Nine studies were included for maternal genotype analysis. The overall odds ratio (OR) is 0.71 for genotype CC versus CG+GG, with a 95% confidence interval (CI) 0.54 - 0.92, showing a significant protective effect of the CC genotype. For women of European descent, the OR is 0.68 (95% CI:0.51-0.91), but the association is not significant in other populations. For fetal genotypes, only three studies could be included, and the overall OR is 0.92 (95% CI 0.66 - 1.28). In conclusion, the IL6 -174 CC genotype is protective against PTB in women of European ancestry. No significant association was observed for fetal genotypes, or for non-European maternal genotypes.

1905W

Gender-Specific Association of the Interleukin 18 Gene with Gallstone Disease. *H. Yang[1], S. Shih[2, 3], T. Chang[1], H. Wang[2, 3], K. Hu[2, 3], C. Chang[2, 3], C. Chang[2, 3], C. Hung[2, 3], H. Chan[1], M. Lin[1], Y. Lee[1, 4, 5, 6].* 1) Medical Research, Mackay Memorial Hospital, Taipei, Taiwan; 2) Division of Gastroenterology, Department of Internal Medicine, Mackay Memorial Hospital, Taipei, Taiwan; 3) Mackay Medicine, Nursing and Management College, Taipei, Taiwan; 4) Department of Pediatrics, Mackay Memorial Hospital, Taipei, Taiwan; 5) Department of Pediatrics, Taipei Medical University, Taipei, Taiwan; 6) Department of medicine, Mackay Medical College, New Taipei City, Taiwan.

Gallstone disease (GSD) induced strong inflammatory responses and affected extrahepatic bile ducts. Although the pathology and environmental risk factors of GSD are well documented, immune or inflammatory responses in GSD development are still inconclusive. Interleukin 18 (IL18) is a proinflammatory cytokine that plays an important role in immune, infectious, and inflammatory diseases due to induction of interferon-γ. In this study, we investigated whether polymorphisms of the IL18 gene were associated with GSD susceptibility. Genomic DNA was isolated from whole blood samples of 445 patients with GSD and 1121 gallstone-free controls. The IL18 rs549908T>G, rs5744247C>G, rs187238G>C, rs1946518T>G, and rs360719A>G polymorphisms were genotyped using pre-developed TaqMan allelic discrimination assay. No statistical differences of genotype, allele, carrier, and haplotype frequencies of these IL18 gene variants were found between patients with GSD and controls. Interestingly, we found IL18 rs5744247C allele conferred susceptibility to GSD in female patients (OR = 1.33, Pc = 0.015). Haplotype analysis revealed that TGGTA protected against GSD development in females (OR = 0.75, Pc = 0.014). Based on our findings, IL18 rs5744247C>G polymorphism could be a potential genetic marker to predict GSD susceptibility in Han Chinese women.

1906T

Search for new modifiers of disease severity in Jamaican sickle cell disease patients using whole-exome DNA sequencing. *G. Lettre[1,2], G. Galarneau[1,2], M. Beaudoin[1,2], K.S. Lo[1,2], M. Reid[3], G. Serjeant[4], I.R. Hambleton[5], J.N. Hirschhorn[6,7], C.A. McKenzie[3].* 1) Montreal Heart Institute, Montreal, Quebec, Canada; 2) Universite de Montreal, Montreal, Quebec, Canada; 3) Sickle Cell Unit, University of West Indies, Mona, Kingston, Jamaica; 4) Sickle Cell Trust, Kingston, Jamaica; 5) University of the West Indies, Cave Hill, Barbados; 6) Children's Hospital Boston, Boston, MA, USA; 7) Broad Institute, Cambridge, MA, USA.

INTRODUCTION: The β-hemoglobinopathy sickle cell disease (SCD) is characterized by extreme clinical heterogeneity: from asymptomatic presentation to life-threatening complications. To better understand this variable clinical course and identify genetic modifiers of disease, we studied Jamaican SCD patients with >20 years of follow-up. METHODS & RESULTS: Using the Agilent SureSelect All-Exon 50 Mb capture library and the Illumina GAIIx next-generation DNA sequencer, we sequenced the whole-exome of 19 Jamaican SCD patients with a low burden of SCD-related complications. For each patient, we generated ~11 Gb of sequence, providing a mean coverage of 129X, or 84% of the targeted genomic regions re-sequenced at ≥20X. We identified 41,829 high-quality non-synonymous (nonsense, missense) or splice site variants, 25% of which are not reported in public databases. The transition-to-transversion (Ti/Tv) ratio, a metric used to assess the quality of the DNA sequence variants identified, is 3.14. These numbers are consistent with other large-scale next-generation DNA re-sequencing projects, suggesting that the quality of our data is high. To prioritize genes for follow-up in large SCD patient collections, we developed a gene score based on the number and type of mutations found while taking into account important confounders such as gene length. This preliminary analysis yielded several genes involved in lipid metabolism and free radical scavenging, two pathways previously linked to SCD severity. CONCLUSION: We used whole-exome DNA sequencing to identify new genetic modifiers of SCD severity in a unique SCD birth cohort. Combining these sequence results with other available genome-wide association study datasets will provide new opportunities to understand clinical variability in SCD.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

† Trainee Award Finalist

1907F

Complex trait alleles are enriched for cell-specific chromatin marks. *S. Raychaudhuri[1,2,3,4], G. Trynka[1,2,3,4], H. Xu[5], B.E. Stranger[1,4], X.S. Liu[5].* 1) Division of Genetics, Brigham and Women's Hospital, Boston, MA; 2) Division of Rheumatology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA; 3) Partners Center for Personalized Genetic Medicine, Boston, MA, USA; 4) Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA, USA; 5) Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute and Harvard School of Public Health, Boston, Boston, MA, USA.

Complex trait alleles might act by altering regulatory regions. In this case, chromatin marks highlighting active regulatory regions should overlap alleles in the specific cell-types most relevant to the trait. However, those marks that are most useful in defining cell-types and fine-mapping common variants have not been defined. We hypothesized that marks that have greatest utility must be “phenotypically cell-specific”, that is the overlap of alleles with marks is restricted by the phenotype to specific cell-types.

To identify these marks, we examined 15 chromatin marks across multiple cell-types (ENCODE) and 510 independent variants associated with 31 phenotypes. H3K4me3 showed the highest phenotypic cell-specificity ($p<10^{-6}$). We observed only those marks highlighting active gene regulation, such as DNase HS I and H3K9ac, were phenotypically cell-specific. Conditional analysis demonstrated that the phenotypic cell specificity of these marks were largely redundant.

We assessed whether the phenotypic cell-specificity of H3K4me3 was related to the specific location of the peaks or the simple presence of peaks related to cell-specific gene expression. We observed repeatedly that repositioning peaks locally within disease loci reduces phenotypic cell-specificity ($p<0.001$). This suggests that the phenotypic cell-specificity of H3K4me3 is driven by the specific location of cell-specific peaks.

Finally, we examined H3K4me3 marks (NIH epigenome project) to identify key cell-types for four phenotypes. We found that H3K4me3 peaks most significantly overlapped 37 plasma low-density lipoprotein concentration alleles in the liver ($p<10^{-6}$), 40 rheumatoid arthritis alleles in CD4+ memory T-cells ($p=1x10^{-4}$), 32 body mass index alleles in pancreatic islet cells ($p=1x10^{-4}$) and 11 neuropsychiatric disease alleles in the mid frontal lobe ($p=0.015$). We illustrate how these results can be used to fine-map trait-associated variants to causal variants. This study suggests that common, complex trait-associated variants implicate regions involved in cell-specific gene regulation, and that selected high quality chromatin marks can be comprehensively applied to many cell-types to identify critical cell-types and specific regulatory elements in an unbiased fashion.

1908W

A regulatory variant in FZD6 gene contributes to Nonsyndromic Cleft Lip and Palate (NSCLP). *J. Hecht[1], N. Cvjetkovic[1], L. Maili[1], A.M. Letra[2], M. Raia[1], E.C. Swindell[1], J.F. Teichgraeber[3].* 1) Pediatrics, Univ Texas Med Sch, Houston, TX; 2) University of Texas Dental Branch at Houston; 3) Department of Pediatric Surgery, University of Texas Medical School at Houston.

Nonsyndromic cleft lip with or without cleft palate (NSCLP) is a common birth defect with a multifactorial etiology. Despite decades of research, the genetic underpinnings of NSCLP still remain largely unexplained. A genome wide association study (GWAS) of a large NSCLP African American family with seven affected individuals across three generations found evidence for linkage at 8q21.3-24.12 (LOD = 2.98). This region contained three biologically relevant candidate genes: Frizzled-6 (FZD6) (LOD = 2.8), Matrilin-2 (MATN2) (LOD = 2.3), and Solute Carrier Family 25, Member 32 (SLC26A32) (LOD = 1.6). Sequencing of the coding regions and the 5′ and 3′ UTRs of these genes in two affected family members identified a rare intronic variant, rs138557689 (c.-153+432A>C), in FZD6. The rs138557689/C allele segregated with the NSCLP phenotype; in silico analysis predicted and EMSA analysis showed that the 138557689/C allele creates new DNA binding sites. FZD6 is part of the WNT pathway, which is involved in craniofacial development, including midface development and upper lip fusion. Our novel findings suggest that an alteration in FZD6 gene regulation may perturb this tightly controlled biological pathway and in turn contribute to the development of NSCLP in this family. Studies are underway to further define how the rs138557689/C variant affects expression of FZD6.

1909T

Next-Generation sequencing of unresolved Meckel Syndrome pedigrees reveals a complex genetic makeup. *K. Hopp[1], C.M. Heyer[2], J.L. Sundsbak[2], S.J. Koon[1], V.J. Kubly[2], V.E. Torres[2], P.C. Harris[1,2].* 1) Biochem. and Molec. Biology, Mayo Clinic, Rochester, MN; 2) Nephrology and Hypertension Research, Mayo Clinic, Rochester, MN.

Meckel Syndrome (MKS) is an embryonic lethal disorder characterized by polycystic kidney disease (PKD), central nervous system (CNS) defects, polydactyly and liver fibrosis; phenotypic parameters that classify it as a ciliopathy. Genetically, MKS has been linked to autosomal recessive inheritance of variants in 13 ciliogenes, however only five of these are mutated in multiple families, highlighting a high level of genetic heterogeneity driven by privately mutated genes. In concordance, Sanger sequencing of the five commonly associated genes (*MKS1*, *TMEM67*, *CEP290*, *CC2D2A*, *RPGRIP1L*) revealed two mutated alleles in only 51% of our MKS cohort (49 families). Interestingly, in 20% of the cases (10 families) we identified only one pathogenic variant, suggesting a more complex, multi-loci inheritance. To investigate the genetic complexity and to resequence all implicated MKS genes, we performed a targeted *Agilent SureSelect* exon-capture of 707 cilia genes on our genetically unresolved pedigrees. We identified a *CC2D2A* frame-shift mutation (c.3534delG) that was missed by Sanger sequencing in a pedigree that had already one *CC2D2A* change (W1182R). Interestingly, the same patient carried private frame-shifting mutations in three additional ciliogenes and a likely pathogenic missense change in *CEP290* (D433G); highlighting a high level of genetic complexity. In other patients we identified multiple (range 1-5), private and potentially pathogenic variants not present in dbSNP135, the 1000 Genomes or the HapMap project. A number of these are frame-shifting variants in ciliogenes with known homology domains to already identified ciliopathy targets; we are currently evaluating them for their disease association using *in silico* tools (particularly for missense variants), patient cell lines, and over-expression systems. Together, these findings suggest that simple, recessive and more complex, oligogenic inheritance may associate with MKS. In addition, we have developed an *Agilent HaloPlex* gene panel including all known identified ciliopathy genes plus ~200 strong candidates to sequence 96 single indexed samples; mimicking diagnostic settings. In suitable families we are also performing whole exome analysis to identify the full mutation load beyond the ciliary complex.

1910F

Role of the PRICKLE genes in neural tube defects in humans. *R. Allache[1,2], V. Capra[3], M.Q. Wang[2], C.M. Bosoi[2], P. Drapeau[1], A.G. Bassuk[4], Z. Kibar[2].* 1) Pathology and Cell Biology, University of Montreal, Montreal, Quebec, Canada; 2) Department of Obstetrics and Gynecology, CHU Sainte Justine Research Center and University of Montréal; 3) Laboratorio del Servizio di Neurochirurgia, Instituto G. Gaslini, Genova, Italy; 4) Department of Pediatrics, University of Iowa, Iowa City, Iowa, USA.

PK1 and PK2 are core members of the planar cell polarity (PCP) pathway that controls the process of convergent extension (CE) during gastrulation and neural tube closure in vertebrates. This pathway has been implicated in the pathogenesis of neural tube defects (NTDs) in animal models and human cohorts. In this study, we analyzed the role of PK1 and PK2 in these malformations by re-sequencing analysis of their open reading frames and exon-intron junctions in a cohort of 810 unrelated NTD patients. We identified 9 and 7 rare missense heterozygous mutations in each of PK1 and PK2 respectively that were absent in all controls analyzed. Polyphen and SIFT analyses predicted 10 of these mutations to be functionally deleterious. In a yeast-two-hybrid (Y2H) assay, 4 mutations identified in PK1 were found to affect the interaction of PK1 with its binding partners, VANGL2, DVL3 and DIV. In overexpression studies in Zebrafish, 4 zPk2 variants were found to produce a significantly less severe phenotype compared to the wild-type Pk2 suggesting a loss of function mutations. Functional validation of PK2 variants with Y2H assay is currently underway. Our study demonstrates that PK1 and PK2 could act as predisposing genetic factors to NTDs in a subset of patients and further expands our knowledge on the role of PCP genes in the pathogenesis of these malformations.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1911W

Investigation of Complex Copy Number Polymorphisms and Age-related Macular Degeneration. *S. Cantsilieris[1,2], S. White[2], R. Guymer[1], P. Baird[1].* 1) Centre for Eye Research Australia, University of Melbourne, The Royal Victorian Eye and Ear Hospital, Melbourne, Victoria, Australia; 2) Centre for Reproduction and Development, Monash Institute of Medical Research, Melbourne, Victoria, Australia.

Age-related Macular Degeneration (AMD) is a leading cause of vision loss in the developed world. There is compelling evidence that AMD is associated with several gene regions involved in immunity, lipid metabolism and angiogenesis. The role of inflammation, particularly complement regulation, appears to be involved in the disease process. We have undertaken a study to investigate the role of three complex copy number polymorphisms (CNPs) containing, CCL3L1, FCGR3B and C4 involved in encoding chemokines and chemotatic cytokines, chemotaxis and complement regulation. The aim was to determine the potential role of these CNPs in AMD. We employed Multiplex Ligation-dependent Probe Amplification (MLPA) to measure CNV at these three loci. We analysed 346 AMD cases and 377 controls using MLPA, then genotyped a further 230 AMD cases and 44 controls for the C4A and C4B genes using agarose gel electrophoresis, for a total of 576 AMD cases and 421 control samples. We achieved clear separation of copy number integers, ranging between 0-6 for CCL3L1, 1-4 for FCGR3B, 0-4 for C4A and 0-3 for C4B. The frequency distributions for both low and high copy number for CCL3L1 and FCGR3B were not significantly different between cases and controls. In the case of C4A, we observed a 2-fold difference in the number of homozygous deletions between cases and controls (3.1% vs 1.4%) which was nominally significant (p=0.06), t-test. For homozygous deletions of C4B, we observed an almost 4-fold difference between cases and controls (3.5% vs 0.9%) which was statistically significant (p=0.005), t-test. These results suggest that CNPs encompassing C4A and C4B are associated with risk of AMD, and support the role of complement dysregulation in AMD pathogenesis.

1912T

Identification of a pleiotropic effect locus associated with a composite CVD risk trait in the genetic isolate of Norfolk Island. *L.R. Griffiths[1], M. Benton[1], H. Cox[1], C. Bellis[2], M. Carless[2], M. Hanna[1], D. Eccles[1], J. Blangero[2], R.A. Lea[1].* 1) Genomics Research Centre, Griffith Health Institute, Griffith University, Southport, Australia; 2) Department of Genetics, Texas Biomedical Research Institute, San Antonio, Texas, USA.

Many heritable cardiovascular disease (CVD) risk traits tend to be correlated suggesting the underlying presence of pleiotropic effect genes. Genome-wide association studies (GWAS) aimed at composite CVD traits may reveal such genes, which would remain latent using single phenotype analyses. Analysis of large pedigrees offers the added advantage of assessing the heritability of composite traits, which can help prioritise genetically influenced phenotypes for GWAS analysis. In this study we have utilized the Norfolk Island genetic isolate to investigate composite CVD risk traits. This population originates from a small number of British and Polynesian founders who are descendents of the Bounty mutiny with the majority of the population forming a very large multigenerational pedigree. In this study we performed a principal components analysis (PCA) and subsequent heritability (H2) estimation of n=37 CVD-related phenotypes in 330 related individuals from the Norfolk Island genetic isolate. We then performed a pedigree-based GWAS using Illumina 610K quad chips to search for pleiotropic effect loci. PCA revealed 13 components explaining >75% of the total variance within the sample space. Nine components yielded statistically significant H2 values ranging from 0.22 to 0.54 (P<0.05). The most heritable component was comprised of 7 phenotypic measures including; % body fat, waist-to-hip ratio, systolic and diastolic blood pressure, creatine, urea, uric acid. A GWAS of this composite phenotype revealed statistically significant associations for 3 adjacent SNPs on chromosome 1p22 (P<1x10-8). These SNPs form a 42kb haplotype block and explain 11% of the genetic variance indicating a major susceptibility locus for this phenotype. No associations were observed for the phenotypes assessed individually, nor were any significant associations observed for the other 6 heritable component traits. Our results support the `phenome scanning' approach to search for pleiotropic effect loci associated with correlated CVD phenotypes. Further research is now underway to elucidate the causative gene variant(s) within the identified susceptibility region and explain the functional impact on this composite CVD phenotype.

1913F

Association of BMI and waist-related genetic variants with visceral adipose tissue volume in the Family Heart Study. *A. Justice[1], M. Graff[1,2], N. Franceschini[1], A.P. Reiner[3], M.F. Feitosa[4], J.J. Carr[5], J.G. Terry[5], P. Gordon-Larsen[2], M.K. Wojczynski[4], I.B. Borecki[4], K.E. North[1,2].* 1) Epidemiology, University of North Carolina, Chapel Hill, NC, USA; 2) Carolina Population Center, University of North Carolina at Chapel Hill, Chapel Hill, NC; 3) Department of Epidemiology, University of Washington, Seattle, WA; 4) Department of Genetics, Washington University School of Medicine, St Louis, MO; 5) Department of Radiology, Wake Forest University School of Medicine, Winston Salem, NC.

Genome-wide association (GWA) studies have identified multiple single nucleotide polymorphism (SNP) associations with body mass index (BMI) and anthropometric waist measures (waist circumference [WC] and waist to hip ratio [WHR]). Computed tomography (CT) assessed visceral adipose volume (VAT) is a more sensitive and specific measure of central adiposity, which is correlated to both BMI and anthropometric waist measures, and their shared disease risk. However, VAT measurements are generally unavailable in large population-based studies. The extent to which BMI and central adiposity SNPs reflect visceral, general central adiposity or body mass is unknown. We evaluated the association of 38 established BMI and 14 WC/WHR SNPs with VAT and VAT corrected for BMI (VAT/BMI) using GWAS data from the Family Heart Study (FamHS), which was comprised of 2,658 European-American individuals measured by CT scans. Within the FamHS cohort, high correlations exist between VAT and BMI (r^2=0.70), WC (r^2=0.73), and WHR (r^2=0.58). From the 38 BMI SNPs, we found three SNPs for VAT and three for VAT/BMI that had nominally significant ($p \leq 0.05$) associations. Of these, five were non-overlapping between phenotypes, but none of these remained significant after correction for multiple testing. One SNP, rs2241423 was nominally associated with both VAT and VAT/BMI but the BMI increasing allele was negatively associated with VAT and VAT/BMI (β=-8.78±2.80 and -4.92±1.99 respectively). Of 14 WC/WHR SNPs, one was nominally associated with VAT (rs1294421, β= 5.51±2.43) and two with VAT/BMI (rs6795735, β= 3.316±1.66; rs6861681, β= 4.908±1.86), with the waist increasing alleles displaying directionally consistent association with VAT and VAT/BMI. While our findings support the hypothesis that BMI, WC/WHR, and VAT share common genetic influences the inconsistent effect estimates demonstrate that the previous established SNP-BMI associations are expectedly crude and may represent various markers of adiposity, such as overall body size, general adiposity, and more specific measures of central adiposity. These genes deserve further investigation to elucidate the relationship among these BMI genes and actual central adiposity. In contrast, the WC/WHR loci nominally associated with VAT and VAT/BMI display effect estimates that are directionally consistent with previous WC/WHR studies, suggesting that these loci are capturing central adiposity and not overall adiposity.

1914W

Follow up analyses of the APOL1/MYH9 locus and non-diabetic end stage renal disease (non-DM ESRD) in African Americans: the FIND Consortium. *M. Li, FIND consortium.* Epidemiology, Johns Hopkins Univ, Baltimore, MD.

Previous studies identified an admixture-generated signal on chromosome 22 with non-DM ESRD in African Americans (AA). Subsequent work showed that the signal was due to strong associations between the APOL1 G1 and G2 alleles and the MYH9 E-1 haplotype. To further understand the genetic architecture of this locus, we conducted follow-up analyses in the FIND AA non-DM samples: 1) sequencing of exons and the 3′- and 5′-flanking regions of MYH9 in 186 individuals, followed by sequence-enhanced genotyping in 529 non-DM ESRD cases and 727 controls; 2) a haplotype-based association analysis of both MYH9 and APOL1 variants to differentiate the effects of APOL1-only at-risk haplotypes from MYH9-only at-risk haplotypes; and 3) an admixture mapping analysis in 1058 AA individuals who did not have 2 copies of the at-risk APOL1 alleles. In the sequenced samples, 377 SNPs were identified (242 with frequencies<5%). Five exonic SNPs were associated with non-DM ESRD (p<0.05), but the Madsen Browning test was not significant (p=0.68). In the samples with the follow-up genotyping, only one exonic SNP (rs2269529) remained associated with non-DM ESRD (OR= 1.71, p=0.002); however, the association signal shrank to insignificance after adjustment for APOL1-MYH9 haplotypes (OR=1.06, p=0.80). Meanwhile, individuals with 2 copies of the APOL1-only haplotype were about 8 times more likely to have non-DM ESRD compared to those with the low-risk haplotype at APOL1 and MYH9 (OR=8.34;p=4.94 x 10-12). In contrast, those with 2 copies of the MYH9-only at-risk haplotypes were not significantly more likely to have non-DM ESRD (OR=1.03, p=0.92). Finally, in a genome-wide admixture scan of AA individuals without 2 copies of APOL1 at-risk alleles, the genome-wide LOD score on chromosome 22 was -0.07. Sequence analyses suggest that common exonic variants in MYH9 are not associated with non-DM ESRD in AA, and that the APOL1 G1 and G2 risk alleles are significantly associated with non-DM ESRD in a recessive manner and account for the admixture-generated signal on chromosome 22.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1915T

HLA typing of a large type 1 diabetes cohort using SNP genotyping of the extended MHC. *J. Bradfield[1], D.S. Monos[2], H.Q. Qu[3], C. Kim[1], M. Rossman[4], K.D. Rosenman[5], S.F.A. Grant[1,6], C. Polychronakos[7], H. Hakonarson[1,6].* 1) Center for Applied Genomics, Abramson Research Center, The Children's Hospital of Philadelphia, Philadelphia, PA; 2) Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania and The Children's Hospital of Philadelphia, Philadelphia, PA; 3) Division of Epidemiology, Human Genetics and Environmental Sciences, University of Texas School of Public Health, Brownsville Regional Campus, Brownsville, TX; 4) Department of Medicine, University of Pennsylvania School of Medicine, Philadelphia, PA; 5) Department of Medicine, Michigan State University, East Lansing, MI; 6) Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA; 7) Endocrine Genetics Laboratory, McGill University Health Center (Montreal Children's Hospital), Research Institute, Montreal, Quebec, Canada.

Type 1 diabetes is strongly associated with specific HLA polymorphisms and with the MHC region in general. The HLA genes contribute over 50% of the genetic risk for the disease. Due to the highly polymorphic nature of the region and its long-range linkage disequilibrium, identifying additional variants with independent effects, besides those that have already been described, is very challenging. Towards this objective, knowing the HLA type is essential, as any type of analysis needs to both account and stratify the population for this HLA information. For this purpose, and in the absence of HLA typing information, we imputed the classical HLA alleles (HLA-A, HLA-B, HLA-C, HLA-DRB1, HLA-DQB1, and HLA-DPB1) in a cohort that was recently used in a meta-analysis of T1D with approximately 10,000 cases and 17,000 controls. The imputation was performed using Beagle with a reference set of approximately 1,500 individuals who were HLA genotyped and were genotyped on SNP platforms from Illumina and Affymetrix used in the analysis. The exception to this is the HLA-DPB1 reference set, which was only genotyped on the Illumina platform for ~430 individuals. We also imputed all variation (~133,000 SNPs and indels) in the extended MHC from the 1000 genomes project to fill in gaps between the classical HLA alleles. The 1000 genomes imputation was performed with Imputev2. We have identified association with a high degree of resolution of each of the classical HLA alleles, as well as all common diallelic SNPs and indels contained within the extended MHC. While imputing classical HLA alleles may have its shortcomings, we believe that it is a meaningful way to economically analyze large numbers of individuals and to properly control for prior discoveries in the extended MHC.

1916F

Systematic evaluation of validated type 2 diabetes and glycemic trait loci for association with insulin clearance. *M.O. Goodarzi[1], X. Guo[1], J. Cui[1], M.R. Jones[1], T. Haritunians[1], A.H. Xiang[2], Y.I. Chen[1], K.D. Taylor[1], T.A. Buchanan[3], W.A. Hsueh[4], L.J. Raffel[1], J.I. Rotter[1].* 1) Cedars Sinai Medical Center, Los Angeles, CA; 2) Kaiser Permanente Southern California Medical Group, Pasadena, CA; 3) University of Southern California Keck School of Medicine, Los Angeles, CA; 4) Methodist Hospital Research Institute, Houston, TX.

GWAS and GWAS meta-analyses have identified over 50 loci for type 2 diabetes (T2D), as well as a several loci for fasting (FG) and 2-hour glucose (2hrG), fasting insulin (FI), and hemoglobin A1c (HbA1c). Follow-up studies, mainly utilizing fasting or oral glucose tolerance measurements, suggest the great majority of these loci affect insulin secretion, with only a minority appearing to influence insulin sensitivity. Whether validated T2D and glycemic trait loci affect insulin clearance has not been systematically evaluated. We analyzed the Cardio-Metabochip finemapping content for T2D (34 finemapped loci), FG (19 loci), 2hrG (3 loci), FI (2 loci), and HbA1c (5 loci) in two family-based Hispanic-American cohorts, the Mexican-American Coronary Artery Disease (MACAD) cohort, n=676, and the Hypertension-Insulin Resistance (HTN-IR) cohort, n=638. Insulin clearance was measured via the steady state plasma insulin (SSPI) level during euglycemic hyperinsulinemic clamps in both cohorts. ~15,200 SNPs were sufficiently polymorphic and passed quality control for analysis. Association between SNPs and SSPI was evaluated using general estimating equations (GEE1) implemented in the GWAF program, using an additive model (adjusted for age, sex, BMI, diabetes status, and principal components). 23 variants had a $p<0.05$ for association with SSPI in both cohorts, with consistent beta values. We generated meta-analysis p values on these 23 SNPs, which represent 6 loci. The top meta-analysis signals for each locus were rs10241087, $p=1.5\times10^{-5}$ (*DGKB/TMEM195*); rs849334, $p=2.0\times10^{-4}$ (*JAZF1*); rs2380949, $p=3.6\times10^{-4}$ (*GLIS3*); rs55903902, $p=1.3\times10^{-3}$ (*FADS1*); rs35749, $p=4.0\times10^{-4}$ (*IGF1*); and rs9204, $p=4.7\times10^{-3}$ (*HNF1A*). To account for multiple testing, we conducted LD pruning of the 15,200 SNPs, using a LD cutoff of 0.5, yielding 3827 independent SNPs and thus a correction factor of $0.05/3827=1.3\times10^{-5}$. This cutoff was nearly met by the SNP at rs10241087, which lies in an intron of DGKB (diacylglycerol kinase beta, which regulates intracellular levels of the second messenger diacylglycerol, DAG, which may affect hepatic insulin sensitivity). It is possible that additional T2D and glycemia loci discovered subsequent to design of the MetaboChip may also influence insulin clearance. While apparently not the case for the majority of loci, these data raise the possibility that some validated loci for T2D, FG, 2hrG, FI, and HbA1c might exert their effects via insulin clearance in Hispanics.

1917W

Whole Genome Sequencing to Identify Functional Variants that Contribute to Type 2 Diabetes and Obesity in Pima Indians. *K. Huang, C. Bogardus, L. Baier.* Phoenix Epidemiology and Clinical Research Branch, NIDDK, NIH, Phoenix, AZ.

To identify functional variants that affect risk for type 2 diabetes (T2D) or obesity in Native Americans, we are analyzing whole genome sequence data on Pima Indians. Sequencing is being performed by Complete Genomics, Inc., and to date, 30 genomes have been completed. Among these initial genomes, 9,235,137 variants were found (7,283,045 SNPs, 800,901 insertions, 795,389 deletions and 255,802 substitutions) of which 32,596 variants were nonsynonymous (28,041 were SNPs). Multiple prediction programs were used to predict that 12,795 nonsynonymous variants were "functional" (damaging or damaging). These "functional" variants underwent data cleaning: variants located within segmental duplication or non-conserved regions were removed, variants with genotypes called in <80% of the samples were removed, and variation observed in fewer than 3 alleles (MAF<5% in the 30 samples) were removed. The 30 subjects who were sequenced had also been part of a prior genome-wide association study (GWAS) and had been genotyped using the Affymetrix 1 million SNP chip. Therefore, many of the remaining 1,528 "functional" variants could be tagged ($r2 \geq 0.8$) by SNPs previously genotyped in 954 Pima Indians as part of our GWAS. Several of the GWAS tag SNPs (e.g. SNPs near ACTRT2, VNN3 and POC5) had been modestly associated with T2D in the GWAS (tag SNPs' P value $\approx$ 10E-3 - 10E-4). Other GWAS tag SNPs (e.g. SNPs near CCDC67, PKD1L3 and C17orf103) had been modestly associated with BMI in the GWAS (tag SNPs' P value $\approx$ 10E-4). The predicted "functional" variants within these genes, as well as untagged "functional" variants located within physiologic candidate genes for type 2 diabetes and obesity, are currently being genotyped in a population-based sample of 3500 Pima Indians to determine whether they are associated with type 2 diabetes or obesity.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1918T

Tracking elusive modifier loci using identity-by-descent and influential alleles. *E. Marchani[1], E. Wisjman[1,2,3]*. 1) Div Med Gen, University of Washington, Seattle, WA; 2) Dept Biostat, University of Washington, Seattle, WA; 3) Dept Genome Sci, University of Washington, Seattle, WA.

It is not simple to identify genetic loci influencing complex traits, which are often genetically heterogeneous. As modifier and risk alleles are neither necessary nor sufficient to cause disease, they may not be shared by all cases and can hide among unaffected individuals. This is especially true for late-onset traits with censored phenotype data. A rigorous approach is required.

We previously identified three regions linked to modifiers of age-at-onset among Volga German families sharing a *PSEN2* variant causing early-onset Alzheimer's disease. Our goal is to find the biologically relevant variant(s) driving these linkage signals. Sequencing-based analyses at this early stage are prohibitive: the regions are too broad (27.9 - 72.8 Mb) for cost-effective candidate gene screening, and we hesitate to assume the type (copy number variant, missense) or frequency (common, rare) of the underlying variants necessary for exclusion mapping using next-generation sequence data. Instead, our strategy is to find the smallest chromosomal segment with the strongest evidence of linkage, then focus on validating variants co-segregating with that segment in a large pedigree.

For each linkage region, we have tested thousands (6,512 - 11,574) of markers to find those that tag the chromosomal segment driving the linkage signal. We prioritized markers with large effect sizes that eliminate the linkage signal when included as a covariate (4 - 11 markers). We deduced the tagging haplotype with Markov chain Monte Carlo segregation and linkage analyses performed by Loki 2.4.7. Each haplotype has estimated effect sizes comparable to *APOE*. We determined which relatives carry that haplotype using the MORGAN package's computationally savvy identity-by-descent estimation.

This process has allowed us to narrow our focus to chromosomal regions 10 to 30 times smaller than the original linkage signals: 0.9 - 7.6 Mb. We are sequencing carriers of these modifier haplotypes, and will soon identify co-segregating variants. We will also estimate the frequency of these haplotypes among related populations, such as CEPH Europeans and other Volga German families affected by Alzheimer's disease.

1919F

Genetic dissection of Chiari Type I Malformation using stratified whole genome linkage approaches. *C. Markunas[1], K. Soldano[1], K. Dunlap[1], H. Cope[1], E. Asiimwe[1], J. Stajich[1], D. Enterline[2], G. Grant[3], H. Fuchs[3], S. Gregory[1], A. Ashley-Koch[1]*. 1) Center for Human Genetics, Duke University Medical Center, Durham, NC; 2) Division of Neuroradiology, Department of Radiology, Duke University Medical Center, Durham, NC; 3) Division of Neurosurgery, Department of Surgery, Duke University Medical Center, Durham, NC.

Chiari Type I Malformation (CMI) is a developmental disorder characterized by displacement of the cerebellar tonsils below the base of the skull, resulting in significant neurologic morbidity. Although cerebellar tonsillar herniation (CTH) is hypothesized to result from a small posterior fossa (PF), patients are frequently diagnosed by the extent of CTH without cranial morphometric assessment. The resultant clinically heterogeneous CMI population has subsequently hampered gene identification efforts. We performed a whole genome qualitative linkage screen using 66 non-syndromic CMI multiplex families (367 individuals). Both two-point and multipoint parametric and nonparametric linkage analyses were conducted. Initial results showed minimal evidence for linkage. Thus, additional analytic approaches were applied to reduce potential genetic heterogeneity.

First, due to the frequent co-occurrence of CMI with connective tissue disorders (CTD), families were stratified based on a family history of CTD related conditions. Stratified linkage analyses resulted in a marked increase in evidence of linkage to multiple genomic regions. Of particular interest were two regions (Chr8, LOD=3.0; Chr12, LOD=2.1), both of which harbor genes (GDF6, GDF3) implicated in the development of Klippel-Feil syndrome, which co-occurs with CMI in roughly 5% of patients. Sequencing efforts for both genes are underway and currently two putative causal variants have been identified.

Second, using twenty four cranial morphology measurements derived from the PF we estimated heritability for individual PF traits, performed principal components (PC) analysis to identify the most heritable PCs, and performed ordered subset analysis (OSA) using PF traits that were both heritable and associated with CMI. Several interesting regions were implicated from OSA, including chromosome 22. The region harbors EP300, a histone acetyltransferase associated with Rubenstein-Taybi syndrome (RSTS) which can co-occur with CMI. Interestingly, within a morphometrically similar subset of families we observed a slight increase in evidence for linkage to a region containing a related gene, CREBBP, which is also associated with RSTS. These analyses demonstrate the utility of reducing genetic heterogeneity by implicating credible candidate genes in CMI susceptibility.

1920W

Heritability and Linkage Analysis of the Scale of Aging Vigor in Epidemiology (SAVE) in the Long Life Family Study. *J.L. Sanders[1], J. Singh[2], R.L. Minster[2], C.M. Kammerer[2], M.M. Barmada[2], E.W. Daw[3], M. Feitosa[3], R. Costa[4], N. Schupf[4], J. Walston[5], A.B. Newman[1]*. 1) Departments of Epidemiology, GSPH, University of Pittsburgh, Pittsburgh, PA; 2) Department of Human Genetics, GSPH, University of Pittsburgh, Pittsburgh, PA; 3) Division of Statistical Genomics, Department of Genetics, Washington University School of Medicine, St. Louis, MO; 4) Taub Institute for Research on Alzheimer's Disease and the Aging Brain, Columbia University, New York, NY; 5) Center on Aging and Health, Johns Hopkins Medical Institutions, Baltimore, MD.

Frailty, a complex phenotype of older age, is characterized by weakness, slowness, weight loss, low energy, and fatigue. Frailty is strongly associated with increased morbidity, hospitalization, and mortality - thus, frailty is a heavy burden to individuals and communities. Preventing frailty is a prime concern, but its etiology is unknown. There is evidence that unfavorable levels of serum inflammatory markers and hormones, metabolic factors, and short telomeres, all of which are heritable, may spur frailty. However, there are few studies of the possible genetic components of frailty. We used a new trait, Scale of Aging Vigor in Epidemiology (SAVE), a measure of vigor (similar to frailty) that is informative across age groups and comprises measures of strength, speed, and liveliness, to assess possible genetic components of vigor in the Long Life Family Study (LLFS). LLFS is a two-generation, family-based cohort designed to study sibships of exceptionally healthy older individuals and their adult offspring and spouse controls. LLFS consists of 4451 individuals in 480 families; 1512 in the older generation (mean age = 89.4 yr), 2199 in the offspring generation (mean age = 60.5 yr), and 740 spouses (mean age = 61.0 yr). The SAVE trait comprises grip strength, gait speed, physical activity, fatigue, and weight loss; each scored 0, 1, or 2 using approximate tertiles; and summed from 0-10 to create a continuous, normal vigor distribution in both generations, thus vigor decreases with an increasing score. Using maximum likelihood methods, heritability of this trait was estimated overall (age-adjusted $h2=0.33\pm0.04$, $p<10^{-21}$). We next performed variance components linkage analyses using multipoint IBD probabilities estimated using SNPs with approximately 1.0cM spacing. Although we did not detect any significant signals (i.e., LOD > 3.0), we obtained suggestive evidence that a quantitative trait locus for SAVE was located near the end of chromosome 18q (maximum LOD = 2.9 and the chromosomal region of interest encompassed 118 - 131 cM). These findings support further evaluation of the genetic and environmental factors contributing to this phenotype of vigor in the elderly.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1921T

Genomewide linkage analyses of hematological phenotypes in Long Life Family Study. *J. Singh[1], R.L. Minster[1], M.M. Barmada[1], E.W. Daw[2], B. Thyagarajan[3], K. Christensen[4], A.B. Newman[5], C.M. Kammerer[1].* 1) Department of Human Genetics, GSPH, University of Pittsburgh, Pittsburgh, PA; 2) Division of Statistical Genomics, Department of Genetics, Washington University School of Medicine, St. Louis, MO; 3) Department of Laboratory Medicine and Pathology, University of Minnesota, Minneapolis, MN; 4) Institute of Public Health at the University of Southern Denmark, Odense, Denmark; 5) Department of Epidemiology, GSPH, University of Pittsburgh, Pittsburgh, PA.

Hematological phenotypes such as white blood cells (WBC), red blood cells (RBC) and platelets play important roles in immune response, oxygen carrying and blood clotting. These traits are heritable and tightly regulated within narrow physiological limits. Abnormalities in these traits are associated with a number of diseases such as anemia. To detect possible quantitative trait loci (QTL) involved in regulation of these traits, we performed genomewide linkage analyses using data from the Long Life Family Study (LLFS) - a two-generation family-based cohort study designed to elucidate the genes and environmental factors that influence exceptional aging. Data on hematological phenotypes (including hematocrit, hemoglobin, lymphocytes, neutrophils, RBC, and WBC counts) were available on 4520 individuals belonging to 480 families of European ancestry. Based on the literature, we used maximum likelihood variance components methods to assess traits for effects of the following covariates (field center, age, sex, age-squared, BMI, smoking status, drinking habits, and menopausal status); residuals of each trait were obtained for significant ($p \leq 0.05$) covariates and used in subsequent analyses. Similar to previous reports, residual additive genetic heritability ranged from 0.20±0.03 (for hemoglobin) to 0.42±0.04 (for platelet count); all p-values $<10^{-21}$. Variance component linkage analyses were performed using multipoint identical by descent probabilities (IBD) estimated from SNPs with approximately1cM spacing. We obtained evidence for a QTL influencing RBC count on chromosome 11p15.1 to 11p15.2 (maximum LOD score = 3.2 at 37 cM, genomic p-value <0.05). Interestingly, the Framingham Heart Study previously reported linkage of RBC count to a QTL at 11p15.2. The hemoglobin beta-cluster located at 11p15.5 (>10 Mbp away) is outside our region of interest. We also obtained evidence for a QTL for platelet count on chromosome 8p22 (maximum LOD = 3.4 at 31cM, genomic p-value<0.05). Reports from large consortium genomewide association studies have not implicated either chromosomal region, indicating that we may have identified additional genes influencing RBC count and platelets. Our findings illustrate the utility of long-lived families to enhance our understanding of biology involved in genetic control of hematological traits.

1922F

Effect of DNA polymorphisms of apolipoprotein B gene on lipid homeostasis in obese Egyptians. *G. El-Kannishy[1], R. Elbaz[2], A. Wafa[1], H. AbdElHafez[1], A. Settin[3].* 1) Internal Medicine Dept., Mansoura Faculty of Medicine, Mansoura University, Mansoura, Egypt; 2) Genetic Unit, Faculty of Medicine, Mansoura University; 3) Peadiatric Dept., Faculty of Medicine, Mansoura University.

Background: Numerous polymorphisms of Apoliporotein B (ApoB) gene may predispose to obesity, particularly a variable number of tandem repeats (VNTR) polymorphism. association between the 3′APOB-VNTR polymorphism and serum lipid levels was observed in many ethinic groups. Objective: The aim of this study is to investigate the association of the 3′VNTR polymorphisms of the ApoB gene in obese Egyptians with normal and abnormal lipid profile. Methods: The 3′APOB-VNTR alleles were determined and classified according to the number of repeats in 180 subjects with simple obesity as well as 100 age and sex matched healthy controls. Obese group was further subdivided according to lipid profile into normolipidemic and dyslipidemic groups. Results: Different numbers of alleles were identified in the studied groups. The most frequent allele in all the studied groups was VNTR 31, followed by VNTR 38. VNTR 38 repeats were significantly higher in control group than total obese and dyslipidemic groups. Dyslipidemia was associated with significantly higher frequency of the homozygous VNTR-SS genotype; on the otherhand, normolipidemic group had a significantly lower frequency of the heterozygous VNTR-LS genotype compared to healthy controls. Conclusions: VNTR 38 repeats seems to be protective from lipid abnormality in normal weight and obese. High VNTR-SS genotype among subjects with dyslipidemia possibly indicate that VNTR-SS genotype could be considered as a genetic risk factor for lipid abnormality among obese Egyptians.

1923W

Identification of genes influencing serum levels of brain-derived neurotrophic factor in large Mexican American pedigrees. *M.A. Almeida[1], J.M. Peralta[1], J.W. Kent[1], J.E. Curran[1], T.D. Dyer[1], G. Juan[2], T.M. Teslovich[2], C. Fuchsberger[2], A.R. Wood[3], T.M. Frayling[3], P. Cingolani[4], T.W. Blackwell[2], R. Sladek[4], G. Atzmon[5], J. Laramie[6], S. Lincoln[6], D.M. Lehman[7], G. Abecasis[2], L.A. Almasy[1], R. Duggirala[1], D.C. Glahn[8], J. Blangero[1].* 1) Department of Genetics, Texas Biomedical Research Institute, San Antonio, TX; 2) Center for Statistical Genetics, University of Michigan, Ann Arbor, MI 481092; 3) Peninsula Medical School, University of Exeter, Exeter UK; 4) McGill University, Montreal, Quebec, Canada; 5) Departments of Medicine and Genetics, Albert Einstein college of medicine, Bronx NY, 10463; 6) Complete Genomics; 7) University of Texas Health Science Center at San Antonio; 8) Yale University, New Haven, CT.

Brain-derived neurotrophic factor (BDNF), a member of the neurotrophin family, is a 14 kDa polypeptide involved in regulating neuronal survival, differentiation, and outgrowth. Serum levels of the BDNP protein have previously been associated with a number of psychiatric diseases such as depression and also with neurodegenerative diseases. We measured serum BDNF levels in a sample of 860 San Antonio Family Heart Study (SAFHS) individuals who are members of extended Mexican American pedigrees. To localize genomic regions influencing BDNF, we performed quantitative trait linkage analysis utilizing high density SNPs to calculate identity-by-descent probabilities. Genome-wide linkage analysis localized four genomic regions harboring quantitative trait loci (QTLs). The observed LOD scores were 4.45 (chromosome 22), 3.29 (chromosome 1), 2.41 (chromosome 19) and 2.39 (chromosome 5). We then examined existing whole genome sequence (WGS) data for a subset of 491 individuals to search for causal variants/genes in these genomic regions. WGS data (either directly sequenced commercially by Complete Genomics or accurately imputed using familial data as part of the T2D-GENES Consortium) were available for 491 total individuals. Each variant within a QTL-specific 1-LOD support interval was tested for association with BDNF levels using a measured genotype model in a variance component framework. Variants were also prioritized for testing by their functional potential. Three of the QTL regions yielded substantial evidence for causal gene identification. For the chromosome 22 QTL, we identified a non-synonymous variant (N759D) in the PLXNB2 gene that was significantly associated with BDNF. This variant is predicted to be strongly deleterious. The PLXNB2 gene is involved in axon growth and guidance. For the chromosome 5 QTL, we identified multiple significant associations ($p = 1 \times 10\text{-}6$) including a private novel variant in an extremely evolutionarily conserved region in the IRX1 gene. This gene represents an excellent causal candidate with a known major role in neurodevelopment. For the chromosome 19 QTL, our top hits ($p < 5 \times 10\text{-}7$) occurred in rare variants in/near GNG7 that has been previously associated with epilepsy. These results empirically demonstrate the potential of deep sequence data to follow up QTL localization studies to identify novel likely causal variants/genes.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1924T

First genome-wide analysis in pediatric multiple sclerosis (MS) confirms a role for adult MS risk variants and reveals new candidates. *L.F. Barcellos[1], J.R. Oksenberg[2], E. Elboudwarej[1], H. Quach[1], F. Briggs[1], A. Belman[3], A. Chokkalingham[1], P.A. Buffler[1], L. Krupp[3], E. Waubant[2].* 1) Div Epidemiology-SPH, UC Berkeley; 2) Dept Neurology, UC San Francisco; 3) Dept Neurology, Stony Brook University, NY.

Multiple sclerosis (MS), a complex autoimmune disorder, is the leading cause of non-traumatic neurological disability in young adults. Disease onset typically occurs between the ages of 20 and 40; however, up to 10% of all MS patients have symptoms before 18 years. Pediatric MS (onset before age 18) affects between 10,000 and 20,000 patients in the United States. Disease course in children is similar to adults with some exceptions. Compared to adults, young children with MS are more likely to develop seizures and ataxia during the course of the disease; males are equally affected compared to females within the youngest pediatric MS cases. In addition, cases almost exclusively have a relapsing-remitting course, and are less likely to be of European ancestry. Exact mechanisms involved in pediatric MS pathogenesis remain unclear; however, similar to adult MS, both genetic and environmental factors have been implicated in disease susceptibility. A role for DRB1;k415:01 within the MHC has been established for MS, regardless of onset age. The largest GWAS ever performed in adult MS recently identified 52 non-MHC MS variants (WTCCC and IMSGC; Nature 2011). Here, we describe the first GWAS study in pediatric MS using the Illumina platform. Whole genome profiles for 200 pediatric MS cases and 400 race/ethnicity, gender and age matched pediatric control individuals were studied. Genetic imputation was performed using IMPUTE2, using the 1000 Genomes Phase 3 integrated variant set (March, 2012). The C-alpha method was used for rare variant analysis. Genes identified through GWAS results for both pediatric and adult MS were subjected to pathway bioinformatics analyses and compared for similarities and differences using DAVID (v 6.7). Variants within calcium signaling and axon guidance pathway genes were significantly over-represented among associated hits in pediatric MS ($p<10-4$; based on KEGG). Top established non-MHC adult MS risk variants associated with pediatric MS included rs949143 within MPHOSPH9 (OR= 1.82, 95% CI 1.38-2.40, p=2.3 x 10-5), rs229370 within TMEM39A (OR= 0.55, 95% CI 0.39-0.80, p=0.001) and rs3118470 within IL2RA (OR=0.67, 95% CI 0.51-0.90, p=0.007). Results suggest variants within genes involved in adult-onset MS are also risk factors in children and suggest that similar biological processes are present in both groups. Thus, disease-modifying therapies approved for adults with MS may similarly benefit children.

1925F

Metabochip analysis in over 12,000 African Americans identifies several variants associated with Metabolic Syndrome: results from the Population Architecture using Genomics and Epidemiology (PAGE) Study. *C.L. Carty[1], J. Haessler[1], I. Cheng[2], V. Aroda[3], T. Thornton-Wells[4], C.N. Hsu[5], S. Liu[6], R. Jackson[7], C. Carlson[1], L.A. Hindorff[8], L. LeMarchand[2], J.S. Pankow[9], U. Peters[1], K.E. North[10], C. Kooperberg[1].* 1) Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA; 2) Epidemiology Program, University of Hawaii Cancer Center, Honolulu, HI; 3) MedStar Health Research Institute, Hyattsville, MD; 4) Center for Human Genetics Research, Vanderbilt University, Nashville, TN; 5) Information Sciences Institute, University of Southern California, Marina del Rey, CA; 6) Department of Epidemiology, School of Public Health, University of California, Los Angeles, CA; 7) Department of Internal Medicine, Ohio State University, Columbus, OH; 8) Office of Population Genomics, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD; 9) Division of Epidemiology and Community Health, University of Minnesota, Minneapolis, MN; 10) Department of Epidemiology and Carolina Center for Genome Sciences, University of North Carolina, Chapel Hill, NC.

Metabolic syndrome (MetS) is characterized by a clustering of cardiometabolic risk factors including dyslipidemia, central obesity, high blood pressure and high glucose. We sought to identify MetS susceptibility loci and investigate whether genetic variants with pleiotropic effects account for some of the correlated architecture of the MetS traits. Using the Metabochip genotyping array (n=161,098 SNPs passing quality control), we investigated SNP associations with MetS and its component traits in African Americans from the PAGE Study. Based on NCEP ATP III guidelines requiring ≥3 of 5 criteria to define MetS, we had 4989 cases and 7576 controls. We contrast findings from a subset-based meta-analysis method, ASsociation analysis based on subSETs (ASSET), (Bhattacharjee *et al.* 2012) with those from 1)logistic regression modeling MetS as a binary outcome, and 2)traditional meta-analysis of the 5 individual components of MetS coded as binary variables. Unlike the traditional meta-analysis which may lack power when associations for each of the components are null or have effects in opposite directions, ASSET uses two one-sided tests to detect positive and negative results for the components separately and combines the tests taking into account correlations among the components. Using ASSET, we identified 28 SNPs in 5 loci (chromosomes 8, 10, 11, 16 and 19) significantly associated with MetS, all $p< 2.5e-7$, the Bonferroni adjusted p-value. The chromosome 10 and 16 SNPs were associated with multiple MetS criteria: glucose and waist circumference, and HDL and waist circumference, respectively. SNPs on chromosomes 8,11 and 19 were associated with both lipid criteria. In contrast, using standard logistic regression or meta-analysis of all components we detected only one low density lipoprotein cholesterol SNP in the 19q13.32 locus, both $p< 2.5e-7$.

We identified several variants associated with MetS components in the ASSET analysis. While we found evidence of pleiotropy for some of these significant SNPs, none were significantly associated with ≥3 traits, as required for MetS. This study highlights the value of using novel methods to increase power and efficiency in large scale genomic association analyses. Importantly, it also brings into question the MetS construct, which is inherently heterogeneous by definition, and its utility in genetic studies in which SNPs may be associated with one MetS trait and not with another.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1926W

Power of population diversity and positive selection in probing the biology of asthma disparities. *G. Dunston[1,2,4], T. Mason[2], L. Ricks-Santi[2,3,4].* 1) Dept MicroBiol, Howard Univ, Washington, DC; 2) National Human Genome Center; 3) Dept Pediatrics Child Health; 4) Howard University Cancer Center.

Purpose: Asthma, a common chronic inflammatory disease caused by gene(s) and environmental interactions, has a worldwide distribution and is associated with alterations in airway smooth muscle reactivity and remodeling. GWAS have revealed several SNPs associated with asthma. However, the role of many of these SNPs in the etiology of asthma is unknown. We investigated asthma associated SNPs in genes with signatures of positive selection to evaluate population diversity in genetic susceptibility to asthma. Methods: The NHGRI GWAS Catalog was queried for signals associated with asthma and atopy. SNPs in genes were interrogated for strong signatures of recent positive selection via SNP;caEvolution, a database using a sliding window method that measures heterozygosity and the fixation statistic to locate genomic regions under selection. Regional results reflect geographical adaptation, founder effects and fixed or unfixed selections. The selection profile of genes surrounding those with signatures of positive selection was also examined. For genes with strong signatures of recent positive selection, the literature was mined and queried for information on their disease association, function, pathway analyses, and molecular processes relevant to the biology of asthma. Results: Forty-four SNPs were significantly associated with asthma and atopy. Twenty-one of these SNPS were in genes and 2 of these genes (i.e., RAD50 and CDK2) showed strong signatures of recent positive selection. While RAD50 was positively selected in the Gujarati Indians from Houston, TX, CDK2 showed strong signatures of recent positive selection in all the African descent populations in the International Haplotype Map Project. Four additional genes with a similar positive selection profile as CDK2 were discovered in the scanned region surrounding CDK2. Commonalities were noted between asthma associated genes with signatures of positive selection and schistosomiasis. Conclusion: Positive selection can be used as a tool to interrogate population diversity in GWAS signals associated with the biology of common complex diseases of public health interest (i.e., health disparities), like asthma. Additionally, this study highlights the possible effect of adaptation to infectious disease in GWAS signals associated with asthma and atopy in African descent populations. Research supported in part by NIH Grants NCMHD 5P20 MD000198; NIGMS S06 GM08016; NCRR 2 G12 RR003048 from the RCMI Program.

1927T

Conditional Analysis Identifies Three Novel Major Histocompatibility Complex Loci Associated With Psoriasis. *J. Knight[1], S.L. Spain[2], F. Capon[2], J.N. Barker[3], M.E. Weale[2], R.C. Trembath[4], WTCCC2, Genetic Analysis of Psoriasis Consortium, I-chip for Psoriasis Consortium.* 1) Neuroscience Research, Centre of Addiction and Mental Health, Toronto, ON, Canada; 2) Division of Genetics and Molecular Medicine, King's College London School of Medicine, London, United Kingdom; 3) St John's Institute of Dermatology, King's College London, London, United Kingdom; 4) Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, United Kingdom.

Many genetic loci have now been shown to influence psoriasis and these implicate a number of genes providing evidence for an integrated model for the pathogenesis of psoriasis that combines skin barrier function, innate immune pathogen sensing, adaptive immunity and Th17 cell responses. There is one major disease susceptibility locus located within the major histocompatibility complex (MHC), in which the signal is thought to be driven by *HLA-C*. Other independent association loci have been identified within the MHC, but determining the number and location of these has been hampered by extensive linkage disequilibrium across the region. We leveraged the power of large discovery and replication datasets of European ancestry to look for further association signals in the MHC region. In addition to the major loci at *HLA-C* ($p=2.20 \times 10^{-236}$), we found and replicated four additional hits, three of which are novel. The results from the discovery sample are as follows; rs2507971 $p= 6.73 \times 10^{-14}$, rs9260313 $p=7.93 \times 10^{-09}$, rs66609536 $p=3.54 \times 10^{-07}$ and rs380924 $p=6.24 \times 10^{-06}$), all replicate in an independent sample ($p<=0.01$). All lie within the MHC Class I region. The previously identified loci lies near *MICA*, the other three lie near *MICB*, *HLA-A* and *HCG9* (a non-coding RNA). The identification of disease associations with both *MICA* and *MICB* is particularly intriguing, since each encodes an MHC Class I-related protein with potent immunological function.

1928F

Prevalence of MYH9 common variants associated with non-diabetic end-stage renal disease and chronic kidney disease in the diverse National Health and Nutrition Examination Surveys as part of the Epidemiologic Architecture for Genes Linked to Environment (EAGLE). *S. Wilson, K. Brown-Gentry, N. Gillani, H. Jin, B. McClellan, J. Boston, C. Sutcliffe, H. Dilks, D. Crawford.* Vanderbilt University Medical Center, Nashville, TN.

Common genetic variants in the myosin, heavy chain 9, non-muscle (MYH9) gene were initially identified as associated with non-diabetic end-stage renal disease in African Americans. ESRD MYH9-associated variants rs4821480, rs4821481, and rs2032487 were demonstrated to be very frequent in African Americans compared with European-descent populations. Subsequent studies have suggested that MYH9 common variants also associated with chronic kidney disease and other measures of kidney function in various populations. To further characterize the prevalence of MYH9 variants and their association with kidney related traits, we have genotyped intronic rs4821480, intronic rs4821481, and intronic rs2032487 in 7,389 participants from the National Health and Nutrition Examination Surveys (NHANES) 1999-2002 as part of the Epidemiologic Architecture for Genes Linked to Environment (EAGLE). NHANES is a cross-sectional population-based survey of the United States conducted by the National Center for Health Statistics at the Centers for Disease Control and Prevention and consists of three major groups: non-Hispanic whites (NHW; n=4,003), non-Hispanic blacks (NHB; n=1,350), and Mexican Americans (MA; n=1,877). In this general adult (18 years old or greater) population ascertained regardless of health status excluding pregnant participants, rs4821480-G, rs4821481-C, and rs2032487-C alleles were most frequent among non-Hispanic blacks (0.63, 0.63, and 0.64) compared with either NHW (0.05 each) or MA (0.07 each). In linear regressions assuming an additive genetic model adjusted for age and sex, none of the MYH9 variants tested was associated with serum creatinine, serum albumin levels, or urinary albumin levels in this general population sample stratified by race/ethnicity ($p>0.05$). The lack of associations is most likely due to power highlighting the importance of the availability of large, diverse populations for genetic associations studies of complex diseases and traits.

1929W

Both classical and non-classical human leukocyte antigen (HLA) loci in the HLA region contribute to Graves disease (GD) susceptibility. *P. Chen[1,2,3], W. Yang[2,4,5], T. Chang[2,5].* 1) Department of Medical Genetics, National Taiwan University Hospital, Taipei, Taiwan, Taiwan; 2) Department of Internal Medicine, National Taiwan University Hospital, Taipei, Taiwan, Taiwan; 3) Graduate Institute of Clinical Genomics and Proteomics, National Taiwan University, Taipei, Taiwan, Taiwan; 4) Graduate Institute of Clinical Medicine, National Taiwan University, Taipei, Taiwan, Taiwan; 5) Department of Internal Medicine, National Taiwan University, Taipei, Taiwan, Taiwan.

GD (MIM 27500), the worldwide leading cause of hyperthyroidism and thyroid eye disease, is a common organ-specific autoimmune disorder with genetic predisposition. Our team previously published a GD-HLA association study using the direct genotypes of the 6 classical HLA alleles in 1704 individuals of ethnic Han Chinese (Chen *et al*. *PLoS One* 6, e16635 (2011)). Later Chu *et al*. (*Nat Genet* 43, 897-901 (2011)) reported a SNP genome-wide association study of GD, in which HLA demonstrated the highest association signals. It would be important to clarify whether the association signals in the HLA region from these two studies represented the same signals or not. In order to make comparison between our results and the report from Chu *et al*., we further genotyped those 5 SNPs (rs4947296, rs1521, rs6903608, rs6457617 and rs2281388) in the HLA region reported by Chu et al. in 499 unrelated GD cases and 504 unrelated controls in our collection from whom we have comprehensive genotypes of 6 classical HLA loci (*HLA-A*, *-B*, *-C*, *-DPB1*, *-DQB1* and *-DRB1*). We demonstrated that rs2281388 is in high linkage disequilibrium (LD) with the *HLA-DPB1;k405:01* allele (D' = 0.98, r-square = 0.86) in the ethnic Chinese Han population. This SNP was the SNP showing the strongest signal (odds ratio [OR] = 1.64, $P = 1.50 \times 10^{-65}$) in the GD association study by Chu *et al*., and it is also the SNP repeatedly being shown association with chronic hepatitis B infection in Japanese and in Chinese. This result implicates that *DPB1;k405:01* is the genuine susceptibility allele in the DPB1 region. Three out of 4 other SNPs also showed statistically significant associations in our cohort (rs1521: OR = 3.04, $P = 3.10 \times 10^{-3}$; rs6903608: OR = 1.83, $P = 5.02 \times 10^{-11}$; rs6457617: OR = 1.48, $P = 3.09 \times 10^{-5}$). However, there were no classical HLA alleles corresponding to the SNP associations. The LD between these 3 SNPs and any of our significant classical HLA alleles was quite low (all r-square < 0.09). This study illustrates that direct classical HLA genotype-based and indirect SNP-based association studies are complementary approaches. It also implicates that both classical and non-classical HLA loci in the HLA region contribute to GD susceptibility.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1930T

Fine mapping on chromosome 10q24.2 implicates *ADD3* in biliary atresia. *M.-M. Garcia-Barcelo[1], G. Cheng[1], C. S. M. Tang[2], X. L. Liu[1], R.Z. Zhang[1], M. T. So[1], E. H. M. Wong[2], P. H. Y. Chung[1], I. H. Y. Chan[1], J. Liu[3], W. Zhong[4], H. Xia[4], J. Yu[4], K. K. Y. Wong[1], S. S. Cherny[2], P. C. Sham[2], P. K. H. Tam[1].* 1) Dept Surgery, Univ Hong Kong, Hong Kong, Hong Kong; 2) Dept Psychiatry, Univ Hong Kong, Hong Kong; 3) Department of Pediatric Surgery, First Affiliated Hospital of Sun-Yatsen University, Guangzhou, China; 4) Department of Surgery, Guangzhou Women and Children's Hospital, Guangzhou, China.

Biliary Atresia (BA [OMIM 210500]) is a major cause of neonatal cholestasis (arrest of the normal flow bile) and is characterized by progressive fibrosclerosing and inflammatory obliteration of the extrahepatic biliary system during the first few weeks of life. The only current treatment is surgical (portoenterostomy -Kasai operation). The condition presents mainly sporadically, with only a few familial cases reported. BA is a rare congenital disorder whose incidence varies widely among populations (from 1 in 5,000 in Asians to 1 in 18,000 in Caucasians), being most common in East Asia and Polynesia. Through a genome-wide association study (GWAS) on Chinese individuals we identified the association of a 129Kb BA-associated region on chromosome 10q24.2 encompassing two biologically plausible genes, *ADD3* and *XPNPEP1*. The strongest association was for a Single Nucleotide Polymorphism (SNP; rs17095355; p=6.94x10-9), intergenic between ADD3 and XPNPEP1. To determine the genetic architecture within this region we increased the genotype density by genotyping 117 tag-SNPs in 339 Chinese BA patients and 389 Chinese controls and also used imputation. rs17095355 remained the most associated marker within the region (p=6.06 x10-9). To elucidate the risk-contribution model of regional variation, we performed stepwise regression and sequential conditional test of the genotyped SNPs with BA phenotype. The disease risk model included 6 SNPs (i.e. rs975442, rs17095355, rs10509906, rs11194975, rs2501577 and rs11194997), in which rs975442, rs10509906, rs11194975, rs11194997 were variants newly included in the fine mapping study but not in GWAS. In this model, rs10509906 was independently associated with BA (p=0.018 conditioned on rs17095355). This model explains 8.9% of the phenotypic variation (R2= 0.031 in GWAS) and indicates that multiple variants may contribute to BA susceptibility. The risk alleles of rs17095355 and rs10509906 were ancestral, suggesting an ancestral-susceptibility model of BA pathogenesis. We found that the haplotype comprising the risk allele of rs17095355 was associated with reduced *ADD3* gene expression in liver of BA patients. Thus, our data suggests that *ADD3* expression level in BA liver may contribute to BA pathogenesis. *ADD3* is a component of the actin-adducin-spectin complex which is known to be associated with pathologic cholestasis.

1931F

Whole exome sequencing and subsequent association studies reveals novel candidate variants associated with severe types of alopecia areata. *S. Lee[1,2], C. Park[1], OS. Kwon[3], J-I. Kim[1,2], J-S. Seo[1,2].* 1) Biomedical Sci, Seoul National Univ, Seoul, South Korea; 2) Genomic Medicine Institute (GMI), Medical Research Center, Seoul National University, Seoul, South Korea; 3) Department of Dermatology, Seoul National University College of Medicine, Seoul, South Korea.

Alopecia areata (AA) is a common autoimmune disorder mostly presented as round patches of hair loss and subclassified into alopecia totalis/alopecia universalis (AT/AU) based on the area of alopecia. Although AA is relatively common, AT/AU is exceptionally rare and only 5% of AA patients progress to AT/AU. To determine genetic determinants of this orphan disease, we undertook an exome-wide association study using 6 samples from AU patients, which were compared with 155 Asian controls. Using Fisher's exact test, our study revealed 27 immune-related candidates; 8 candidates were replicated in 14 independent AU samples (P < 0.05). Linkage disequilibrium was observed between some of the most significant SNPs, including rs41559420 of HLA-DRB5 (P < 0.001, OR 44.57) and rs28362679 of BTNL2 (P < 0.001, OR 30.21). While BTNL2 was reported as a general susceptibility gene of AA, HLA-DRB5 has not been implicated in AA. In addition, we have found several genetic variants on novel genes (HLA-DMB, TLR1, PMS2, LTF, and DMBT1), and replicated an additional 2 loci (HLA-A, C5). This study provides further evidence for association of previously reported immune-related genes with AA and novel findings such as HLA-DRB5, which might represent a hidden culprit gene of the disease.

1932W

Next-generation sequencing of 11 targeted genes show serum urate levels and gout age of onset are influenced by both rare and common variants. *A. Tin[1], L. C. Shimmin[2], A. Kottgen[1,3], E. Boerwinkle[2], J. Coresh[1], J. E. Hixson[2], W. H. Kao[1].* 1) Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, U.S.A; 2) Department of Epidemiology, Human Genetics and Environmental Sciences, School of Public Health, The University of Texas Health Science Center at Houston, Houston, TX, U.S.A; 3) Department of Internal Medicine, Renal Division, University Medical Center Freiburg, Freiburg, Germany.

Background. Elevated serum urate level is an important risk factor for gout, which affects 3 million Americans. Genome-wide association studies (GWAS) have identified multiple loci associated with serum urate, but functional mutations in most, except for ABCG2, remain to be found. Method. We sequenced the exons, promoter regions, and 3′ flanking regions of 11 genes identified from serum urate GWAS (PDZK1, GCKR, SLC2A9, ABCG2, RREB1, SLC17A1, SLC16A9, SLC22A11, SLC22A12, R3HDM2, INHBC) in 819 European Americans (EA) and 287 African Americans (AA) from the Atherosclerosis Risk in Communities (ARIC) study. Within each population, we selected cases with self-reported gout (n=446 for EA and n=188 for AA) and controls without gout and serum urate levels <5th population-specific percentile (n=372 for EA and n=99 for AA). Variants with minor allele frequency (MAF) >5% were tested in single-SNP analysis using linear regression for associations with serum urate levels and self-reported age of onset of gout. For SNPs with MAF <5%, we conducted gene-based association analysis using the SKAT package. Results. We identified a total of 1009 SNPs with read depth >20 and mean quality score >=30. Of these, 73% have MAF <5%, and 43% were singletons within EA or AA. We confirmed previously reported associations of serum urate levels with nonsynonymous substitutions in ABCG2 (rs2231142, Gln142Lys, p=3.9x10-2) and SLC2A9 (rs16890979, Val253Ile, p=1.8x10-4) in EA. We did not identify associations of novel coding SNPs in either EA or AA. However, two novel missense variants were nominally associated with later age of onset of gout in EA (beta~2 year per copy of allele, p<=3.7x10-2). In addition, a novel promoter SNP (p=4.6x10-2) and a novel 3′ UTR SNP (p=4.3x10-2) were associated with serum urate in AA. Using gene-based analyses, we found rare variants in three of the genes showing suggestive associations (p<0.05) with serum urate levels. Conclusion. Next generation sequencing of 11 targeted genes associated with serum urate levels in a selected subsample of the ARIC cohort identified both common and rare novel variants potentially associated with serum urate levels or age of onset of gout; however, no novel SNPs appeared to account for the original GWAS signals. Further studies in the entire ARIC cohort are needed to clarify the associations of these novel variants.

1933T

Genetic variation in vitamin D pathway genes impacts serum vitamin D deficiency and cardio-metabolic traits in individuals with and without type 2 diabetes. *T.R. Braun, L.F. Been, P.R. Blackett, D.K. Sanghera.* Pediatrics, Section of Genetics, Univ Oklahoma Health Sci Center, Oklahoma City, OK.

Recent studies have examined the physiological functions of vitamin D beyond its well- established role in musculo-skeletal health. In addition to reports of oncologic and immunologic associations, vitamin D deficiency is associated with cardiovascular risk factors including obesity and type 2 diabetes (T2D). The purpose of this investigation was 1) to examine the distribution of vitamin D (25(OH)D) in a Punjabi diabetic cohort from Northern India; and 2) to replicate and validate the role of genetic variation in vitamin D pathway genes associated with serum 25(OH)D levels and cardio-metabolic risk. We measured serum 25(OH)D in 1,765 Punjabi participants (887 T2D cases, 878 NG controls) who were part of the Sikh Diabetes Study and were genotyped using Illumina's 660 Quad BeadChip arrays (Illumina, San Diego, USA).

A total of 76% of individuals were deficient (<50 nmol/L) in vitamin D. A higher percentage of diabetic patients (83%) were vitamin D deficient compared to controls (68%) (p<0.0001). A significantly increased prevalence of vitamin D deficiency was associated with increases in obesity (body mass index (BMI)<23kg/m2, 65%; BMI 23-27.5kg/m2, 75%; and BMI>27.5kg/m2, 81%). Low serum 25(OH)D was significantly associated with T2D (β= -0.41, p=2.8x10-20), elevated fasting glucose (β= -0.021, p=0.019), BMI (β= -0.71, p=1.4x10-7), and systolic blood pressure (β= -1.69, p=0.006). A positive association of 25(OH)D was observed with HOMA-B (β=0.16, p=8.0x10-6), and serum C-peptide (β=0.09, 0.017). Five SNPs in the vitamin D pathway genes IVL, VDR, CYP2R1, GC, and NADSYN1 showed significant associations with 25(OH)D levels. The strongest association with low serum 25(OH)D was observed in the IVL (rs11586313, p=6.1x10-8) gene. Risk alleles on VDR (rs11574143, p=0.03), CYP2R1 (rs12794714, p=0.02), GC (rs2282679, p=0.0019) and NADSYN1 (rs1792226, p=0.03) genes showed moderately significant association with decreased 25(OH)D levels.

This study confirms that low serum 25(OH)D may be a significantly predictor of cardio-metabolic risk. Our findings also suggest that the genetic variation in the vitamin D pathway genes may have a significant contribution in the observed deficiency of vitamin D, which could be synergistically contributing to increased T2D and cardiovascular risk in this population.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1934F

Coding variants at a single multiallelic amino acid position of a HLA class II gene influence follicular lymphoma risk in Europeans and Asians. *J.N. Foo[1], K.E. Smedby[2], P.I.W. de Bakker[3], I.D. Irwan[1], H. Darabi[4], X. Jia[3], L. Padyukov[5], D.E.K. Tan[1], H. Hjalgrim[6], A. Seow[7], K. Humphreys[4], J.J. Liu[1].* 1) Human Genetics, Genome Institute of Singapore, Singapore, Singapore; 2) Department of Medicine, Clinical Epidemiology Unit, Karolinska Institutet, Stockholm, Sweden; 3) Broad Institute of Harvard and MIT, Cambridge, MA, USA; 4) Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden; 5) Rheumatology Unit, Department of Medicine, Karolinska Institutet and Karolinska University Hospital Solna, Stockholm, Sweden; 6) Department of Epidemiology Research, Statens Serum Institut, Copenhagen, Denmark; 7) Saw Swee Hock School of Public Health, National University of Singapore, Singapore, Singapore.

Non-Hodgkin lymphoma represents a diverse group of blood cancers, of which follicular lymphoma (FL) is a common subtype. In a genome-wide association study of 379 FL cases and 791 controls, we identified two independent SNPs (rs10484561 and rs2647012) in the HLA class II region significantly associated with FL risk. Both associations were firmly validated across European datasets and each remained genome-wide significant when we adjusted for the effects of the other; rs10484561:C is a risk allele while rs2647012:T is a protective allele. To fine-map these signals and determine if coding variants in HLA genes are responsible for the associations, we imputed classical HLA alleles and coding variants in our discovery samples using SNP array genotypes. We used a phased reference panel consisting of 2767 unrelated founders of European descent, with genotypes for 2537 SNPs in the HLA region, classical HLA types determined to 4-digit resolution and all variable amino acid positions in the coding regions of the HLA genes encoded for each founder. A total of 263 classical HLA alleles were imputed to 2- and 4-digit resolution and 372 amino acid positions were imputed, out of which 86 were multiallelic. We conducted biallelic trend tests on all the imputed variants, testing every possible combination of 1, 2 or 3 amino acids against the rest at multiallelic sites. From this analysis, the top signal came from a single multiallelic amino acid position within the peptide binding groove of a HLA class II gene. This variant showed stronger association ($P=7.8x10^{-9}$, omnibus $P=9.8x10^{-8}$) than any other HLA variant or SNP and could fully account for both association signals in this region, with rs10484561:C partially tagging the high risk alleles (OR=1.37-1.63) and rs2647012:T partially tagging the low risk alleles (OR=0.61-0.80). There was a 3.8-fold difference in risk between subjects carrying two high risk alleles and those carrying two low risk alleles. We experimentally validated the imputed genotypes and found a 97.3% concordance, demonstrating high accuracy of the imputation. In a similar GWAS and imputation analysis on FL in 363 Singaporean Chinese samples, we observed a similar pattern of association at the same amino acid position. Our current findings suggest that coding variants at a single multiallelic amino acid position of a HLA class II gene may influence FL risk in both Europeans and Asians, and may explain the complex associations revealed by SNP analyses.

1935W

Common polymorphisms of dopamine D2 receptor (DRD2) gene are not associated with adolescent obesity, but with hyperinsulinemia. *N. Col Araz[1], M. Nacak[2], S. Oguzkan Balci[3], N. Benlier[2], S. Pehlivan[3], A. Balat[4], M. Araz[5].* 1) University of Gaziantep, Faculty of Medicine, Pediatrics, Gaziantep, Turkey; 2) University of Gaziantep, Faculty of Medicine, Pharmacology, Gaziantep, Turkey; 3) University of Gaziantep, Faculty of Medicine, Medical Biology and Genetics, Gaziantep, Turkey; 4) University of Gaziantep, Faculty of Medicine, Pediatric Nephrology, Gaziantep, Turkey; 5) University of Gaziantep, Faculty of Medicine, Endocrinology and Metabolism, Gaziantep, Turkey.

Objectives: Obesity is a multifactorial, complex disease of appetite regulation and energy metabolism which is influenced by genetic and environmental factors. Among the genetic factors, there are neuromodulators and multiple neurotransmitter systems which play an important role of feeding behavior. The dorsal striatum plays a n important role on consummatory food reward, and striatal DRD2 receptors has been found to be related to obesity and increased body mass index (BMI). Adolescent obesity may contribute increased prevalence of adult obesity, type 2 diabetes and metabolic syndrome. The aim of this study was to examine the possible relation between common polymorphisms of the DRD2 gene with adolescent obesity and hyperinsulinemia. Methods: Sixty-seven obese adolescents and 100 age- and sex-matched healthy controls were included. Hyperinsulinemia was determined as a basal insulin level ≥15 µIU/mL. Two polymorphisms (Taq1A and Taq1B) in DRD2 gene were genotyped by PCR-RFLP. Results: For DRD2 Taq1A polymorphism, the distribution of A1/A1, A1/A2, and A2/A2 genotypes was 7.5%, 23.9% and 68.6% in obese adolescents compared with 4.0%, 31.0% and 65.0% in controls. The distribution of B1/B1, B1/B2 and B2/B2 genotypes for Taq1B polymorphism was 7.5%, 20.9% and 71.6% for cases and 3.0%, 27.0% and 70.0% for the controls. There were no statistically significant differences in DRD2 Taq1A and DRD2 Taq1B genotypes or allelic frequencies between the obese adolescents and controls ($p>0.05$). Taq1A1/A1 genotype was associated with hyperinsulinemia in adolescent obese subjects ($p<0.05$). Conclusion: No association was found between the Taq1A and Taq1B polymorphisms in the DRD2 gene and adolescent obesity. But hyperinsulinemia was associated with A1/A1 genotype. Preliminary results of the ongoing study are presented here and these findings should be supported in larger obese adolescents groups.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1936T

Association analysis of chromosome 3q markers in nonsyndromic cleft lip/palate. *A. Letra[1], M. Cooper[2], T. McHenry[2], E. Czeizel[3], F.W.B. Deleyiannis[4], L. Ma[5], E.E. Castilla[6], F. Poletta[7], L.L. Field[8], A.R. Vieira[2], R.M. Silva[1], M.L. Marazita[2].* 1) Dept Endodontics and Pediatric Research Center, Univ Texas Health Science Center at Houston, Houston, TX; 2) University of Pittsburgh School of Dental Medicine and Center for Craniofacial and Dental Genetics, Pittsburgh, PA; 3) Center for the Control of Hereditary Diseases, Budapest, Hungary; 4) Departments fo Pediatric Plastric Surgery, Surgery and Otolaryngology, The Children's Hospital and University of Colorado School of Medicine; 5) School of Stomatology, Beijing University, Beijing, China; 6) ECLAMC at INAGEMP-CNPq (National Institute of Population Medical Genetics) in the Department of Genetics, Oswaldo Cruz Foundation, Rio de Janeiro, RJ, Brazil; 7) ECLAMC (Latin American Collaborative Study of Congenital Malformations) at CEMIC (Center for Medical Education and Clinical Research), Buenos Aires, Argentina; 8) Department of Medical Genetics, University of British Columbia, Vancouver, BC, Canada.

Nonsyndromic cleft lip/palate (CL/P) is a common birth defect of complex etiology involving the interplay of genetic predisposition and environmental exposures. Several genes, including but not limited to MSX1, IRF6, CRISPLD2, FOXE1, and members of the WNT and FGF gene families, and additional loci on chromosomes 1p22, 8q24, 10q25.3, 17q22, and 20q12 have been implicated in the etiology of oral clefts. A previous genome wide linkage scan also identified a region on chromosome 3q27-28 with a significant multipoint HLOD peak of ~4.13.0 (α=0.34), warranting further studies in additional populations. Here, we performed a family-based association study of markers in chromosome 3q and CL/P in multiple populations. Data and samples from 885 families ascertained through a proband with CL/P families (1,498 affecteds, 3,468 unaffecteds) were ascertained through the University of Pittsburgh Oral-Facial Cleft study (POFC), including sites in the United States, Argentina, Guatemala, Spain, Hungary, Turkey, Philippines, China, and India. We investigated the association of 996 SNPs comprising 12 genes intertwined with 10 regions of no known genes in chromosome 3q for association with CL/P in our families. Alleles at each SNP were tested for association with CL/P using the Family-based Association Test (FBAT). We performed analyses for each population individually, all populations combined, and in the following groups: Caucasians, Latin Americans, and Asians. Haplotype analysis was also performed. Overall, the most significant association results in all populations and Caucasians combined were seen with markers in LPP, TPRG1, TP63 and FGF12 (P≤0.009) although none reached genome-wide significance. Analysis of individual populations showed association of SNPs in TP63 (rs10155037, P=0.000004; rs16864812, P=0.00001) and in CCDC50 (rs293813, P=0.0001) with CL/P in India, whereas a SNP in TMEM207 (rs10513852, P=0.0000003) was associated in South Americans. Under an additive model, an intergenic marker (rs11706540) between TP63 and LEPREL1 showed nominal association with CL/P in Caucasians (P=0.009). Marker haplotypes in the IL1RAP gene showed association with CL/P in all populations (P=0.005). Under a recessive model, marker haplotypes neighboring TPRGR1 and TP63 showed association in Caucasians (P=0.003). In summary, several markers in chromosome 3q had suggestive association with CL/P that warrant a more focused search for candidate loci in these regions.

1937F

Genotype-phenotype correlations of facial shape and asymmetry in unaffected relatives of children with non-syndromic cleft lip/palate and controls. *S. Miller[1], N. Nidey[4], S. Weinberg[3], M.L Marazita[3], J.C Murray[4], G.L. Wehby[2], L.M Moreno[1].* 1) Department of Orthodontics-Dows Institute, College of Dentistry, University of Iowa, IA; 2) Department of Health Management and Policy, School of Public Health, University of Iowa, IA; 3) Center for Craniofacial and Dental Genetics, Department of Oral Biology, University of Pittsburgh, PA; 4) Department of Pediatrics, University of Iowa, IA.

Individuals with Nonsyndromic cleft lip with or without cleft palate (NSCL/P) and their unaffected relatives often present with features of distinct facial morphology related to the phenotypic spectrum of orofacial clefting. The genetic etiology underlying this phenotypic variation is poorly understood. This study examines whether candidate genes associated with NSCL/P and left-right asymmetry are also correlated with the cleft risk related facial variation present in non-affected parents and siblings of children with NSCL/P. Relatives of individuals with NSCL/P (Cases, N= 188, unaffected parents (N=119), and siblings (n=69) and controls (N=194) with no family history of NSCL/P were genotyped for 20 SNPs across 13 candidate genes for NSCL/P (PAX7, ABCA4-ARHGAP29, IRF6, MSX1, PiTX2, 8q, FOXE1, TGFB3, MAFB) and left/right asymmetry (LEFTY1, LEFTY2, ISL1, and SNAI1). 3D facial shape plus asymmetry phenotypes including fluctuating asymmetry (FA), were derived from 32 3D coordinate landmarks digitized in 3D facial images. Patterns of facial shape and asymmetry were obtained separately for each gender, via relative warps analyses (RWA) and Procrustes ANOVA of the coordinate data. Phenotypes that differed between cases and controls (p<0.07) were subjected to multiple regression adjusting for age on the genotypes obtained. Results for the mother subsample (n=163), suggest that cases have larger faces, wider philtrums, more upper face retrusion, increased lower face height and prognathism and higher FA than controls (P<0.05). ABCA4-ARHGAP29, SNAI1 and LEFTY1 were correlated (p<0.05) with shape variation and FA respectively. FA was more significant (p=0.0063) in female relatives of children with unilateral clefts than controls. The father subsample (N=80) indicated that cases have increased lower face height and prognathism, and their philtrums deviate to the left side compared to controls (P=0.07), with significant correlations found for IRF6 and MAFB (p<0.05). Female siblings (N=68) also showed more upper face retrusion in the cases than in controls and suggestive correlations (p<0.07) with FOXE1, MSX1 and TGFB3. Lastly, male siblings (N=71) showed more retrusive upper faces, and a larger deviation of the philtrum to the left than controls, together with significant (p<0.05) correlations for IRF6 and LEFTY2. This study highlights the importance of genotype-phenotype correlations studies in discerning the genetic etiology of NSCL/P. DE016148, CDC5R01DD000295.

1938W

Cannabinoid Receptor-1 (CNR-1) gene 1359G/A polymorphism is related with childhood obesity, but not with insulin resistance. *M. Nacak[1], N. Col Araz[2], S. Oguzkan Balci[3], N. Benlier[1], S. Pehlivan[3], A. Balat[4], M. Araz[5].* 1) University of Gaziantep, Faculty of Medicine, Pharmacology, Gaziantep, Turkey; 2) University of Gaziantep, Faculty of Medicine, Pediatrics, Gaziantep, Turkey; 3) University of Gaziantep, Faculty of Medicine, Medical Biology and Genetics, Gaziantep, Turkey; 4) University of Gaziantep, Faculty of Medicine, Pediatric Nephrology, Gaziantep, Turkey; 5) University of Gaziantep, Faculty of Medicine, Endocrinology and Metabolism, Gaziantep, Turkey.

Objectives: Obesity is becoming global health problem in the world. The cannabinoid receptor (CNR) and its endogenous ligands, the endocannabinoids, are involved in regulation of feeding behavior and energy balance by stimulating appetite, increasing body weight, and implicated in human obesity. Furthermore, activation of the CNR1 may play a role in the development of insulin resistance. The aim of this study was to investigate the influence of the Cannabinoid reseptor-1 (1359G/A) gene polymorphisms on chilhood obesity and insulin resistance. Methods: A hundred obese subjects and 100 age-and sex-matched healthy controls were analyzed for CNR1 1359G/A polymorphisms. Genotyping were performed by PCR and/or RFLP. Homeostasis model assessment (HOMA) was used to consider insulin resistance. Results: The frequency of the A allele of the CNR1 1359G/A polymorphism was significantly higher in obese children than in control subjects (21.0 % versus 13.0 %, p=0.0166). The frequency of genotypes AG and GG of the CNR1 1359G/A SNP was statistically significant difference between the obese and control subjects (for AG: 34.0% versus 22.0%, p= 0.0294; for GG: 62.0% versus 76.0%, p=0.0162, respectively). However, there was no significant difference in the AA genotype distribution among patients and controls (4.0% versus 2.0%, p>0.05). The observed genotype counts were not deviated significantly from those expected according to the HWE. No relationship was found between CNR1 genotypes or alleles and insulin resistance in patients (p<0.05) Conclusion: 1359G/A polymorphism of the CNR1 gene may have a role in obesity. However, further studies will clarify the exact role of this gene on childhood obesity and insulin resistance.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1939T

Altered expression but no genetic association of the epidermal transglutaminases in atopic dermatitis. *M.CG. Winge[1;2], A. Lieden[1], A. Sääf[1], I. Kockum[3], E. Ekelund[1], E. Rodriguez[4], T. Hoppe[5], R. Fölster-Holst[4], A. Franke[6], M. Tengvall-Linder[7], H. Baurecht[4], S. Weidinger[4], CF. Wahlgren[2], M. Nordenskjöld[1], M. Bradley[1;2].* 1) Dept of Mol Med & Surgery, Karolinska Institutet, Stockholm, Stockholm, Sweden; 2) Dermatology Unit, Department of Medicine Solna, Karolinska University Hospital, Stockholm, Sweden; 3) Department of Clinical Neurosciences, Karolinska Institutet, Stockholm, Sweden; 4) Department of Dermatology, University of Kiel, Kiel, Germany; 5) Department of Medical Sciences, Dermatology and Venereology, Uppsala University, SE-75185 Uppsala, Sweden; 6) Institute of Clinical Molecular Biology, Christian- Albrechts-Universität zu Kiel, Kiel, Germany; 7) Clinical Immunology and Allergy, Department of Medicine Solna, Karolinska Institutet, Stockholm, Sweden.

Background: Atopic dermatitis (AD) is a common chronic inflammatory skin disorder where epidermal barrier dysfunction is a major factor and the identification of susceptibility genes in barrier dysfunction is therefore of importance. The epidermal transglutaminases (TGM1, TGM3 and TGM5) encodes essential cross-linking enzymes in the epidermis. Objective: In this study we investigated the expression in AD patients and healthy controls, and whether genetic variability at these loci contributes to AD susceptibility. Methods: Gene expression of the TGM1, TGM3 and TGM5 gene was determined by relative quantification with Real Time PCR (qRT-PCR) and in silico analysis of expression data compared to trans-epidermal water loss (TEWL), filaggrin mutations, pH and disease severity (SCORAD). Immunohistochemical (IHC) analysis was performed to detect TG1, TG3 and TG5 protein expression in the skin of patients and healthy controls. Forty-seven single nucleotide polymorphisms (SNPs) in the TGM1, TGM3 and TGM5 gene region were tested for genetic association with AD using a pedigree disequilibrium test (PDT) in a Swedish material consisting of 1753 individuals from 539 families and a candidate SNP (rs941505) was also tested in silico in a German case-control material consisting of 533 AD cases and 1996 controls. Results: increased TGM1 expression was detected in silico in non-lesional AD skin compared to healthy skin, and in skin from AD patients with a severe phenotype whereas no significant difference was detected for TGM3. Significantly increased TGM3 expression was however detected with qRT-PCR in lesional skin, together with a non-significant increase of TGM1 transcript levels in both non-lesional and lesional skin. Increased protein expression of both TG1 and TG3 was estimated in lesional skin. No differences in TGM1, TGM3 or TGM5 expression were found in relation to TEWL, pH or filaggrin genotype. Although the PDT analysis identified a significant association between the TGM1 SNP rs941505 and AD with allergen-specific IgE in the Swedish AD family material, the association was not replicated in a German case-control material. Conclusion: Expression analysis links the TGM1 and TGM3 gene to the manifestation of AD and may reflect ongoing inflammation and an altered barrier function. Furthermore, the results from the genetic analysis show that genetic variation in the epidermal transglutaminases is not an important modifier of AD susceptibility.

1940F

Genetic variation and epigenetic modification of *TACR1* gene are associated with postoperative nausea and vomiting in surgical patients. *T. Hayase, S. Sugino, M. Yamakage.* Department of Anesthesiology, Sapporo Medical University School of Medicine, Sapporo, Hokkaido, Japan.

Background: Postoperative nausea and vomiting (PONV) is a frequent and debilitating complication in surgical patients, but its occurrence and intensity vary individually. Although this interindividual difference could be influenced by genetic and/or environmental factors, the underlying mechanisms have not been fully elucidated. Neurokinin 1 receptor plays a major role in drug-induced emesis in the central nervous system, and recent studies have suggested an association with PONV. The aim of the present study was to determine whether genetic variations and epigenetic modifications of the *TACR1* gene, which codes neurokinin 1 receptor, are associated with the occurrence and intensity of PONV. **Methods:** Seventy-seven Japanese female patients undergoing gynecological surgery were enrolled in this study. The occurrence of PONV was recorded for 24 hrs after surgery and defined as ≥3 instances of retching and/or vomiting over a period of 24 hrs or nausea persisting for at least 12 hrs after surgery. The intensity of PONV was assessed using a 0 ~ 100 visual analogue scale (VAS). We selected seven tag SNPs of the *TACR1* gene for haplotype analysis (MAF>0.1, r^2<0.8). Genotyping for SNPs was carried out by a PCR-direct sequencing method. In a separate study, we investigated DNA methylation at 11 CpG sites in the -100/+6-bp upstream region of the *TACR1* gene using a methylation-sensitive high-resolution melting protocol as previously described by Wojdacz *et al*. The primer pair was designed using MethylPrimer Express software and amplicons were generated from bisulfite-converted DNA of the patients. Haplotype data were analyzed using a χ^2 test or Fisher exact test as appropriate. Methylation data were analyzed by Spearman's rank correlation coefficient. **Results:** The four most common haplotypes with a frequency of >5% were TAACATA (36%), TAATAGA (14%), TAATATA (7%), and TAATATC (6%). The frequency of PONV in patients with the TAACATA haplotype was significantly higher than that in patients with other haplotypes (P=0.04). Methylation level of the promoter region was inversely correlated with VAS scores of PONV (r = -0.17, P < 0.01). **Conclusions:** The TAACATA haplotype of the *TACR1* gene is associated with the occurrence and severity of PONV. *TACR1* promoter methylation is related to attenuation of the severity of PONV. The results of this study indicate that genetic polymorphisms and epigenetic regulations of the *TACR1* gene might be involved in PONV after surgery.

1941W

Identification of a novel 1p31 locus associated with the combined asthma plus allergic rhinitis phenotype through positional cloning in asthma-ascertained families. *M-H. Dizier[1], P. Margaritte-Jeannin[1], A-M. Madore[2], I. Annesi-Maesano[3], J. Just[4], F. Kauffmann[5], C. Laprise[2], M. Lathrop[6], E. Bouzigon[1], F. Demenais[1].* 1) INSERM U946, Paris, paris, France; 2) Université du Québec, Chicoutimi, Canada; 3) Inserm U707, Paris, France; 4) Centre de diagnostic et traitement de l'asthme, Hôpital Trousseau, Paris, France; 5) Inserm U1018, Paris, France; 6) CEA-CNG, Evry, France.

A genome-wide linkage scan conducted in the French Epidemiological study on the Genetics and Environment of Asthma (EGEA) showed strong evidence of linkage of 1p31 to asthma, allergic rhinitis (AR) and even more to the combined asthma plus AR phenotype (Dizier el al, 2005, 2007). Our purpose was to conduct fine-scale mapping of the 1p31 linkage region using a panel of 2062SNPs (spanning 10 Mb) in 388 EGEA families to identify the genetic variants associated with asthma plus AR. Association analyses were first carried out using the family-based association method (FBAT). We then applied logistic regression to the genetic variants showing significant association with asthma plus AR, to validate the FBAT results. To investigate whether the association was specific to the combined asthma plus AR phenotype, we tested for heterogeneity of association according disease status defined by the presence of the two diseases (asthma plus AR) versus the presence of only one disease ('asthma only' or `AR only') using FBAT. We found significant association of asthma plus AR to one 1p31 SNP using FBAT (P=2.10-5, which is below the multiple testing-corrected threshold of 5x10-5). This result was confirmed by logistic regression (P=2.10-4). That SNP showed significant heterogeneity of association with asthma plus AR vs. asthma alone or AR alone (P=0.02; moreover, there was no evidence for association with asthma alone or AR alone). We replicated our finding, in 253 French Canadian families from the Saguenay-Lac-Saint-Jean study, which, similarly to EGEA were ascertained through asthmatic subjects. Meta-analysis of EGEA and SLSJ results (P=0.03 for association of asthma plus AR to the 1p31 SNP using FBAT) showed strong improvement in both the evidence for association with asthma plus AR (P=4x10-6) and in the evidence for heterogeneity of association according to disease status (P=0.007 for association with asthma plus AR vs asthma alone or AR alone). Our study shows that the co-morbidity, asthma plus AR, is a phenotypic entity, that is influenced by genetic factors, including the 1p31 locus, which are distinct from those underlying asthma alone or AR alone. Although further confirmation of our results is needed, it is of note that the SNP associated with asthma plus AR belongs to a gene involved in the replication of adenovirus type 2, which makes it a promising candidate gene.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1942T

Dense genotyping of six atopic dermatitis and 180 autoimmune risk loci in 2,425 atopic dermatitis patients. *D. Ellinghaus[1], E. Rodríguez[2], H. Baurecht[2], J. Esparza-Gordillo[3,4], Y.-A. Lee[3,4], S. Cichon[5], C. Gieger[6], H.-E. Wichmann[7,8,9], R. Duerr[10,11], C. Bü̈ning[12], S. Brand[13], S. Schreiber[1,14,15], S. Weidinger[2], A. Franke[1].* 1) Institute of Clinical Molecular Biology, Christian-Albrechts-University, Campus Kiel, Kiel, Germany; 2) Department of Dermatology, Allergology, and Venerology, University Hospital Schleswig-Holstein, Campus Kiel, Kiel, Germany; 3) Pediatric Pneumology and Immunology, Charité, Universitätsmedizin Berlin, Berlin, Germany; 4) Max-Delbrück-Centrum (MDC) for Molecular Medicine, Berlin-Buch, Germany; 5) Department of Genomics, Life & Brain Center, University Hospital Bonn, Bonn, Germany; 6) Institute of Genetic Epidemiology, Helmholtz Centre Munich, German Research Center for Environmental Health, Neuherberg, Germany; 7) Institute of Epidemiology I, Helmholtz Centre Munich, German Research Center for Environmental Health, Neuherberg, Germany; 8) Institute of Medical Informatics, Biometry and Epidemiology, Ludwig-Maximilians-University, Munich, Germany; 9) Klinikum Grosshadern, Munich, Germany; 10) University of Pittsburgh School of Medicine, Division of Gastroenterology, Hepatology and Nutrition, Department of Medicine, Pittsburgh, Pennsylvania, USA; 11) University of Pittsburgh Graduate School of Public Health, Department of Human Genetics, Pittsburgh, Pennsylvania, USA; 12) Department of Gastroenterology, Hepatology and Endocrinology, Charité, Campus Mitte, Berlin, Germany; 13) Department of Medicine II - Grosshadern, Ludwig-Maximilians-University (LMU), Munich, Germany; 14) PopGen Biobank, University Hospital Schleswig-Holstein, Kiel, Germany; 15) Department of Internal Medicine, University Hospital Schleswig-Holstein, Kiel, Germany.

Atopic dermatitis (AD), or eczema, is one of the most common chronic inflammatory skin diseases, with a polygenic, multifactorial nature. To date, genome-wide association studies have established six susceptibility loci with genome-wide significance ($P < 5\times10^{-8}$). However, the causal variation at these loci remains unknown. To better define risk variants and identify additional susceptibility loci previously implicated in other autoimmune (AI) diseases, we performed a fine-mapping and association study using 2,425 German AD cases and 5,449 German population controls. All samples were genotyped on the ImmunoChip, a custom Illumina Infinium High-Density array containing 196,524 polymorphisms (195,806 SNPs and 718 small insertions/deletions) across 186 distinct AI risk loci, including the six known AD susceptibility loci and the HLA region. At these loci, the array contains all known SNPs in the dbSNP database, from the 1000 Genomes Project (release of February 2012), and from other AI disease resequencing efforts, therefore providing a powerful means of fine-mapping known AI loci. Following quality control, 128,830 polymorphic markers (SNPs with minor allele frequency > 1%) were available for association analysis, with an individual call rate > 99.9%. Genotypes were called using Illumina's GenomeStudio GenTrain 2.0 algorithm. Principal component analysis revealed no marked differences in ancestry between cases and controls. We observed that 132 SNPs within non-HLA risk loci reached genome-wide significance ($P < 5\times10^{-8}$) in the screening phase. For each locus, the possibility of the existence of multiple independent association signals will be investigated. In addition, we selected the most strongly associated SNPs (n=39) with $P < 10^{-4}$ from each associated locus for replication analysis in independent case-control collections. The final results will be presented at the conference.

1943F

Identification of functional variants in *FOXO3A*, a confirmed candidate gene influencing human longevity: Genetic investigation requires special attention to sequence homology with *FOXO3B*. *F. Flachsbart[1], L. Gentschew[1], C. Däumer[2], N. Badarinarayan[1], A. Caliebe[2], M. Krawczak[2,4], S. Schreiber[1,3,4], A. Nebel[1].* 1) Institute of Clinical Molecular Biology, Christian-Albrechts-University, Kiel, Schleswig-Holstein, Germany; 2) Institute of Medical Informatics and Statistics, Christian-Albrechts-University, Kiel, Germany; 3) Clinic for Internal Medicine I, University Hospital Schleswig-Holstein Kiel, Germany; 4) Popgen Biobank, Christian-Albrechts-University, Kiel, Germany.

Human longevity is influenced by multiple genetic and environmental factors. Approximately 30% of the overall variation in adult lifespan is because of genetic variation that becomes particularly important for survival at advanced age. Up to now, variation in only two genes has been identified, which has an effect on longevity in various populations: the apolipoprotein E gene (*APOE*) and the forkhead box O3A gene (*FOXO3A*). For *FOXO3A* the underlying molecular mechanisms still remain to be elucidated, because most of the longevity-associated single nucleotide polymorphisms (SNPs) analyzed are located in intronic regions and the “functionally relevant” SNP has yet to be identified. To explore potential unknown genetic variants both next generation and Sanger sequencing are being applied. Functional effects of all identified rare and common SNPs are evaluated by in silico analyses. Variants of interest are genotyped in our extensive German longevity sample (~1400 long-lived cases and 1100 younger controls). In our study we could recently demonstrate that the genetic investigation of *FOXO3A* is greatly hampered by the fact that the exonic regions have 99% sequence homology with the *FOXO3B* pseudogene. If unaccounted for, this high degree of homology can cause serious genotyping or sequencing errors. Here, we present an experimental set-up that allows reliable data generation for the highly homologous regions and that can be used for evaluation of assay specificity. Using this design, we could exemplarily show *FOXO3A*-specific results for two SNPs (rs4945816 and rs4946936) that are significantly associated with longevity in our German longevity sample (Peach = 0.0008). Because both SNPs are located in the 3′ untranslated region of *FOXO3A*, they could be of functional relevance for the longevity phenotype. Our experimental set-up can be used for reliable and reproducible data generation for further sequencing and genotyping studies of *FOXO3A* with the aim of discovering new SNPs of functional relevance. In the following, various in vitro and ex vivo assays will be applied for the functional characterization (gene expression profiles, mRNA stabilization, translational efficiency, splicing events) of associated variants. Subsequently, the functional assessment of the *FOXO3A* gene region could serve as the basis for future treatment strategies to achieve healthy ageing in the elderly.

1944W

***PDGFRa* Mutations in Humans with Isolated Cleft Palate.** *V. Shotelersuk[1,2], S. Rattanasopha[1,2], S. Tongkobpetch[1,2], C. Srichomthong[1,2], P. Siriwan[3], K. Suphapeetiporn[1,2].* 1) Center of Excellence for Medical Genetics, Department of Pediatrics, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand; 2) Molecular Genetics Diagnostic Center, King Chulalongkorn Memorial Hospital, Thai Red Cross, Bangkok, Thailand; 3) Division of Plastic Surgery, Department of Surgery, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand.

Objective: *Pdfgra* knock-out zebrafish showed cleft palate (CP). We attempted to identify whether *PDGFRa* was responsible for isolated CP in humans. **Materials and Methods:** We recruited 102 patients with CP under the auspices of the Thai Red Cross, from 33 medical centers throughout Thailand. The entire coding regions and the 3′ UTR of the *PDGFRa* and *miRNA140* were PCR-amplified and sequenced. PCR-RFLP was used to confirm the variants identified in the patients and 500 unaffected ethic-matched controls. We then determined the pathogenicity of the mutations found in the *PDGFRa* 3′ UTR by luciferase reporter system. **Results:** PCR-sequencing identified seven novel single base-pair substitutions in the *PDGFRa* in 9/102 patients (8.8%) compared to 5/500 ethnic-matched unaffected controls (1%) (the two-tailed p value < 0.0001). Of these seven, four were missense mutations in the coding regions and three in the 3′UTR. Frequencies of four changes (three in coding, one in 3′UTR) were statistically different from those of controls (p value < 0.05). The c.;k434G>A was identified in 1/102 cases and 0/500 controls. This position is conserved in primates and located 10 bp away from a predicted binding site for the human microRNA, miR-140. Luciferase assay revealed that, in the presence of miR-140, the c.;k434G>A significantly repressed luciferase activity compared to that of the wild type. **Conclusion:** For the first time, we provided strong evidence supporting a role of *PDGFRa* in human palatal development.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1945T

Copy-number variation in chromosome 5q31-33 is associated with Graves' disease susceptibility. *X. Chu, M. Shen, Y. Bai, F. Xie, W. Huang.* Department of Genetics, Chinese Human Genome Ctr, Shanghai, China.

Graves' disease is a common multifactorial disease with familial aggregation. The common variants identified using single-nucleotide polymorphisms (SNPs) as markers only explained a modest proportion of the known heritability of Graves' disease, and could not give an explanation for familial aggregation in this disease. We and Japanese researchers have previously identified chromosomal region 5q31-33 for contribution to the genetic susceptibility to Graves' disease in East-Asian populations from linkage analysis. These results implicated high-penetrance variants associated with Graves' disease might be located in this region. Our further candidate gene analysis in 5q31-33 and recent genome-wide association analysis (GWAS) revealed SNPs associated with Graves' disease, but these low-penetrance variants could not account for the linkage signal in this region. Copy-number variation (CNV) is another type of variation widely existed in human genome and hereditary CNVs might be the genetic basis of familial aggregation in diseases. In our previously GWAS for Graves' disease, we genotyped 1,536 Graves' disease cases and 1,516 sex-matched controls using Illumina Human 660-Quad BeadChips, which included 657,314 SNP markers and 64,527 CNV markers. In this study, we aimed to identify Graves' disease associated CNVs with high-penetrance in the linkage region 5q31-33. CNVs were called by CNVPartition v1.2.1 and PennCNV on genotyping signal intensity data. CNVs were included for further association analysis only if they were detected by both software packages. We compared the frequency of CNVs between cases and controls using Plink. The potentially disease associated CNVs were further validated by quantitative PCR (qPCR), and TaqMan probes were custom-designed using Primer Express 3.0 (Applied Biosystems). We found a rare CNV at 5q31-33 were associated with Graves' disease susceptibility (P = 0.0034). The majortity cases with deletion were females with family history of Graves' disease. We would perform further validation in Graves' disease pedgrees in our further study and try to understand the genetic mechanisms underlying familial aggregation of Graves' disease.

1946F

Genetic analysis of HLA-G gene: association study in a Brazilian cohort relating the 14bp INDEL frequencies in a psoriasis case/control study. *G. Debortoli[1], B.S. Almeida[1], A. Firmino[1], D.H. Nunes[2], M.M. Sens[2], V.P.M. da Silva[2], A.R. Marrero[1], Y.C.N. Muniz[1], I.R. Souza[1].* 1) LAPOGE, Federal University of Santa Catarina, Florianopolis, Santa Catarina, Brazil; 2) HU, Federal University of Santa Catarina, Florianopolis, Santa Catarina, Brazil.

Psoriasis is an autoimmune disease which affects the skin and articulations. The HLA-G protein can influence the disease course, because your expression can act as an immune surveillance mechanism in tissue protection against autoimmune inflammatory responses. Researchers have reported in literature the InDel 14 bp polymorphism located in the 3′UTR of the HLA-G gene, where appears to influence the expression, generating an mRNA more stable. The main objective was the genotyping of 100 patients and 100 controls (matched for age and gender), analyzing the influence of the Indel 14pb in the psoriasis manifestation. We conducted a survey of risk factors for gender, age of manifestation onset, family history, smoking, and involvement of the joints. The DNA samples were extracted from whole blood, and the genotyping was carried out by PCR, whereas visualization by PAGE and stained with silver nitrate. The results showed that 57 of patients were women, 44 individuals had family history, 43 manifested the disease before the age of 30, 46 were smokers, and 16, expressed psoriatic arthritis. Allele frequencies were calculated resulting in a frequency of 0.25 InDel 14 bp in patients with psoriatic arthritis and 0.50 in patients without psoriatic arthritis. The results was not significant, but corroborates published studies, which show a protective trend of InDel 14 bp polymorphism (mRNA more stable), in this particular disease. For a better understanding of the gene influence and the role in the immune system in the development of psoriasis, the sample number will be raised as well as other markers of influence on the stability of mRNA (Financial support: PNPD CAPES, FAPESC and CNPq).

1947W

Lung eQTL in the major histocompatibility complex. *M. Lamontagne[1], C. Couture[1], M. Laviolette[1], Y. Bossé[1,2], The Merck-Laval-UBC-Groningen Lung eQTL consortium.* 1) Centre de recherche Institut universitaire de cardiologie et de pneumologie de Québec, Québec, Canada; 2) Department of Molecular Medicine, Laval University, Quebec, Canada.

The major histocompatibility complex (MHC) is a key region in the human genome for the immune system. The extended MHC in humans (xMHC) covers 7.6 Mb on the short arm of chromosome 6. Variants in MHC genes are associated with hundreds of diseases, particularly autoimmune diseases. The genetic mechanisms regulating the MHC complex are still unclear. The lungs are extensively exposed to environmental pathogens and are the most common route of infections. The MHC is thus a critical genomic region to control infections in the lung such as pneumonia and tuberculosis. The MHC is also known to play a critical role in the development of chronic lung diseases (e.g. COPD and asthma). The goal of this study is to further our understanding of the mechanisms regulating this important region by using lung expression Quantitative Trait Loci (eQTL). Genome-wide gene expression profiles of 500 non-tumor lung specimens were obtained from patients undergoing lung surgery. Blood-DNA from the same patients were genotyped for 1,2 million SNPs. Following genotyping and gene expression quality control filters, 409 samples were analyzed. A total of 3948 cis-eQTLs and 507 trans-eQTLs were identified in the xMHC. This represents 13 fold more eQTLs than any other genomic region of the same size. The expression of a number of genes were regulated by multiple independent SNPs. Notably, the expression of HLA-DRB5 was regulated by more than 20 SNPs. rs35366052 was the most significant SNP regulating the expression of HLA-DRB5 (p = 2,488E-297). Regulatory genetic variants were also found to be pleiotropic. For example, rs2647012 located in the HLA-DQ loci was significantly associated with the expression of 6 genes. This study is important in increasing our understanding of the genetic mechanisms regulating the MHC region and its role in pulmonary infections and chronic respiratory diseases.

1948T

Rare TLR pathways gene variants and susceptibility to invasive pneumococcal disease: a population-based study. *K.S. Elliott[1], M.K. Ellis[1,2,3], A. Rautanen[1], M.A. Rivas[1], D.W. Crook[4], A.V.S. Hill[1], S.J. Chapman[1,5].* 1) Wellcome Trust Centre for Human Genetics, University of Oxford, United Kingdom; 2) Queensland Institute of Medical Research, Brisbane, Australia; 3) Australia Centenary Institute, Sydney, Australia; 4) Department of Microbiology, John Radcliffe Hospital, Oxford, UK; 5) Respiratory Medicine, Churchill Hospital Site, Oxford Radcliffe Hospital, Oxford, UK.

Streptococcus pneumoniae (the pneumococcus) is the most common cause of community-acquired pneumonia in Europe and the United States and a leading cause of death in children worldwide. In addition to pneumonia, infection may be manifested as invasive pneumococcal disease (IPD), defined by the isolation of *Streptococcus pneumoniae* from a normally sterile site. Host genetic variation is widely believed to influence susceptibility to IPD, although only a limited number of common polymorphisms have been identified in association with IPD to date. Recent study of very rare human primary immunodeficiencies associated with selective predisposition to IPD has identified causative mutations within genes in the Toll-like receptor (TLR) pathway which result in impaired NF-κB activation. A key question is the relative contribution of multiple rare variants in these genes to IPD susceptibility at the population level. We therefore performed sequencing of the genes *MYD88*, *IKBKG*, *IRAK4*, and *TIRAP* in a population-based case-control association study of 164 adults with IPD and 164 controls. For both cases and controls, individuals were divided into 4 pools containing 41 genomic DNAs each. From these pools 20 amplicons (median length 4kb) were amplified using long range PCR and sequenced using next generation sequencing. We found several variants to be present in one or more case pools and absent from control pools, some of which are novel (defined by absence from dbSNP B135). Using syzygy software we estimated that case pools contained one copy of the missense variant S229L and a single copy of a novel splice variant in *MYD88*; one copy of a novel missense variant within *IRAK4* and 2 copies of a novel missense variant and one copy of the published loss of function variant rs8177400 which encodes D96N in *TIRAP*. Direct sequencing of these variants in each individual in the case and control pools will provide accurate genotypes for further analysis. These rare TLR pathway gene mutations in adults with IPD, suggest that at least a proportion of adults with this major infectious disease phenotype may have a previously unidentified, selective primary immunodeficiency state. This result provides an insight into the genetic architecture of infectious diseases as well as the nature of inherited immunodeficiency in humans.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1949F

Association of VPS26 variants with susceptibility to infection by Leishmania in Brazil, Iran and Sudan. *A. Romano[1,2], H. Dessein[1,2], A. Muller[1,2], F. Santoro[1,2], A. Salhi[1,2], S. Rafati[3], S. El Saafi[4], JP. Gorvel[5], L. Argiro[1,2], A. Dessein[1,2].* 1) INSERM, UMR906, Marseille, France; 2) Aix Marseille University, Faculty of Medicine, Marseille, France; 3) Pasteur Institute of Teheran, Teheran, Iran; 4) University of Karthoum, Karthoum, Sudan; 5) Aix Marseille University, CIML, Marseille, France.

Leishmaniasis are a group of parasitic diseases that affect millions of people in southern countries. Leishmania parasites multiply in macrophages and cause a wide spectrum of diseases including cutaneous lesions (CL) and lethal visceral disease (VL). The reasons why certain patients remain asymptomatic following infection whereas others develop severe lesions are unclear. We have performed a Genome Wide Linkage Study in affected siblings from a Brazilian sample living in an endemic region for L.braziliensis. This GWLS indicated a susceptibility locus for cutaneous lesions on chromosome 10, within the 10q21-q23 region. Fine mapping with additional families yield a maximum MLB lod-score of 2.39 for D10S522. Analysis of candidate genes in this region revealed a protein of the retromer complex VPS26 (Vacuolar Protein Sorting 26), involved in the maturation of endosome / lysosome. Thirty SNPs localised in VPS26 and surrounding genes were genotyped in 84 families corresponding to 141 trio (at least 1 affected child and his two parents). Most SNPs in VPS26 and in the flanking region were comprised in seven correlation (r2>0.8) bins. Transmission Disequilibrium Test analysis of the data indicated that SNPs of 3 correlation bins were significantly ($p<0.01$) associated with CL in the Brazilian sample. In addition, we found that several SNPs associated with CL in Brazil, were also associated with cutaneous lesions caused by L.tropica in an Iranian sample and with visceral leishmaniasis caused by L.donovani in a Sudanese sample. Immunostainning was done for LPG, Lamp1 and VPS26 on mouse macrophage cell line RAW infected by leishmania. We observed that VPS26 colocalised with LPG positive-leishmania very rapidly from the beginning of the phagocytosis. In addition, VPS26 staining persisted up to the late stage when leishmania containing vacuoles expressed LAMP-1. Another protein members of the retromer complex (VPS35) also gave a positive signal on the phagocyted leishmania. We hypothesize that polymorphisms in VSP26 gene modify expression of the retromer complex on the leishmania-phagocytic vacuoles resulting in an aggravation of the infection. Ongoing experiments are conducted to explore this hypothesis.

1950W

MTHFR 677 C>T polymorphism does not influence the risk of polycystic ovarian syndrome (PCOS). *S. Rajender[1], J. Carlus[2], K. Singh[3], K. Thangaraj[2].* 1) Endocrinology, Central Drug Research Institute, Lucknow, UP, India; 2) Centre for Cellular and Molecular Biology, Hyderabad, India; 3) Banaras Hindu University, Varanasi, India.

Statement of purpose: Polycystic ovarian syndrome (PCOS) is one of the most common forms of ovarian dysfunction accounting for a large proportion of infertile women. Sufficient folate intake and an optimum efficiency of folate pathway are considered essential for normal ovarian function. Elevated levels of homocysteine have been correlated with PCOS, and these patients respond to folic acid or vitamin B complex therapy. The homocysteine/methionine cycle runs efficiently only upon assistance from folate cycle. 677 C>T polymorphism in the 5, 10-methylene tertrahydrofolate reductase (MTHFR), an important enzyme of folate pathway, compromises the activity of this enzyme by about 50%. PCOS risk associated with this polymorphism has been analyzed in few small case-control studies. Studies from South Asian/Asian region are lacking and a meta-analysis on published data is urgently required to settle the discordance in results. Methods used: In the present study, we have undertaken analysis on a fairly large sample size (261cases and 256 controls) to address the level of association between this polymorphism and PCOS and to figure out the effect of ethnic variations on the possible association. For the latter, two ethnically divergent case-control groups (North Indian and South Indian) from the Indian sub-continent were recruited for analysis. MTHFR 677 C>T polymorphism was analyzed by direct DNA sequencing. Data were analyzed to find if rare allele (T) or the rare genotypes (CT and TT) conferred increased risk of PCOS. We also conducted a meta-analysis on the data published between 1999 and 2011. Results: Independent (group-wise) and combined analysis of allele/genotype data showed no association between this polymorphism and PCOS. No significant difference was seen when the patients were grouped according to the BMI. Similar results in the two ethnically divergent groups rules out the influence of ethnic variations on the level of association. The consensus emerged from the meta-analysis also suggest no significant association of this polymorphism with PCOS. Conclusions/significance: MTHFR 677 C>T polymorphism does not seem to affect the risk of PCOS. No influence of this polymorphism is seen even in the individuals with high BMI. This inference is not influenced by ethnic variations across study populations.

1951T

Mapping of a novel locus modifying glaucoma severity within the *GLC1K* region at 20p12. *P. Belleau[1], S. Dubois[1], K. Lebel[1], R. Arseneault[1], E. Shink[1], J.L. Anctil[2], G. Côté[2], M.A. Walter[3], M. Amyot[4], V. Raymond[1].* 1) Molecular Genetics of Sensory Systems, CHUL Research Centre, Quebec City, QC, Canada; 2) Ophthalmology, Universite Laval, Quebec City, QC, Canada; 3) Medical Genetics, University of Alberta, Edmonton, AB, Canada; 4) Ophthalmology, Universite de Montreal, Montreal, QC, Canada.

Chronic open-angle glaucoma (COAG, which refers to juvenile- and/or adult-onset open-angle glaucoma) is a major cause of blindness characterized by retinal ganglion cell death. Although COAG is primarily considered a complex genetic disorder, the disease may also segregate as a Mendelian trait in some families. Using such kindreds, 12 *GLC1* loci and 4 disease-genes have been identified. One of these genes is *myocilin* (*MYOC*) that causes autosomal dominant COAG. For the past 50 years, we have been investigating the huge French-Canadian CA pedigree (749 people) in which the $MYOC^{K423E}$ mutation expresses a very wide phenotypic variability. Recently, we identified a novel class of modifier elements that accounted for variable ages-at-onset of the glaucoma phenotype within the family. The modifiers might alter the outflow of aqueous humor thereby increasing intraocular pressures above normal. To decipher the molecular basis by which these modifiers caused variable ages-at-onset, we performed a genome-wide linkage analysis using a total of 184 CA family members. Out of the 133 individuals who were heterozygotes for the $MYOC^{K423E}$ mutation, 114 were affected while the 17 remaining persons were asymptomatic even if 9 of them were older then 35 years of age at their last exam. A total of 408 microsatellite markers were genotyped at an average interval of 10 cM. To prioritize *GLC1* loci previously linked to open-angle glaucoma (OAG), markers density was increased within these regions. Linkage values were then calculated using a Bayesian MCMC method implemented in Loki. This procedure calculates L-score values by estimating the posterior probability of linkage divided by the prior probability. Using this strategy, we detected a 5 cM region with an L-score value of 7.1 located between D20S189 and D20S898 at 20p12. Interestingly, both markers mapped within the *GLC1K* locus that contains a gene for early-onset OAG. Even if this L-score value was at the lower limit for such type of analysis, it clearly confirms that *GLC1K* encodes at least one gene (or non-coding DNA element) that is linked to early-onset glaucoma and/or modifies the severity of the glaucoma phenotype when the disorder is caused by the myocilin gene. Our data suggest that the modifier at *GLC1K* acts by altering the first episode when intraocular hypertension occurs in our $MYOC^{K423E}$ heterozygotes.

1952F

Fine-mapping of eight psoriasis susceptibility loci. *S. Das[1], P.E. Stuart[2], J. Ding[1], T. Tejasvi[2], Y. Li[1], L.C. Tsoi[1], V. Chandran[3], J. Fischer[4], C. Helms[5], K.C. Duffin[6], J.J. Voorhees[2], A.M. Bowcock[5], G.G. Krueger[6], G.M. Lathrop[4], R.P. Nair[2], P. Rahman[7], G.R. Abecasis[1], D. Gladman[3], J.T. Elder[2,8].* 1) Center for Statistical Genetics, Department of Biostatistics, University of Michigan, Ann Arbor, MI; 2) Department of Dermatology, University of Michigan, Ann Arbor, MI; 3) Department of Rheumatology, University of Toronto, Toronto, Ontario, Canada; 4) Centre National de Génotypage, Institut Génomique, Commissariat à l'énergie Atomique, Evry, France; 5) Division of Human Genetics, Department of Genetics, Washington University at St. Louis, St. Louis, MO; 6) Department of Dermatology, University of Utah, Salt Lake City, UT; 7) Department of Medicine, Memorial University, St. John's, Newfoundland, Canada; 8) Ann Arbor Veterans Affairs Hospital, Ann Arbor, MI.

Psoriasis is a common immune-mediated disease with a genetic basis. Genome-wide association studies have identified more than 20 loci for this disease. In 2009 we designed a custom genotyping array and initialized a follow-up study to fine-map the eight genome-wide significant susceptibility loci known at that time (*IL23R*, *IL13*, *IL12B*, *TNIP1*, *MHC*, *TNFAIP3*, *IL23A* and *RNF114*). After quality control, the dataset consisted of genotypes of 1,962 SNPs from the 8 loci for 2,847 psoriasis cases and 2,208 unaffected controls of European Caucasian ancestry. To increase marker density, genotype data were imputed using the latest 1000 Genome reference haplotypes, which include both indels and SNPs, yielding 49,120 genetic variants in the 8 loci imputed with high quality. Stepwise conditional association analysis was used to identify independent signals from each of the 8 loci. We used a sequential forward selection method where, at each iteration, we identified the most significant variation in each locus conditional on signals from previous iterations. We identified 13 independent signals among the 8 loci and two of these were deletions. Among the 11 independently associated SNPs, rs848 and rs114175464 were found to be in strong LD with non-synonymous SNPs in *IL13* and *RNF114*, respectively. Three of the loci had more than one independent signal: *IL12B* (2), *MHC* (4), and *TNIP1* (2). These results add further definition to the genetic map of psoriasis.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1953W

Association of markers in DOCK6 with total cholesterol concentration in American Indians. *J.K. DiStefano[1], M. Khosroheidari[1], E. Eddy[1], S. Kobes[2], W.C. Knowler[2], R.L. Hanson[2].* 1) Diabetes, Cardiovascular Meta, TGen, Phoenix, AZ; 2) Diabetes Epidemiology and Clinical Research Section, NIDDK, NIH, Phoenix, AZ.

We have previously identified a locus linked to fasting serum total cholesterol concentration (TC) in Pima Indians on chromosome 19p (LOD=3.89). This locus has been linked to both total and LDL cholesterol in at least 15 independent studies across four different ethnic groups, and is one of the most consistently observed regions of linkage for lipid traits. Subsequent fine-mapping of this region using 346 single nucleotide polymorphisms (SNPs) genotyped in Pima Indians reduced the linkage interval from 90 cM to 15.7 cM, which spans 19p13.1-p13.3 and contains 229 genes. To identify trait-associated SNP alleles, we genotyped 1536 markers within the region of interest in in two study samples derived from the Pima Indian population, including 1) a family-based sample comprised of individuals who participated in the original genome scan for total cholesterol (N= 958) and 2) a population-based sample representing individuals in whom TC levels have been measured (N=926). We analyzed log values of TC using measurements from the last examination at which diabetes status was known and age was at least 20 years. Adjustment was performed for age, sex, diabetes status, and examination date. The strongest evidence for association with TC concentration was found with markers located in the DOCK6 (dedicator of cytokinesis 6) gene, including rs4804576 (increase of 4.2% per copy of the C allele; $P=2.0 \times 10^{-7}$), rs2116876 (4.0% per copy of the G allele; $P=1.1 \times 10^{-6}$), rs2278436 (4% per copy of the G allele; $P=5.0 \times 10^{-6}$), and five additional SNPs (all $P<0.0005$). Conditional analyses among the four most strongly associated markers showed that rs4804576 provided most of the evidence for association; however, rs12608933 also provided independent evidence for association (P=0.02), suggesting the potential presence of multiple functional variants with effects on TC levels or a single functional variant that crosses the haplotype of these four SNPs. DOCK6 is guanine nucleotide exchange factor belonging to group C of the DOCK family, and is involved with actin cytoskeletal reorganization via activation of RAC1 and CDC42; however, its role in cholesterol metabolism is not known. However, it is worth noting that rs4804576 is highly concordant with rs737337, which was previously found to be associated with HDL-C ($P=3 \times 10^{-9}$) in >100,000 individuals of European ancestry. These findings suggest that genetic variation in DOCK6 influences TC levels in Pima Indians.

1954T

Fine-scale genetic mapping reveals multiple signals of association in waist-hip ratio loci and shows evidence of sexually dimorphic genetic effects. *T. Ferreira[1], D. Shungin[2,3,4], D.C. Croteau-Chonka[5], T.W. Winkler[6], A.E. Locke[7], R. Mägi[1,8], R.J. Strawbridge[9], T. Workalemahu[10], K. Fischer[8], J. Wu[11], A.U. Jackson[7], F. Day[12], M.C. Zillikens[13], A.E. Justice[14], H. Völzke[15], I. Barrsos[16,17], C.S. Fox[18], E. Ingelsson[19], M.I. McCarthy[1,20], E.K. Speliotes[21,22], P.W. Franks[2,4,10], L. Qi[10], J.N. Hirschhorn[23], R.J.F. Loos[12,24], K.E. North[14], I.M. Heid[6], L.A. Cupples[11], A.P. Morris[1], K.L. Mohlke[5], C.M. Lindgren[1] on behalf of the Genetic Investigation of ANthropometric Traits (GIANT) consortium.* 1) WTCHG, University of Oxford, Oxford, United Kingdom; 2) Genetic and Molecular Epidemiology Group, Department of Public Health and Clinical Medicine, Section for Medicine, Umeå University, Umeå, Sweden; 3) Department of Odontology, Umeå University, Umeå, Sweden; 4) Department of Clinical Sciences, Skåne University Hospital, Lund University, Malmö, Sweden; 5) Department of Genetics, University of North Carolina, Chapel Hill, NC; 6) Public Health and Gender Studies, Institute of Epidemiology and Preventive Medicine, Regensburg University Medical Center, Regensburg, Germany; 7) Department of Biostatistics, University of Michigan, Ann Arbor, MI; 8) Estonian Genome Center, University of Tartu, Tartu, Estonia; 9) Cardiovascular Genetics and Genomics Group, Karolinska Institute, Stockholm, Sweden; 10) Department of Nutrition, Harvard School of Public Health, Boston, MA; 11) Department of Biostatistics, School of Public Health, Boston University, Boston, MA; 12) MRC Epidemiology Unit, Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge, UK; 13) Department of Internal Medicine, Erasmus MC Rotterdam, the Netherlands; 14) Department of Epidemiology and Carolina Center for Genome Sciences, University of North Carolina, Chapel Hill, NC; 15) Institute for Community Medicine, Ernst-Moritz-Arndt-University Greifswald, Greifswald, Germany; 16) Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, UK; 17) University of Cambridge Metabolic Research Labs, Institute of Metabolic Science Addenbrooke's Hospital, Cambridge, UK; 18) National Heart, Lung, and Blood Institute, Framingham, MA; 19) Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden; 20) OCDEM, University of Oxford, Oxford, United Kingdom; 21) Department of Internal Medicine, Division of Gastroenterology, and Center for Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI; 22) Broad Institute, Cambridge, MA; 23) Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA; 24) Charles R. Bronfman Institute of Personalized Medicine, Child Health and Development Institute, Department of Preventive medicine, Mount Sinai School of Medicine, New York, NY 10029, USA.

Waist-hip ratio (WHR) is a heritable measure of body fat distribution and a significant predictor of metabolic and cardiovascular risk independent of overall obesity, as measured by body mass index (BMI). To further our understanding of the genetic architecture of WHR adjusted for BMI (WHRadjBMI), we performed fixed-effects meta-analysis of 67,325 individuals from 28 studies genotyped with Metabochip and 142,762 individuals from 57 studies with genome-wide association imputed data, all of European ancestry. Given our previous reports of sexually dimorphic genetic effects on WHRadjBMI, we also performed sex-stratified meta-analyses. To search for multiple signals of association within WHRadjBMI loci, we performed approximate conditional analyses using GCTA, estimating the linkage disequilibrium between SNPs using genotype data from 949 Swedish individuals from PIVUS. We identified multiple independent signals of association at genome-wide significance ($p<5\times10^{-8}$) at 9 WHRadjBMI loci, many of which showed a complex pattern of sexual dimorphism in genetic effects. For example, at the *COBLL1* locus, all 4 association signals were female-specific (male-specific $p>0.4$ for all 4 signals), and at the *VEGFA* locus, both association signals were stronger in women than men (1-degree of freedom test of heterogeneity between sex-specific allelic effects $p=2.7\times10^{-10}$ and 2.2×10^{-2}). However, at the *WARS2/SPAG17* locus, we observed one male-specific, one female-specific, and two sex-combined association signals.

We also assessed fine-mapping resolution across 14 WHRadjBMI loci that are densely covered by Metabochip. Using the sex-combined meta-analysis results, we constructed credible sets of SNPs that account for 99% of the probability of being causal (or tagging an unobserved causal variant). For 6 of the loci, the interval covered by the credible set was <20kb, and at *VEGFA* contained just one SNP. We also observed that the addition of 14,340 individuals of non-European ancestry did not noticeably improve fine-mapping resolution, despite expected differences in LD structure between such diverse populations. Our results highlight the importance of fine-mapping to fully elucidate the complex underlying genetic architecture of body fat distribution, and show promise for extending our knowledge of the biological mechanisms contributing to the trait.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1955F

The FTO gene polymorphisms are associated with obesity in the Chinese Population: the SAPPHIRe follow-up study. *C.-M. Hwu[1, 2], M.-W. Lin[3, 4], P.-T. Tsai[3], C.-A. Hsiung[5], L.-M. Chuang[6], W.H-H. Sheu[7], Y.-R. Hung[8], L.-T. Ho[2, 4].* 1) Section of Endocrinology and Metabolism,, Taipei Veterans General Hospital, Taipei, Taiwan; 2) Faculty of Medicine, National Yang-Ming University School of Medicine, Taipei, Taiwan; 3) Institute of Public Health, National Yang-Ming University, Taipei, Taiwan; 4) Department of Medical Research and Education, Taipei Veterans General Hospital, Taipei, Taiwan; 5) Department of Bioinformatics and Biostatistics, National Health Research Institutes, Zhunan, Taiwan; 6) Graduate Institute of Clinical Medicine, National Taiwan University College of Medicine, Taipei, Taiwan; 7) Department of Internal Medicine, Taichung Veterans General Hospital, Taichung, Taiwan; 8) Division of Endocrinology and Metabolism, Tri-Service General Hospital, Taipei, Taiwan.

Obesity is an important risk factor for type 2 diabetes and cardiovascular diseases. The fat mass and obesity-associated gene (FTO) was recently identified as a susceptibility locus for both obesity and type 2 diabetes by genome-wide association studies in several European populations. To investigate the association between FTO gene polymorphisms and obesity and metabolic syndrome in the Chinese population, we genotyped three single nucleotide polymorphisms (SNP) of the FTO gene (rs1421085, rs9939609, and rs9930506) in 847 subjects of Chinese origin from the Stanford Asia-Pacific Program for Hypertension and Insulin Resistance (SAPPHIRe) family study. The rs1421085 and rs9939609 were found to be associated with risk of increased waist circumference in the SAPPHIRe cohort (p= 0.017 for rs1421085, p=0.031 for rs9939609). Moreover, The rs1421085 and rs9939609 were also revealed to be associated with the increase of body mass index (OR=1.71, 95%C.I. 1.07-2.73 for rs1421085 and OR=1.68, 95%C.I. 1.05-2.71, for rs9939609) after five-year follow-up of the study subjects. In conclusion, our results suggest the polymorphisms of the FTO gene are significantly associated with obesity as well as the progression of obesity. Our study validated the association between FTO gene polymorphisms and obesity in the Chinese population.

1956W

Consanguinity in Pakistani families reveals new loci in a complex genetic trait: stuttering. *M.H. Raza[1], R. Amjad[2], S. Riazuddin[2], D. Drayna[1].* 1) LCD, SSBCD, NIDCD, National Institute of Health; 2) Allama Iqbal Medical College, Lahore, Pakistan.

Stuttering is a complex genetic disorder which is characterized by involuntary disruption of the flow of speech. Although the twin and adoption studies have demonstrated high heritability of stuttering, incomplete penetrance and the absence of complete Mendelian segregation contribute to the limited success of linkage studies and next generation sequencing for this disorder. Linkage and gene finding studies of consanguineous stuttering families previously led to the identification of mutations in the GNPTAB gene associated with stuttering in Pakistani families. We have performed additional SNP linkage scans in 32 medium-to-large size consanguineous stuttering families ascertained from Pakistan, using the Human Linkage-12 and Human Linkage-24 chips (Infinium Assay) from Illumina that carries 5913-6090 SNPs occurring at an average density of 0.58 cM. Two-point linkage analysis was performed under different modes of inheritance using SuperLink in the EasyLinkage package. Pakistani stuttering family PKST77 generated a LOD score of 3.23 at SNP rs7631540 on chromosome 3q13.31 under a recessive mode of inheritance. Another family designated PKST58 produced suggestive linkage scores at loci on chromosomes 16, 6, 11, 9 and 21. Based on these results, targeted microsatellite genotyping was performed in families PKST77 and PKST58. Two point analysis using MLINK generated a maximum LOD score of 4.23 in PKST77 at marker D3S1310, and a score of 4.35 in PKST58 at marker D16S3025. Additional evidence of linkage was obtained through multipoint analyses, which produced LOD scores of 4.92 in PKST77 and 4.42 in PKST58 under recessive modes of inheritance. After including microsatellite genotype information on chromosomes 6, 11, 9 and 21, LOD scores fell under all modes of inheritance in PKST58. Our results demonstrate significant evidence of linkage to chromosome 3q and chromosome 16q in consanguineous stuttering families PKST77 and PKST58 respectively that harbor new stuttering genes. These studies indicate the power of consanguineous families for identifying rare variants of large effect that underlie common complex genetic disorders.

1957T

Genome-wide linkage scan of male sexual orientation. *A.R. Sanders[1,2], K. Dawood[3], G. Rieger[4], J.A. Badner[2], E.S. Gershon[2], R.S. Krishnappa[5], A.B. Kolundzija[6], S. Guo[7], G.W. Beecham[7], E.R. Martin[7], J.M. Bailey[8].* 1) NorthShore University HealthSystem, Evanston, IL; 2) University of Chicago, Chicago, IL; 3) Pennsylvania State University, University Park, PA; 4) Cornell University, Ithaca, NY; 5) East Tennessee State University, Johnson City, TN; 6) Columbia University, New York, NY; 7) University of Miami, Miami, FL; 8) Northwestern University, Evanston, IL.

As with many other psychological traits, sexual orientation appears genetically complex, with important genetic and environmental factors. Family and twin studies have consistently demonstrated that genetic contributions are important for the development of sexual orientation in men. Previous research has found evidence for linkage to chromosome Xq28 in some studies focusing on pairs of homosexual brothers, but not in other such studies. To map loci for sexual orientation, and to appraise for potential replication of previous findings, we have conducted a genome-wide linkage scan using genotypes from the Affymetrix 5.0 SNP array on 410 independent pairs of homosexual brothers in 385 families, a sample set independent from previous studies. We classified subjects as homosexual based on both their self-reported sexual identity ("homosexual") and sexual feelings (Kinsey score of 5-6, indicating predominant or exclusive attraction to men). After rigorous quality control steps, we performed non-parametric linkage analysis using MERLIN. The strongest linkage peak, with multiple two-point LOD scores over 3.5 (maximum 4.0) and multipoint support, was in the pericentromeric region of chromosome 8, overlapping with the second strongest linkage peak in the next largest reported linkage scan of 155 independent pairs of homosexual brothers in 146 families. Our second strongest linkage region was on chromosome X with a maximum two-point LOD score of 2.65 at Xq28, the previously reported linkage region. Our findings, taken in context with previous work, suggest that genetic variation in each of these regions contributes to development of the important psychological trait of male sexual orientation.

1958F

Linkage studies in a large stuttering family indicate multiple novel loci and possible assortative mating. *A.A. Schaffer[1], M.H. Raza[2], E.M. Gertz[1], J. Mundorff[2], J. Lukong[2], J. Kuster[3], D. Drayna[2].* 1) National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD USA; 2) National Institute on Deafness and Other Communication Disorders, National Institutes of Health, Bethesda, MD USA; 3) Department of Speech, Hearing, and Rehabilitative Services, Minnesota State University, Mankato, MN.

Stuttering is a common speech disorder that is a complex trait with high heritability. While linkage studies in consanguineous populations have produced highly significant evidence for linkage at a number of loci, such studies in outbred populations have been less successful. We have ascertained a large family with many cases of persistent stuttering from Cameroon, West Africa in which consanguinity is denied. An initial genome-wide linkage scan using microsatellites that tested the complete pedigree for linkage produced a suggestive score on chromosome 1. However subsequent linkage scans with SNPs that analyzed the complete family failed to confirm this linkage, and no significant linkage scores were observed genome-wide. In contrast, dividing the pedigree into five branches and analyzing individual branches and combinations of branches produced strong evidence for linkage with both SNP and microsatellite markers. In this family, we found evidence for linkage to previously reported loci on 3q and 15q, and to novel loci on 2p, 3p, 14q, and a different position on 15q. While the mechanism by which these multiple alleles came together in the same family is unknown, we found little evidence for consanguinity, and we suggest assortative mating as a likely explanation. Such assortative mating has been shown to be important in the epidemiology of hereditary deafness, and may play a role in the epidemiology of stuttering.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1959W

Fine mapping of 6q27 for association with blood pressure. *B. Tayo[1], B. Salako[2], A. Luke[1], X. Zhu[3], A. Adeyemo[4], C. Rotimi[4], A. Ogunniyi[2], R. Cooper[1].* 1) Prev Med & Epidemiology, Loyola Univ Chicago, Maywood, IL; 2) University of Ibadan, Ibadan, Nigeria; 3) Case Western Reserve University, Cleveland, OH; 4) NIH Intramural Center for Research on Genomics and Global Health, National Human Genome Research Institute, Bethesda, MD.

Hypertension is the most common cardiovascular condition in the world and accounts for a substantial proportion of adult mortality. Although elevated blood pressure (BP) has similar heritability to many other traits related to cardiovascular risk, genetic susceptibility loci have been difficult to localize. In a follow-up family-based association analysis of regions on chromosomes 6 and 7 linked to BP among Nigerians we observed strong preliminary evidence that the regions may influence susceptibility to elevations in BP. To refine identified association in 6q27, we conducted fine mapping analysis in a sample of 1614 unrelated adult Nigerians genotyped on Affymetrix 6.0 chip. The study sample comprised of 940 females and 674 males of which 797 were hypertensives. To increase single nucleotide polymorphisms (SNPs) density for the fine mapping, we performed genotype imputation using combined data from the HapMap and 1000 Genomes projects. A total of 3330 typed or imputed SNPs passed quality control in 6q27 and these were subsequently tested for association with BP and hypertension with covariates adjustment for sex, age and body mass index. After correction for multiple testing through Bonferroni adjustment, 7 intergenic SNPs within 70 kilobase pairs in 6q27 retained statistical significance (P<0.005) for association with hypertension. These data reinforce previous findings that variants in chromosome 6 influence susceptibility to elevated BP.

1960T

Efficient Pooled Sequencing of 1,077 Candidate Genes in Individuals With Short Stature. *S.R. Wang[1,2,3], D.B. Mirel[2], A. Dauber[2,3], J.N. Hirschhorn[1,2,3].* 1) Department of Genetics, Harvard Medical School, Boston, MA, USA; 2) Broad Institute, Cambridge, MA, USA; 3) Division of Endocrinology, Boston Children's Hospital, Boston, MA, USA.

Genome-wide association (GWA) studies have identified many common variants associated with height and other polygenic traits. These common variants typically account for well under half of the heritable variation. Many explanations have been proposed to account for this “missing heritability”, including the contribution of rare variants. In this study, we employ a candidate gene approach to search for rare variants associated with human height, focusing on genes within loci identified by previous GWA studies of height, and supplemented with genes known to underlie syndromic growth disorders or skeletal dysplasias, as well as genes involved in growth plate biology. In total, a list of 1077 genes were compiled and the exons (~2Mb in total) were targeted for resequencing. As an initial experiment, we resequenced 192 patients with short stature (>2 SD below the mean for age and gender). To increase sequencing efficiency, we applied a matrix pooling design. The short stature samples were arranged into a 14x14 matrix of 28 pools, with 13-14 patients in each pool (4 empty “holes” were included into the matrix for assessing false positive rate). After hybrid selection of each pooled sample against the baits representing the 1077 genes, barcoded libraries were created. The matrix design allows us to obtain allele frequency estimates for each pool and to ascertain variant carriers for singleton variants appearing once in the matrix. To call variants within pools, we used Syzygy software (Rivas et al. 2011) and then applied a new likelihood-based secondary calling strategy that integrated the extra information from our matrix design. We estimated the false negative and false positive rates of this strategy by comparison with exome sequencing performed in a subset of these samples and by looking for singleton variants that map to one of the holes in the matrix. The overall false negative rate was ~0.2%, while the false positive rate of singletons was ~0.8%. We identified a list of very rare variants (not observed in any currently available sequencing dataset) that appear in short stature samples. Of note, one individual who did not have a known genetic diagnosis previously, was a compound heterozygote for mutations in a gene known to cause a rare inborn error of metabolism (SLC35C1, congenital glycosylation disorder type IIC) that has short stature as a feature. Biochemical analysis of the individual and functional assessment of these mutations are now in progress.

1961F

Molecular analysis of common tag polymorphisms in Hemochromatosis (HFE) genes in Iranian patients affected with PCOS. *S. Matoo, A. Yasari Mazandarani, N. Hatamnejadian, B. Sedaghati Khayat, M. Moghaddam, A. Ebrahimi.* PARSEH GENETIC CENTER, TEHRAN, Iran.

Abstract: Polycystic ovary syndrome (PCOS) is a heterogeny disease digit between 4-7% of women of childbearing age will be involved. The most original Endocrine disorders is in women . One of the main mechanisms involve in pcos insulin resistance or mechanisms that cause diabetes. On the other hand, several epidemiological studies have shown a positive association between increased body iron stores and the risk of type 2 diabetes and of other insulin resistant states such as the metabolic syndrome, pregnancy diabetes and polycystic ovarian syndrome. Here is evaluated two different types of hemochromatosis, including: the classic hemochromatosis, caused by a mutation in a gene designated HFE and type 2A Hemochromatosis, caused by mutations in the HJV gene. Thus, using molecular techniques the relationship between HFE and HJV genes involve in insulin resistance and PCOS are reviewed Methods: To identify the molecular changes common HFE and HJV genes 100 affected patients clinically diagnosed as PCOS were selected. The blood samples were collected and DNA was extracted. And then using molecular techniques ARMS-PCR and PCR-RFLP variations were common, were investigated. Also, 10 percent of the samples were analyzed by direct sequencing method. Results: Our study showed that mutations in HFE and HJV are polymorphic and significant in Iranian population of affected with PCOS. Which polymorphic site could be gene biomarker for prediction of PCOS early signs and genotype-phenotype correlationS studies.

1962W

Molecular analysis of Hepcidin common polymorphisms in Iranian patients affected with PCOS. *A. Yasari Mazandarani, S. Matoo, N. Hatamnejadian, B. Sedaghati Khayat, M. Moghaddam, A. Ebrahimi.* PARSEH GENETIC CENTER, TEHRAN, Iran.

Introduction: The polycystic ovary syndrome (PCOS) is possibly the most common endocrine disorder, affecting several million women worldwide with approximately 6-7% prevalence. Several studies have reported that high body iron stores have a positive association with risk of diabetes and polycystic ovarian syndrome (PCOS).Since hereditary hemochromatosis (HH) is a common autosomal recessive disorder of iron overload, it can lead to insulin resistant disorders such as PCOS. On the other hand, one of the forms of HH is caused by mutations in the hepcidin antimicrobial peptide (HAMP) gene on chromosome 19, which makes iron overload by decreasing in hepcidin synthesis and thus reduced levels of hepcidin. Consequently, by considering the significant role of hepcidin in regulation of iron metabolism, we studied hepcidin common polymorphisms in order to find the association between the HAMP and PCOS in Iranian patients. Material: To identify the molecular common changes of HAMP genes, 100 patients affected with PCOS (100 PCOS patients and 100 control diagnosis samples) were selected. The blood samples were collected and DNA of samples were extracted. Then HAMP gene was amplified using primers and the PCR products were analyzed by direct sequencing. Results: Our study showed that mutations in HAMP are polymorphic and significant in Iranian population of affected with PCOS. Which polymorphic site could be gene biomarker for prediction of pcos early signs and genotype-phenotype correlations studies.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1963T

Mitochondrial haplogroups and age-related maculopathy. *Y.P. Conley[1,2], Y. Jiang[3], M.C. Kenney[4], N. Udar[4], R.E. Ferrell[2], D.E. Weeks[2,3], M.B. Gorin[5].* 1) Hlth Promotion & Development, University of Pittsburgh, Pittsburgh, PA; 2) Department of Human Genetics, University of Pittsburgh, Pittsburgh PA; 3) Department of Biostatistics, University of Pittsburgh, Pittsburgh, PA; 4) Gavin Herbert Eye Institute, University of California, Irvine, CA; 5) Jules Stein Eye Institute, UCLA, Los Angeles, CA.

Age-related maculopathy (ARM) is a complex multifactorial condition where nuclear genomic variation plays a role in susceptibility; however the potential role for mitochondrial variation in susceptibility has been less frequently studied. The purpose of this study was to investigate whether mitochondrial haplogroups H and T play a role in susceptibility to ARM in our familial ARM cohort. TaqMan allele discrimination assays were conducted to classify haplogroup H (C702T) and haplogroup T (A4917G) and ad hoc tests for association conducted. For haplogroup H, we determined the status for 1,009 Caucasian individuals in 587 families, including 860 affecteds and 126 unaffected controls. Haplogroup H was associated with ARM in the total sample (p=0.019) but only in females in the sex-specific subgroup analyses (p=0.019). Haplogroup T indicated no significant association in our analyses. Haplogroup H has been previously associated with ARM; however this is the first study conducted within a cohort of ARM families and the first to find a sex-specific association. Additionally, our findings indicate that the presence of haplogroup H increases risk for ARM, particularly in females, which is counter to previous findings that haplogroup H is associated with reduced risk. This may be an artifact due to our sampling scheme. In order to investigate the functional consequences of the mitochondrial DNA variants that define the ancestral haplogroups, we also developed cybrids (cytoplasmic hybrids) using human retinal pigment epithelial cells. Our findings show that the different cybrids, including the haplogroup H containing cybrid, have different modes of energy production, rates of cell growth, and altered expression of nuclear genes involved in critical molecular pathways and cellular functions. These data support that different haplogroups have unique features which may relate to disease development.

1964F

Candidate genes as a modifier of age at Huntington disease motor onset. *E.M. Ramos[1,2], J. Latourelle[3], J.-H. Lee[1], T. Gillis[1], J.S. Mysore[1], J.F. Gusella[4, 1], J.-M. Lee[1], I. Alonso[2,5], J. Sequeiros[2,5], R.H. Myers[6,3], M.E. MacDonald[1].* 1) Center for Human Genetic Research, Massachusetts General Hospital, Boston, MA; 2) UnIGENe, IBMC - Institute for Molecular and Cellular Biology, University of Porto, Porto, Portugal; 3) Department of Neurology, Boston University School of Medicine, Boston, MA, USA; 4) on behalf of the COHORT study of the HSG; 5) CGPP, IBMC - Institute for Molecular and Cellular Biology, University of Porto, Porto, Portugal; 6) on behalf of the HD-MAPS Investigators.

Huntington's disease (HD) is a dominantly inherited neurodegenerative disorder characterized by motor, cognitive, and behavioral disturbances, caused by the expansion of the *HTT* CAG trinucleotide repeat. The expanded CAG repeat is the major determinant of age-at-onset of motor symptoms (AO), such that the longer the repeat the earlier the onset. CAG size accounts for 70% of the variability in onset age. The remaining variance in AO (residual AO) is highly heritable but remains unexplained. In the search for modifiers, polymorphisms in about 20 loci have been implicated in AO or progression of HD, though with different results in different studies. We have now critically evaluated some of these polymorphisms located either in the mitochondrial regulator *PPARGC1A*, N-methyl D-aspartate receptor subtype genes (*GRIN2A* and *GRIN2B*) or in genes in the dopamine pathway (*DAT1*, *DRD4*, *DRD2* and *COMT*). We genotyped genomic DNA from a total of 1,727 HD patients with 40 to 53 CAGs and known genetic/self-reported ancestry and familial relationship. Multivariate analyses using GEE assessed the effect on HD residual motor onset, adjusting for familial component and ancestry. For the *PPARGC1A* gene, in the entire cohort, rs7665116 genotype had a significant effect on AO, both in additive (p=0.009) and dominant models (p=0.008). However, when examined by origin, the Southern European and the Northern European and North American HD samples, differed for rs7665116 MAF, consistent with a SNP tagging-ancestry, as, despite similar mean CAG sizes, Southern European samples had an older mean AO, revealing phenotype stratification. Population-specific regression analysis did not show a significant association of rs7665116 genotype with AO, in either additive or dominant models. Using this same approach, our analysis of the implicated polymorphisms in *GRIN2A*, *GRIN2B*, *DRD2*, *DRD4*, *DAT1* and *COMT* were also found not to significantly modify motor AO, either in a dominant or additive model, after adjusting for ancestry. Our results, therefore, do not support any of the previously implicated genes as modifiers of AO of motor symptoms. We found evidence for a dramatic effect of phenotypic (AO) and genotypic (MAF) stratification among European cohorts that was not considered in previously reported association studies. A significantly older AO in Southern Europe may reflect population differences in genetic or environmental factors that warrant further investigation.

1965W

Variations in Potassium Channel Genes are Associated with Persistent Breast Pain after Breast Cancer Surgery. *D.J. Langford[1], B. McCann[1], T. Koetters[1], C. Baggott[1], C. West[1], J.D. Levine[2], C. Elboim[3], G. Abrams[2], D. Hamolsky[2], L. Dunn[2], H. Rugo[2], M. Dodd[1], S.M. Paul[1], J. Neuhaus[2], B.A. Cooper[1], B. Schmidt[4], J. Cataldo[1], A. Dhruva[2], B.E. Aouizerat[1,5].* 1) Physiological Nursing, University of California, San Francisco, San Francisco, CA; 2) School of Medicine, University of California, San Francisco, San Francisco, CA; 3) Redwood Regional Oncology Group, Santa Rosa, CA; 4) School of Dentistry, New York University, New York, NY; 5) Institute for Human Genetics, University of California, San Francisco, CA.

Background/Purpose: Persistent pain after breast cancer surgery is common and prevalence estimates vary widely. Evidence suggests a role for common genetic variation in nociception (i.e., perception of noxious stimuli). Given their putative role in nociceptive transmission, we hypothesized that variation in potassium (K+) channel genes may be associated with persistent breast pain after surgery. Therefore, the purpose of this study was to test for associations of single nucleotide polymorphisms (SNPs) and inferred haplotypes among 10 K+ channel genes with persistent breast pain after breast cancer surgery.**Methods:** We recently used growth mixture modeling in an effort to identify latent classes (i.e., subgroups) of patients with distinct trajectories of worst pain scores using a 0-10 numeric rating scale (NRS). Pain was rated prior to and monthly for 6 months after breast cancer surgery. The comparisons were made solely between the two largest classes (i.e., No Pain and Mild Pain). Using a tag-SNP and literature-driven approach to SNP selection, genotyping was conducted using a custom genotyping array and blinded to clinical status. Only common SNPs (minor allele frequencies>.05) that passed quality control filters were analyzed. Haplotypes were constructed using a Bayesian approach. Contingency table tests identified significant differences in genotype or haplotype frequency between the pain groups. **Results:** Four subgroups were identified: 126 patients (31.7%) with "No Pain", 173 (43.4%) with "Mild Pain" (NRS of ~3 that remained constant for 6 months), 53 (13.3%) with "Moderate Pain" (NRS of ~2 that increased over 6 months), and 46 (11.6%) with "Severe Pain" (NRS of ~8 that persisted for 6 months). Compared to the No Pain group, patients in the Mild Pain group were younger (p<.01), had lower functioning (p<.05), had more lymph nodes removed (p=.02), more frequently had reconstruction at the time of surgery (p<.01), were pre-menopausal (p<.05), and had strange sensations in the breast prior to surgery (p<.001). We found significant associations between variations in 6 genes [KCNA1 (n=1, p<.01), KCND2 (n=1, p<.05), KCNJ3 (n=6, all p<.05), KCNJ6 (n=10, all p<.05), KCNK9 (n=3, all p<.05) and KCNS1 (n=2, both p<.05)] and pain group membership.**Conclusion:** K+ channel gene variation and important clinical and demographic characteristics are associated with novel phenotypic characterization of persistent breast pain after breast cancer surgery.

1966T

HLA Haplotypes Influence Hearing in Old Age. *T. Solomon[1], D.L. Newman[2].* 1) College of Health Science and Technology, Rochester Institute of Technology, Rochester, NY; 2) Thomas H. Gosnell School of Life Sciences, Rochester Institute of Technology, Rochester, NY.

Age-related hearing impairment (ARHI), or presbycusis, is characterized by reduced hearing sensitivity and speech understanding in noisy environments, slowed central processing of acoustic information, and impaired localization of sound sources. While both environment and genetic factors affect hearing loss, evidence has been shown that up to 50% of the variance in presbycusis is genetic. Thus far, only a few genes have been implicated in ARHI, including *GRM7*, *GRHL2*, and *NAT2*. There have been several studies correlating the human leukocyte antigen locus (HLA) with other inner ear and neurological diseases, specifically an extended HLA haplotype that contains the A1/B8 haplotype. We are investigating the A1/B8 haplotype for over 600 subjects with varying degrees of hearing capabilities. We are genotyping the patients for four SNPs that represent this abbreviated haplotype by using restriction fragment length polymorphism (RFLP) analysis and real-time PCR. Preliminary results suggest that the HLA-B8 haplotype is associated with worse pure tone detection and worse hearing in a noisy background (p = 0.024 and p = 0.017, respectively). In contrast, the HLA-A1 haplotype is associated with an improved ability to detect small temporal gaps in sound (p = 0.007). In the future, we would like to further investigate this extended HLA haplotype, particularly the DR3 haplotype.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1967F

A Y chromosome association study in inflammatory bowel disease. *L. Jostins, C.A. Anderson, J.C. Barrett, The International IBD Genetics Consortium.* Wellcome Trust Sanger Institute, Cambridge, United Kingdom.

Studies have suggested a link between sex chromosomes and immunity. Autoimmune diseases are more prevalent in females than in males, and are more common in individuals with Turner syndrome (absence of one sex chromosome). There is also evidence that Y haplogroups affect progression of HIV infection in males. However, there are few additional large studies of the effect of Y chromosome variation on human immune-mediated diseases.

We performed an association analysis of Y chromosome haplogroups in inflammatory bowel disease (IBD) using data from the International IBD Genetics Consortium's Immunochip project. This dataset contains 12,318 male IBD cases and 9,811 male controls from 15 European countries typed on a custom genotyping chip that includes 1735 Y chromosome variants. Using a novel method to assign Y haplogroups, we showed that these samples represent 11 major haplogroups, including 6 with a frequency above 1%. The frequency spectrum of these 6 haplogroups shows evidence of differing between IBD cases and controls after controlling for country-of-origin, genotyping batch and principal components (logistic likelihood-ratio $\chi 2 = 14.2$, d.f. = 5, p = 0.014), driven by an enrichment of haplogroup N in IBD cases (OR = 1.53, p = 0.006). Restricting the analysis to a homogeneous subset of 674 cases and 343 controls from populations with a high frequency of haplogroup N did not reduce the strength of the association (OR = 1.90, p = 0.003), suggesting that the association is not due to population stratification. We used sequence data from the 1000 Genomes Project to identify haplogroup N specific mutations that may account for this association. Mutations in or near 15 genes, including a variant 3kb upstream of CD24 (a cell adhesion gene upregulated in IBD) and a missense mutation in KDM5D (a MHC antigen involved in graft rejection) are among possible causal variants.

To extend our analysis, we generalized our haplogrouping method to include probabilistic sequence data, and will use cases and controls sequenced at low depth to further clarify Y chromosome associations in IBD. Additionally, we will analyze other immune-related diseases with varying sex ratios, such as primary sclerosing cholangitis (male-biased) and primary biliary sclerosis (female-biased), to investigate broader relationships between Y chromosome variation and immunity.

1968W

Prevalence of alpha-1 antitrypsin deficiency and hereditary hemochromatosis gene mutations in Algarve, Portugal. *V. Gaio[1], A. Fernandes[2], F. Mendonça[3], F. Orta Correia[3], A. Beleza[2], A. Gil[1], M. Bourbon[1], A.M. Vicente[1], C.M. Dias[1], M. Barreto da Silva[1].* 1) Instituto Nacional de Saúde Doutor Ricardo Jorge, Lisboa, Portugal; 2) Laboratório de Saúde Pública Dra. Laura Ayres, Faro, Portugal; 3) Administração Regional de Saúde do Algarve, Faro, Portugal.

Alpha-1 antitrypsin (AAT) deficiency and hereditary hemochromatosis (HH) are two of the most fatal genetic disorders in adult life, affecting million individuals worldwide. They are often under-diagnosed conditions and diagnosis is only made when the patient is already in the advanced stages of damage. AAT deficiency results from mutations in one highly pleiomorphic gene located on chromosome 14, *SERPINA 1*, being Z and S mutations the most relevant clinically. These mutations will lead to an AAT deficit that compromises the lungs protection, originating emphysema, chronic bronchitis, asthma or even chronic obstructive pulmonary disease (COPD) and it is also strongly associated with various liver diseases. On the other hand, C282Y and H63D mutations in the *HFE* gene, located on chromosome 6, are reported to be mostly responsible for the iron accumulation in HH disorder, leading to severe damage in different organs. Disease manifestations include cirrhosis, hepatic fibrosis, diabetes mellitus, arthropathy and hepatocarcinoma. Given the insufficient population-based information about the prevalence of these gene variants in the Portuguese population, the aim of this study was to assess their frequency in a representative sample from São Brás de Alportel, in the South of Portugal. To achieve our goal, we have genotyped a total of 208 adult subjects, including 118 females and 90 males (mean age: 58 years, range: 26-91). Regarding AAT deficiency, we found 4,3% MZ, 0,5% SS and 15,4% MS genotypes. The calculated frequency for the Z allele was 2,2% (95% CI: 0-11,7%) and for the S allele was 8,2% (95% CI: 0-17,4%). About HH, we found 1,4% C282Y/H63D, 2,4% H63D/H63D, 5,8% C282Y/N and 23,6% H63D/N genotypes. Frequencies of C282Y and H63D alleles were 3,6% (95% CI: 0-13%) and 14,9% (95% CI: 6-23,8%), respectively. The observed allele frequencies were in Hardy-Weinberg Equilibrium and no association was found with related diseases likely due to the smaller sample available. Our findings show the highest prevalence of Z allele from *SERPINA1* gene found, when compared to other populations. The remaining findings are in agreement with previously published studies. Future studies involving a larger sample size will be necessary to evaluate the penetrance of the studied gene mutations and to assess gene-environment interactions that influence disease risk, contributing to reduce the burden of these diseases which can have a great public health impact.

1969T

Menarche timing-related variants identified by recent GWAS are associated with growth outcomes in male and female children. *C. He[1], E. Wagner[1], G. Eckert[2], Z. Yu[2], H. Pratt[3], W. Tu[2].* 1) Dept. of Public Health, Indiana Univ Sch Med, Indianapolis, IN; 2) Dept. of Biostatistics, Indiana Univ Sch Med, Indianapolis, IN; 3) Dept. of Medicine, Indiana Univ Sch Med, Indianapolis, IN.

Weight and height increase in an accelerated fashion during puberty. Previous studies identified novel genetic variants that were associated with age at menarche, a marker for female puberty timing. However, the effects of these variants on growth related outcomes have not been investigated. We examined the growth effects of 31 menarche-related SNPs newly identified by genome-wide association studies (GWAS) in white women of European ancestry. We analyzed longitudinal data collected from 601 adolescents (263 African Americans and 338 Whites) using mixed effect models. Subjects between ages 4 and 17 were enrolled and followed for an average of 6 years. After adjusting for age and multiple testing, we found 2 menarche-related SNPs, rs1398217 in *SKOR2* and rs17268785 in *CCDC85A*, were respectively associated with more rapid weight gain and greater weight in African American males and females (all P< 0.0022); and two other SNPs, rs7642134 near *VGLL3* and rs9635759 near *CA10*, were respectively associated with increased weight and faster weight gain in White males and females (all P<0.0022); and rs7642134 and rs7821178 near *PEX2* were associated with greater body mass index (BMI) and greater BMI gain in Whites (all P<0.0032). None of the 16 menarche-related SNPs were associated with changes in height. Of the 8 menarche and weight-related SNPs, none was associated with changes in height; but 3 SNPs, rs2815752 near *NEGR1*, rs7647305 near *ETV5* and rs6548238 near *TMEM18*, were associated with increased weight and BMI in Whites and African Americans (all P<0.0034); and rs7138803 near *FAIM2* and rs9939609 in *FTO* were respectively associated with increased weight and BMI in Whites (all P<0.0059). Among the 7 menarche and height-related SNPs, we found rs1042725 near *HMGA2* and rs4549631 near *LOC728666* were associated with increased height and more rapid height gain in African Americans, and rs4800148 in *CABLES1* was associated with height increase in Whites (P=0.0018), whereas rs757608 near *TBX2* was associated with faster height gain in both Whites and African Americans (all P<0.0007); In addition, rs757608 and rs6440003 in *ZBTB38* were also associated with a faster gain in weight in African Americans (all P=0.00003). In conclusion, novel genetic variants associated with age at menarche were found to correlate with various growth measures in male and female children, although the genetic effects appear to be sex and race- specific.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1970F

Haplotypes of the inducible nitric oxide synthase gene are strongly associated with levels of exhaled nitric oxide in adults: a population-based study. *S. Dahgam[1], L. Modig[2], AT. Naluai[3], AC. Olin[1], F. Nyberg[1,4].* 1) Occupational and Environmental Medicine, University of Gothenburg, Gothenburg, Sweden; 2) Department of Public Health and Clinical Medicine Occupational and Environmental Medicine, University of Umeå, Sweden; 3) Department of Medical and Clinical Genetics, Sahlgrenska Academy at University of Gothenburg, Gothenburg, Sweden; 4) AstraZeneca R&D, Mölndal, Sweden.

Introduction: Inducible nitric oxide synthase (iNOS; NOS2) is one of three distinct isoforms of NOS enzymes and it is encoded by the *NOS2* gene. NOS2 appears to be the primary source of nitric oxide in exhaled breath. Haplotypes and several genetic variants in the *NOS2* gene have been associated with levels of fractional exhaled nitric oxide (FENO), a biomarker of airway inflammation, in American and European subjects [1,2]. Aim: To fine map previous findings by us and others using an extended study population and enlarged SNP set, to further understand the contribution of genetic variation within the *NOS2* gene for levels of FENO. Method: In total 11 SNPs in the *NOS2* gene located between 26.07 and 26.12 Mb on chromosome 17 were genotyped by a Sequenom MassArray platform or a competitive allele specific PCR. Seven were from our previous study (rs4795051, rs9901734, rs2297514, rs2248814, rs12944039, rs3729508 and rs2779248)[2] and four from the Southern California Children's study (rs4796017, rs2297520, rs9895453 and rs10459953)[1]. We measured FENO at an exhalation flow-rate of 50 m/s for each subject. Haplotype association of the *NOS2* gene with levels of FENO was investigated in 5912 Swedish adults aged 25-75 years, assuming an additive genetic model. Results: The analyses show that haplotypes of the two SNPs rs9901734 (C/G) and rs2779248 (T/C) are significantly associated with FENO. Compared with the baseline CT haplotype (frequency 45%), haplotypes CC (31%), GC (7%) and GT (16%) were associated with an increase in FENO levels of approximately 6.2% (95CI 4.1-9.4%; p=4.1E-07), 11.6% (95CI 7.2-17.3%; p=7.8E-07) and 10.5%. (95CI 7.2-15%; p=8.7E-10), respectively. Conclusion: The fine-mapping analyses have enabled us to narrow down the relevant genetic region of the *NOS2* gene and provide a strong foundation for further work to identify causative polymorphisms which may affect FENO in the region by e.g. sequencing. References: 1. **Salam MT**, Bastain TM, Rappaport EB, Islam T, Berhane K, Gauderman WJ, Gilliland FD. Genetic variations in nitric oxide synthase and arginase influence exhaled nitric oxide levels in children. Allergy 2011;66:412-9. 2. **Dahgam S**, Nyberg F, Modig L, Naluai AT, Olin AC. Single nucleotide polymorphisms in the NOS2 and NOS3 genes are associated with exhaled nitric oxide. J Med Genet 2012;49:200-5.

1971W

Identification of genetic factors for chronic otitis media with effusion using a targeted resequencing approach. *E. Allen[1,2], J. Mychaleckyj[1,3], X. Hou[1], W. Chen[1,3], A. Quinlan[1,2,3], K. Keene[1,3], S. Rich[1,3], K. Daly[5], M. Sale[1,2,4].* 1) Center for Public Health Genomics, University of Virginia, Charlottesville, VA; 2) Dept Biochemistry and Molecular Genetics, University of Virginia, Charlottesville, VA; 3) Dept Public Health Sciences, University of Virginia, Charlottesville, VA; 4) Dept Medicine, University of Virginia, Charlottesville, VA; 5) Dept Otolaryngology, University of Minnesota, Minneapolis, MN.

Objectives: We previously conducted the first Genome Wide Association Study (GWAS) of chronic otitis media with effusion and/or recurrent otitis media (COME/ROM), identifying associations at chromosomes 2 and 15. Additionally, our prior linkage analysis identified a region of chromosome 19q as harboring a susceptibility locus for COME/ROM. The aim of this study is to identify the causal variant(s) at these loci using a targeted massively parallel resequencing approach. Methods: Agilent's SureSelect Target Enrichment assay was utilized to create a custom RNA bait library to capture the regions chr 2: 174286992-174307309, chr 15: 90098150-90200735, and chr 19: 55434878-59128983 in 100 affected COME/ROM individuals (10 pools of 10 individuals per pool). The captured regions were sequenced using 100 base pair (bp) paired end reads on Illumina's HiSeq 2000 platform. Sequence data was aligned to the human reference (Build 37) using Novoalign, and genetic variants were identified using SAMtools mpileup, GATK Unified Genotyper, and Freebayes. Results: Sequencing on the HiSeq 2000 yielded 36.9 Gb of sequence data at an average coverage of 175x per individual. After alignment with Novoalign, quality control measures were performed, excluding duplicates to remove potential PCR bias. Three complimentary variant detection methods, mpileup, Unified Genotyper, and Freebayes, were used to analyze the targeted resequencing data. The use of three variant detection methods allows for confirmation of variants identified by more than one program, along with detection of rare variants by Freebayes which incorporates ploidy of each pool into the analysis. Collectivley, the total number of unique variants detected by the three methods is 12,866. Conclusion: We have successfully performed targeted sequencing of three regions (chromosome 2, 15, and 19) that contain genetic variants influencing COME/ROM. The sequence data generated in this study is of high quality with coverage that will allow for discovery of rare variants. Subsequent analyses will include the prioritization of SNPs with the most relevant functional role in disease pathogenesis. Individual level genotyping and replication is required to determine the causal variant(s). These data will facilitate the discovery of rare and novel variants contributing to the susceptibility of chronic otitis media with effusion.

1972T

Genetic polymorphisms of long pentraxin-3 are associated with polypoidal choroidal vasculopathy and exudative age-related macular degeneration. *J.H. Chen[1,2], Y. Yang[1], Y. Zheng[1], C. Xu[1], W. Chen[1], M. Zhang[1], H. Chen[1,2], C.P. Pang[1,2].* 1) 1. Joint Shantou International Eye Center, Shantou University & the Chinese University of Hong Kong, Shantou, China; 2) 2. Department of Ophthalmology and Visual Sciences, the Chinese University of Hong Kong, Hong Kong, China.

Purpose: To evaluate the association of polymorphisms in the vascular inflammation marker gene long pentraxin-3 (*PTX3*) with polypoidal choroidal vasculopathy (PCV) and exudative age-related macular degeneration (AMD). **Methods**: The study subjects were unrelated and included 68 PCV, 78 AMD patients, and 346 controls recruited at Joint Shantou International Eye Center, Shantou, China. Five single nucleotide polymorphisms (SNP) of *PTX3*, including a SNP rs9289983 at the 5′ upstream, three intronic SNPs rs2305619, rs3845978, rs1840680, and a SNP rs4680367 at the 3′ downstream were selected as tag SNPs based on the linkage disequilibrium (LD) data in HapMap Han Chinese. Genotyping was conducted using Taqman assays, and genotype data were analyzed by Haploview version 4.2 and logistic regression. 10000 Permutation tests were used to correct multiple comparisons. **Results**: The homozygous minor genotype of rs1840680 conferred risk to PCV (OR = 2.84, P = 0.009), but no single SNP association with AMD was detected. A LD decay between rs9289983 and rs3845978 was observed in PCV and AMD. There was significant association of the rs9289983-rs3845978 haplotypes with PCV and combined AMD+PCV (P = 0.001 and 0.025 respectively). The A-T haplotype of rs9289983-rs3845978 conferred the highest risk to PCV (OR = 19.44, P = 0.017), and also high risk to AMD (OR = 10.22, P = 0.002). **Conclusions**: The current study for the first time provided evidence that *PTX3* was associated with PCV and AMD, suggesting a common role played by *PTX3* in the two choroidal neovascularization diseases.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1973F

Contribution of the reading disability risk locus DYX2 and dopamine signaling factors ANKK1/DRD2 to language impairment and brain imaging phenotypes. *J. D. Eicher[1], N. R. Powers[1], L. L. Miller[2], K. L. Mueller[3], J. B. Tomblin[3], S. M. Ring[2], J. R. Gruen[1,4], Pediatric Imaging Neurocognition Genetics Study*. 1) Department of Genetics, Yale University, New Haven, CT; 2) School of Social and Community Medicine, University of Bristol, Bristol, Bristol, United Kingdom; 3) Department of Speech Pathology and Audiology, University of Iowa, Iowa City, IA; 4) Departments of Pediatrics and Investigative Medicine, Yale University, New Haven, CT.

Children with Language Impairment (LI) have deficits in verbal language despite otherwise normal development. LI is a common impairment in children (prevalence: 5-8%) and is frequently comorbid with reading disability (RD)=munexpected difficulty in reading despite adequate opportunity and instruction=mwith 50% of children with LI also developing RD. Past studies have identified genes on the DYX2 locus, including KIAA0319 and DCDC2, as conferring risk to RD. Our lab also recently identified prenatal nicotine exposure as a potential risk factor for both RD and LI, suggesting that nicotine-related pathways may contribute to language development. The overall aims of this study are to determine whether DYX2 and nicotine dependence (ND) genes contribute to language deficits and how these genes may impact structural brain imaging phenotypes. We genotyped 209 SNPs designed to capture common variation in the DYX2 locus, and 31 variants in ND genes (e.g. ANKK1, CHRNA4) in subjects from the Avon Longitudinal Study of Parents and Children (ALSPAC) (n=5579). We tested single marker and haplotype associations of DYX2/ND variants, and then replicated significant SNPs in an independent cohort (Iowa LI) (n=428). Replicated variants were then interrogated using genetic and MRI data from the Pediatric Imaging, Neurocognitive, Genetics (PING) study (n=839). We observed associations of DYX2 genes, including KIAA0319, DCDC2, THEM2, and C6orf62, in ALSPAC (p<0.0004). Associations were also seen in ND genes ANKK1/DRD2 (p<0.0004). These associations with DYX2 and ANKK1 were replicated in the Iowa LI cohort (p<0.006). Variation in ANKK1 was associated with differences in occipital lobe volume (p=0.025). KIAA0319 variants were found to show differences in volumetric asymmetry in inferior temporoparietal and postcentral cortexes (p=0.0007), while DCDC2 variants associated with differences in cortical thickness in various brain regions of interest (p=0.0007). We found associations of DYX2 and ANKK1/DRD2 with LI and language performance with these variants also associating with structural imaging phenotypes. Our investigation demonstrates the influence of DYX2 and nicotine pathways on core language processes and suggests that they contribute to language deficits seen in distinct neurobehavioral disorders.

1974W

The Maternal Folate Hydrolase Gene Polymorphism Is Associated with Neural Tube Defects in a High-risk Chinese Population. *J. Guo, JH. Wang, H. Xie, HZ. Zhao, F. Wang, C. Liu, L. Wang, XL. Lu, YH. Bao, JZ. Zou, GL. Wang, B. Niu, T. Zhang.* Capital Institute of Pediatrics, Beijing, China.

Folate hydrolase 1 (FOLH1) gene encodes intestinal folate hydrolase, which regulates intestinal absorption of dietary folate. Previous studies on the association between polymorphisms rs202676 and rs61886492 and the risk of neural tube defects (NTDs) were inconclusive. A case-control study of women with NTD-affected pregnancies (n=160) and controls (n=320) was conducted in the Chinese population of Lvliang, a high-risk area for NTDs. Prior to performing the experiments, this study was reviewed and approved by the ethic evaluation committee of Capital Institute of Pediatrics, Beijing, China. We genotyped the polymorphic sites rs202676 and rs61886492 and assessed maternal plasma folate and total homocysteine (tHcy). Our results showed that in case group, plasma folate concentrations was 18% lower compared with control group (8.32nmol/l vs 6.79nmol/l, p=0.033) and tHcy concentrations was 17% higher (10.47μmol/l vs 12.65μmol/l, p=0.047). Almost all samples had the rs61886492 GG genotype (99.78%). The result showed that the frequency of GG genotype in rs202676 was significantly higher in group with multiple NTDs than controls (p = 0.030, OR = 2.157, 95% CI, 1.06-4.38). The multiple NTD group showed significantly higher maternal plasma concentrations of tHcy(10.47μmol/l vs 13.96μmol/l, p= 0.024). The GG genotype of rs202676 had a lower maternal folate and higher tHcy concentrations than other genotypes with no significant differences. The result of structural prediction indicated this variation may change the spatial structure of the protein. These results suggest that the maternal polymorphism rs202676 was a potential risk factor for multiple NTDs in this Chinese population. The allele G may affect maternal plasma folate and tHcy concentration. Jin Guo and Jianhua Wang contributed equally to this manuscript. Correspondence: Ting Zhang and Bo Niu.

1975T

Association study of NPR3 polymorphisms with aspirin exacerbated respiratory disease. *J.H. Kim[1], B.L. Park[2], C.S. Park[3], H.D. Shin[1,2].* 1) Sogang Unoversity, Seoul, South Korea; 2) SNP Genetics, Inc., Seoul, South Korea; 3) Soonchunhyang University Bucheon Hospital, Bucheon, South Korea.

Aspirin exacerbated respiratory disease (AERD) is a clinical syndrome characterized by persistent and severe inflammation of the upper and lower respiratory tracts when taken with non-steroidal anti-inflammatory drugs (NSAIDs) including aspirin. Natriuretic peptides have been implicated in modulating asthma and inflammation. Among natriuretic peptide receptors (NPRs), NPR3 has recently been highlighted to have a role in blood pressure. To investigate the association between NPR3 polymorphisms and AERD, a total of 17 common single nucleotide polymorphisms (SNPs), with minor allele frequency (MAF) over 0.05, were genotyped in 163 AERD and 429 aspirin-tolerant asthma (ATA) subjects in a Korean population. Regression analysis showed no associations between NPR3 polymorphisms and FEV1 decline after aspirin provocation in asthmatics (P > 0.05). However, logistic analysis revealed that NPR3 polymorphisms and haplotypes might be potential risk factors for the development of AERD (P = 0.006-0.03). Despite the needs for replications in large cohorts and further functional evaluations, our preliminary findings suggest that NPR3 polymorphisms might be associated with aspirin hypersensitivity in asthma.

1976F

Identification of DNA variants in 11 candidate genes for developmental dyslexia by next generation sequencing. *H. Matsson[1], E. Tiraboschi[1], M. Huss[2], J. Noppola-Hemmi[3], H. Lyytinen[4], P.H.T. Leppanen[5], N. Neuhoff[6], A. Warnke[7], G. Schulte-Körne[8], M.M. Nöthen[9], J. Schumacher[10], M. Peyrard-Janvid[1], J. Kere[1,2,11].* 1) Department of Biosciences and Nutrition, Karolinska Institutet, Huddinge, Sweden; 2) Science for Life Laboratory, Karolinska Institutet, Solna, Sweden; 3) Division of Child Neurology, Department of Gynecology and Pediatrics, Jorvi Hospital, University Central Hospital, Espoo, Finland; 4) Department of Psychology and Child Research Center, University of Jyväskylä, Jyväskylä, Finland; 5) Department of Psychology, University of Jyväskylä, Jyväskylä, Finland; 6) Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University of Munich, Munich, Germany; 7) Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University of Wuerzburg, Wuerzburg, Germany; 8) Department of Child and Adolescent Psychiatry and Psychotherapy, University of Marburg, Marburg, Germany; 9) Department of Genomics, Life & Brain Center, University of Bonn, Bonn, Germany; 10) Institute of Human Genetics, University of Bonn, Bonn, Germany; 11) Department of Medical Genetics, Biomedicum, University of Helsinki, Helsinki, Finland.

Developmental dyslexia is present in 5-10% of the population. Dyslexia manifests with specific reading problems despite normal learning opportunities, intellectual ability and independently of socio-economical status. Previous studies have shown a large genetic burden for dyslexia and at least 10 genetic loci are described. A handful of genes, DCDC2, DYX1C1, KIAA0319, MRPL19, C2orf3 and ROBO1, show significant association replicated in several independent samples. Genome-wide association studies have not been fruitful, probably due to limited power and the exclusion of rare variants. An in-depth analysis of candidate genes in order to find causative variants has not yet been reported. Previously, we used genomic DNA from 100 selected Finnish dyslexics in 10 pools with the aim to identify and analyse all single nucleotide variants (SNVs) in exons of the dyslexia candidate genes C2orf3, CYP19A1, DCDC2, DIP2A, DYX1C1, KIAA0319, MRPL19, PCNT, PRMT2, ROBO1 and S100B. The enrichment was designed to capture 96% of all protein coding and UTR sequences. After next generation sequencing by SOLiD 4, we predicted 350 novel and known SNVs using the CRISP algorithm. Here, we re-sequenced 90% of all individuals using SOLiD 5500 resulting in 271 known SNPs of which 236 SNPs (of 259 totally) overlapped with our first run. For verification, we selected 111 SNVs for Sequenom genotyping in DNA from the same 100 dyslexics and in additional samples consisting of German and Finnish cases/controls and families with dyslexia. The family samples included at least one affected proband and both parents of 153 families (156 affected) and 252 families (419 affected) from Finland and Germany, respectively. Additionally, 92/67 cases/controls of Finnish origin were included. We excluded SNPs with success rate <95% and individuals with >25% missing genotypes. Analysis of single allele association is currently on-going. By using deep sequencing, we can overcome limitations of traditional DNA genotyping for the identification of rare and novel DNA variation. In this study we also screen, for the first time, the PCNT, DIP2A, S100B and PRMT2 genes contained in a 21q22.3 deletion and co-segregating with dyslexia in a father and three sons, and believe that we will identify the gene(s) for dyslexia on chromosome 21. We expect our set of verified SNVs to aid future genotyping efforts and to increase our understanding of the DNA variants contributing to an increased risk for dyslexia.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1977W

Sequencing of *IL10* in Behçet's disease patients. *I. Sousa[1,2], J.M. Xavier[1,2], N. Rei[1,2], F. Barcelos[3], P. Abrantes[1,2], J. Vedes[4], G. Jesus[5], M. Salgado[6], F. Fontes[7], J.V. Patto[3], J. Crespo[8], S.A. Oliveira[1,2].* 1) Instituto de Medicina Molecular/ Faculdade de Medicina da Universidade de Lisboa, Lisboa, Portugal; 2) Instituto Gulbenkian de Ciência, Oeiras, Portugal; 3) Instituto Português de Reumatologia, Lisboa, Portugal; 4) Hospital de Sousa Martins, Guarda, Portugal; 5) Hospital Infante D. Pedro, Aveiro, Portugal; 6) Hospital Pediátrico de Coimbra, Coimbra, Portugal; 7) Hospital Curry Cabral, Lisboa, Portugal; 8) Hospitais da Universidade de Coimbra, Coimbra, Portugal.

Behçet's disease (BD) is a chronic inflammatory disorder classified as a vasculitis, and may involve several organs, such as skin, mucocutaneous membranes (oral and genital aphtae), eyes, joints, lungs, gastrointestinal and central nervous systems. Common variants in *IL10* have been recently associated with BD in two independent genome-wide association studies (GWAS) in the Turkish and Japanese populations and replicated by our group in an Iranian dataset. Despite the GWAS success in identifying single nucleotide polymorphisms (SNPs) that contribute to some complex diseases, the majority of genetic variants contributing to disease susceptibility are yet to be discovered. In fact, it has been argued that these variants are not likely to be captured in current GWAS that focuses on common SNPs. Sequencing enables the identification of new rare variants of small/moderate effect and/or mutations that can be related to the BD phenotype, which in turn promise to help explain the missing heritability after the GWAS era.

50 Portuguese cases with BD were sequenced using the traditional Sanger sequencing for the entire *IL10* open reading frame, including also 1Kb upstream and downstream of *IL10*. Additionally, two conserved regions located in the *IL10* promoter were also sequenced.

We identified so far 24 known single nucleotide variants (SNVs) in our sample with minor allele frequencies (MAF) ranging from 0.01 to 0.40, and three new SNVs in our BD cases. These new SNVs were present in a heterozygous state. One of these new variants is located at intron 3 and was detected in two individuals while the other two are located in the promoter region and were identified in one individual each. The absence of these three novel variants has been confirmed in 50 healthy Portuguese individuals and genotyping of 100 additional healthy Portuguese controls is ongoing to assess if these SNVs are specific of Behçet's disease patients or if they are rare variants not yet reported. Genotyping of these three new variants will also be performed in 300 Iranian BD patients and controls to investigate if they are specific to Portuguese BD patients.

1978T

Variants of XBP1 are associated with vitiligo in the presence of autoimmunity. *R.G. Tarlé[1,2], L.M. Nascimento[2], C.C.S. De Castro[1,2], R.I. Werneck[3], V.M. Fava[1], M.T. Mira[1,3].* 1) Group for Advanced Molecular Investigation, Medical School, Pontifical Catholic University of Paraná, Curitiba, Paraná, Brazil Rua Imaculada Conceição, 1155. Curitiba, Paraná, Brazil.80215-901; 2) Department of Dermatology, Santa Casa de Misericórdia Hospital, Pontifical Catholic University of Paraná, Curitiba, Brazil; 3) School of Health and Biosciences, Pontifical Catholic University of Paraná, Curitiba, Brazil.

Vitiligo is a chronic, multifactorial disease clinically characterized by depigmentation of segments of the skin. Genetic studies have shown familial aggregation of cases of vitiligo, following a non-Mendelian pattern, and different strategies of genetic analysis have resulted in the description of several genes associated with the disease. Interestingly, a number of these genes have also been associated with autoimmune diseases that often co-occur with vitiligo. Chromosome 22q12 has been described as a generalized vitiligo susceptibility locus in Chinese families. A follow-up study resulted in epidemiological and functional data implicating variant rs2269577 of the XBP1 gene located at 22q12 in the control of susceptibility to vitiligo. The observation has been replicated later in a Caucasian population sample. The goal of our study was to advance on the understanding of the nature of the association between vitiligo and XBP1 variants, by (i) replicating the association signal using an independent, family-based population sample, and (ii) investigating the impact of autoimmune co-morbidities over the observed association. Our study was performed using a Southern Brazilian population sample of 596 individuals distributed in 212 trios composed by an affected child and both parents. Unbiased selection of markers was performed in order to capture the entire information of the XBP1 locus. Family based association analysis resulted in positive evidence for association between marker rs2239815 and vitiligo (p=0.04), and a trend towards association with rs2269577 (p=0,09), with the two SNPs in strong linkage disequilibrium (r2=91). A second, stratified analysis revealed that the association signal between vitiligo with rs2269577 (p=0.007) and rs2239815 (p= 0.01) was concentrated in the sub-sample of 42 and 45 trios, respectively, presenting autoimmune co-morbidity. XBP1 encodes a transcription factor that regulates the expression of MHC class II genes. Variants of XBP1 have been associated with Chron's disease and ulcerative colitis. Our results support the hypothesis that at least certain forms of vitiligo are of autoimmune nature; it also suggests the existence of a common mechanism of genetic control shared by a group of autoimmune diseases.

1979F

Genetic association of SLE with CD247 (CD3ζ). *M. Martins[1,2], A.H. Williams[3], M. Alarcon-Riquelme[4], J.M. Anaya[5], S.C. Bae[6], S.A. Boackle[7], L.A. Criswell[8], G.S. Gilkeson[9], D.L. Kamen[9], C.O. Jacob[10], J.A. James[4], J.A. Kelly[4], K.L. Moser[4], R.H. Scofield[4], R.P. Kimberly[11], J.C. Edberg[11], L.M. Vila[12], M. Petri[13], R. Ramsey-Goldman[14], M.E. Comeau[3], J. Ziegler[3], M.C. Marion[3], J.T. Merrill[4], T.B. Niewold[15], B.P. Tsao[16], B.I. Freedman[3], S.B. Glenn[4], J.B. Harley[4], C.D. Langefeld[3], C. Fesel[2].* 1) Neurologic Unit of Clinical Investigation (UNIC), Instituto de Medicina Molecular (IMM), Lisbon, Lisbon, Portugal; 2) Instituto Gulbenkian de Ciência, Oeiras, Portugal; 3) Wake Forest University School of Medicine, Winston-Salem, NC, USA; 4) Oklahoma Medical Research Foundation, Oklahoma City, OK, USA; 5) Universidad del Rosario, Bogota, Colombia; 6) Hanyang University Hospital for Rheumatic Diseases, Seoul, Korea; 7) University of Colorado Denver School of Medicine, Aurora, CO, USA; 8) Department of Medicine, University of California, San Francisco, CA, USA; 9) Department of Medicine, Medical University of South Carolina, USA; 10) University of Southern California School of Medicine, Los Angeles, CA, USA; 11) Department of Medicine, University of Alabama, Birmingham, USA; 12) University of Puerto Rico Medical Sciences Campus, San Juan, Puerto Rico; 13) Johns Hopkins University School of Medicine, Baltimore, MD, USA; 14) Northwestern University, Feinberg School of Medicine, Chicago, IL, USA; 15) University of Chicago, Section of Rheumatology, Chicago, IL, USA; 16) Department of Medicine, University of California, Los Angeles, USA.

Downregulation and replacement of the CD3z chain, leading to altered TCR signaling, is a known characteristic of T-cells in SLE patients. However, genetic variation in the CD247 locus that encodes CD3ζ is so far not well documented for association with SLE. Independent replication of association findings is clearly necessary to validate the most significant results in independent and large datasets. Our aim was to test the association of the gene encoding the CD247 (CD3ζ) molecule previously identified as a Systemic Lupus Erythematosus (SLE) risk factor in functional and association studies. 44 contiguous SNPs in CD247 were typed in 8,922 SLE patients and 8,077 controls from four ethnically distinct populations. Population stratification was assessed using a panel of 347 ancestral-informative markers. The strongest associations of CD247 with SLE were found in the Asian samples (11 SNPs in intron 1, $4.99x10\text{-}4<p<4.15x10\text{-}2$). Imputation analysis showed evidence of association ($p<0.005$) at additional 5 SNPs. Examination of haplotypes within the Asian samples identified a five-marker haplotype (rs12141731-rs2949655-rs16859085-rs12144621-rs858554; G-G-A-G-A) that exceeded single SNPs in significance. Cross-ethnic meta-analysis assuming an additive genetic model and adjusted for population proportions showed a significant association after Bonferroni correction for multiple testing, with rs704848 being the most associated single SNP. Our study independently confirmed and extended the association between SLE and CD247. These associations give new directions to understand the molecular mechanisms of disease and to pursue new therapeutic targets.

1980W

Estrogen receptor beta genetic variants are associated with intraocular pressure elevation in women. *F. Mabuchi[1], Y. Sakurada[1], K. Kashiwagi[1], Z. Yamagata[2], H. Iijima[1], S. Tsukahara[1].* 1) Dept Ophthalmology, Univ Yamanashi, Chuo, Yamanashi, Japan; 2) Dept Health Sciences, Univ Yamanashi, Chuo, Yamanashi, Japan.

Purpose: We previously reported that the estrogen receptor beta (*ESR2*) genetic variant was associated with an intraocular pressure (IOP) elevation in female patients with primary open-angle glaucoma (POAG). However, the association between the *ESR2* genetic variant and the maximum IOP was identified for only 1 genetic variant (rs1256031). The aim of this study was to confirm the association between the *ESR2* genetic variants and the IOP in women.

Methods: Eight single nucleotide polymorphisms (SNPs) covering the *ESR2* whole region (rs1256031, rs4986938, rs1256049, rs1273916, rs1256120, rs928554, rs1256059, and rs944461) were genotyped in 326 Japanese female patients, including 205 patients with POAG and 121 patients without glaucoma. The participants were divided to the normal (IOP < 20mmHg, n = 190) and high (IOP ≥ 20mmHg, n = 136) IOP groups. The allele frequency differences between normal and high IOP groups were estimated using the Fisher's exact test. In addition, a logistic regression model was used to study the effects of the 8 genotyped SNPs alleles to compare the normal IOP group with the high IOP group adjusted for age and refractive error. Linkage disequilibrium among 8 genotyped SNPs to compare the normal IOP group with the high IOP group was evaluated using the Haploview program.

Results: There were significant differences in the allele frequencies of 4 SNPs (rs928554, rs944461, rs1256059, and rs1256031) between the normal and high IOP groups. (P = 0.012, P = 0.023, P = 0.012, and P = 0.0006 respectively, Fisher's exact test) An increased risks of high IOP were confirmed with the risk alleles of these 4 SNPs (P = 0.0019 to 0.031, odds ratio 1.43 to 1.67) by logistic regression analysis adjusted for age and refractive error. These 4 SNPs showed significant association were in high linkage disequilibrium, based on the findings obtained using the Haploview program.

Conclusions: This study confirmed that the *ESR2* genetic variants were associated with IOP elevation in women.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1981T

Expression of Alternatively Spliced Transcripts of MAP2K4 Gene in Rheumatoid Arthritis. *K. Shchetynsky, M. Ronninger, L. Padyukov.* Rheumatology Unit, Department of Medicine Solna, Karolinska Institutet./ Karolinska University Hospital, Stockholm, Sweden.

BACKGROUND: MAP2K4 gene encodes a mitogen activated protein kinase kinase 4 (MKK4), important for optimal activation of JNK1-3 and p38 - the two members of the MAP kinase family previously implicated in several rheumatic diseases, including rheumatoid arthritis (RA). Recently, MAP2K4 locus has been identified as being in statistical interaction with the major risk factor for RA - HLA-DRB1 shared epitope (SE) - in patients with anti-citrullinated protein antibody (ACPA). In this study we explore the variety of alternatively-spliced transcripts for MAP2K4 gene and investigate differential expression profile and phenotype associations of MAP2K4 splice variants in RA patients and healthy individuals. **METHODS:** We performed an exploratory study to detect novel splice variants of MAP2K4 in total RNA from human peripheral blood cells (PBCs). Variants of interest were amplified by conventional PCR and cloned using bacterial expression systems; sequence data were obtained for selected variants. The relative expression of MAP2K4 transcripts was accessed by TaqMan real-time PCR in PBCs from 44 RA patients and 44 controls of Caucasian ancestry and analyzed against available genotypic and phenotypic data. **RESULTS:** We detected a novel splice variant of human MAP2K4, lacking one of the exons, which form the catalytic domain of MKK4. The MAP2K4 splice forms were differentially expressed in PBCs from 88 RA cases and controls. Additionally, within the group of RA patients, a correlation was observed between MAP2K4 variants expression and phenotypical data for ACPA, rheumatoid factor, and SE. **CONCLUSIONS:** We described a novel splice variant of human MAP2K4, which may be potentially translated as a MKK4 isoform, prompting further functional studies. Our results show differential expression of the canonical and alternatively spliced MAP2K4 mRNAs in peripheral blood of RA patients compared to healthy controls. The expression ratio of two studied MAP2K4 transcripts may also reflect other genetic and phenotypic features of RA. These data imply that MAP2K4 splicing and expression profile may be associated with RA pathogenesis and should be assessed as a potential biomarker.

1982F

RNA-SEQ survey of human lymphoid and myeloid cells - a resource for identifying causal genes in validated immune-mediated disease loci. *S. Foisy[1], C. Beauchamp[1], A. Alikashani[1], F. Latour[1], M. Ladouceur[1], S. Lessage[2,3], J.D Rioux[1,2].* 1) Montreal Heart Institute, Montreal, Canada; 2) Faculte de Medecine, Universite de Montreal; 3) Centre de recherche, Hopital Maisonneuve-Rosemont.

In the last few years we have witnessed an explosion in the number of validated loci for complex diseases, in large part due to the success of genome-wide association studies (GWAS). This success has been most notable for immune mediated diseases where >200 genetic risk factors have been discovered, some that are phenotype-specific and others that are common to more than one disease. The unbiased nature of GWAS has led to the discovery of biological pathways not suspected to play a role in disease pathogenesis, such as the ATG16L1 and IRGM loci and autophagy in Crohn's disease (CD). Although this progress has been significant, for many of these loci we do not have identified the causal gene. While targeted sequencing of genes within validated loci has begun to help identify the causal gene(s), complementary approaches are clearly necessary. As importantly, for most risk loci we only have very rudimentary functional information (or none at all) on the genes within associated regions. In order to unravel the role played by the different genes from each region in disease pathogenesis (disease mechanism) it is critical to place the newly discovered genes in their proper biological context; more specifically, classify them within functional pathways in a cell-specific and environmental context. We hypothesize that disease-causing genes within risk loci are involved in a limited number of biological pathways that are key to the development of disease and will likely play a role in different immune cell subpopulations. Importantly, cells of the immune system are highly specialized and thus knowing the patterns of expression can provide significant insights into their potential functional role. We therefore are building a survey of global gene expression in circulating immune cells (of lymphoid and myeloid origin), isolated by fluorescence-activated cell sorting (FACS) from healthy and diseased individuals. The resulting RNA-SEQ sequence data allow us to: (1) identify all genes expressed in a given cell type; (2) identify putative new exons, isoforms and new genes; and (3) observe if any of the genes in the associated loci are differentially expressed in normal and/or diseased cells. We will present the approach and results from nine different cell types purified from eight different healthy individuals.

1983W

Preliminary data suggest an upstream IRX1 sequence variant identified in a family with kyphoscoliosis disrupts the wild-type expression pattern in zebrafish. *C.M. Justice[1], K. Bishop[2], B. Carrington[2], P. Cruz[3], K. Swindle[4], R. Sood[2], N.H. Miller[4], A.F. Wilson[1], NISC Comparative Sequencing Program.* 1) Genometrics Section, IDRB, NHGRI, NIH, Baltimore, MD; 2) Zebrafish Core, GMBB, NHGRI, NIH, Bethesda, MD; 3) NISC, NHGRI, NIH, Rockville, MD; 4) University of Colorado, The Children's Hospital, Denver, CO.

When idiopathic scoliosis, a structural lateral curvature of the spine ≥ 10;dg in individuals who are otherwise phenotypically normal, is associated with an increase in the normal thoracic kyphosis in the sagittal plane (≥ 40;dg), the condition is referred to as kyphoscoliosis. A previous model-independent linkage analysis of 7 families (53 individuals) with ≥ 2 family members with kyphoscoliosis identified candidate regions on 2q22, 5p15, 13q and 17q11, and analyses of single nucleotide polymorphism (SNP) markers narrowed the region on 5p15 to 3.5 Mb. This region contains only 3 genes, IRX1, IRX2 and IRX4, all members the Iroquois (IRX) gene family that codes for homeoproteins. The most significant linkage peak was in IRX1. The exons from IRX1, IRX2 and IRX4 were sequenced in all 53 individuals, but no functional sequence changes were identified. In this study, the highly conserved non-coding regions (HCNRs) 500 kb upstream and downstream from IRX1, IRX2 and IRX4 were sequenced in these 53 individuals in search of cis-regulatory elements. Quantitative association analysis revealed several SNPs with p-values < 0.01. The allele of 9 sequence variants (SVs; 2 of them novel) differed from the reference allele in a heterozygous state in 6 of 7 affected individuals in one family, and did not differ from the reference allele in all other individuals genotyped. One of these SVs is in a HCNR that functions as an enhancer in mice (enhancer.lbl.gov) and is 413 kb upstream from IRX1. We observed this SV in a heterozygous state (C/T) in 1 out of 90 controls. In order to determine if variation at this SV causes functional changes in vivo, an HCNR surrounding this SV (198 bp in size) was identified, amplified and cloned into a zebrafish enhancer detector (ZED) vector for Danio rerio transgenesis with the wild-type allele (C), the T allele, and the A allele reported by 1000 Genomes (MAF = 0.008). Preliminary data using transient transgenesis showed that the wild-type (C) allele drove strong expression in the midbrain and along the neural tube, while the fragments with the T and A alleles did not show a consistent definite expression pattern. We are in the process of generating stable transgenic lines for each allele. Our preliminary data suggest that sequence variation in a 198 bp HCNR disrupts the wild-type expression pattern in zebrafish and may be involved in the expression of kyphoscoliosis in humans.

1984T

Phenotypic analysis of peptidylarginine deiminase type 4 knockout mice. *A. Suzuki[1], Y. Kochi[1], H. Shoda[2], K. Fujio[2], E. Kanno[1], T. Sawada[3], R. Yamada[1, 4], K. Yamamoto[1, 2].* 1) Ctr for Genomic Med., RIKEN, Yokohama City, Japan; 2) Department of Allergy and Rheumatology, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan; 3) Department of Rheumatology, Tokyo Medical University, Tokyo, Japan; 4) Center for Genomic Medicine, Kyoto University, Kyoto, Japan.

Rheumatoid arthritis (RA) is well-known as an autoimmune disease and is a chronic inflammatory disorder characterized by the destruction of multiple joints. Many genome wide association studies were performed and multiple RA-susceptibility loci and autoimmune-susceptibility loci have been identified. These studies suggested that multiple genes and its functions were related with disease causing and development. Previouly, peptidylariginine deiminase type 4 (PADI4) was identified as a susceptibility gene for RA in a Japanese population by case-control association study (Ref 1). PADI4 is a member of the PADI gene family and converts arginine residue (peptidylarginine) to citrulline residue (peptidylcitrulline). PADI4 is highly expressed in bone marrow, macrophages, neutrophils and monocyte. Peptidylcitrulline is an interesting molecule in RA, because it is an antigen of ACPA and only PADs (translated protein from PADI genes) can provide peptidylcitrullines, via modification of protein substrates. To evaluate the importance of PADI4 gene in the progression of RA, we generated Padi4 -/- DBA1J mice by speed congenic method. We used Padi4-/- mice to show that PAD4 is affected to development and progression of collagen induced arthritis (CIA), well known as an RA model animal. Padi4-/- DBA1Jand WT mice were immunized with bovine type II collagen (CII) for CIA. Clinical disease score was significantly reduced. In Padi4-/- mice sera, the concentrations of serum anti-CII IgM, IgG, and TNF-alpha also decreased significantly rather than in WT mice. Resulting from these studies, we suggested that Padi4 enhanced collagen-initiated inflammatory responses. 1) Suzuki, A. et al Nat. Genet.34, 395-402 (2003).

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1985F

Rs34424835 PGRN gene variation a possible modifier in Frontotemporal lobar degeneration (FTLD). *E. Vitale[1], A. Iuliano[1,2], A. Polverino[1], V. Agosti[2,5], C. Vitale[2,5], A. Postiglione[4], P. Sorrentino[3], S. Pappatà[6], G. Milan[3], G. Sorrentino[2,5].* 1) Cybernetics, CNR, Pozzuoli (NA), Italy; 2) Department of Institutional Studies Parthenope University Naples, Italy; 3) Geriatric Center "Frullone" ASL Napoli 1 - Naples, Italy; 4) Department of Clinical and Experimental Medicine, University of Naples "Federico II", Naples, Italy; 5) Neurological Institute for Diagnosis and Treatment Hermitage Capodimonte, Naples, Italy; 6) Institute of Bioimaging and Biostructures, CNR, Naples, italy.

Frontotemporal lobar degeneration (FTLD) is a degenerative disorder of the frontal and anterior temporal lobes and is a common form of dementia affecting individuals younger than 65 years. The syndrome is clinically characterized by initial behavioral disturbances, followed by cognitive decline, leading to dementia and death within a median of 7 years from symptom onset. FTLD usually segregates as an autosomal dominant trait. Linkage studies have identified several chromosomal loci and a number of genes, including microtubule-associated protein tau (MAPT), progranulin (PGRN) in about 50% of familial cases, valosin containing protein (VCP) and charged multivesicular body protein 2B (CHMP2B) to be associated with the disease. Recently a large hexanucleotide (GGGGCC) repeat expansion was identified in the first intron of C9ORF72 on chromosome 9p21 in amyotrophic lateral sclerosis-frontotemporal dementia (ALS-FTD). This locus contains one of the last major unidentified autosomal-dominant genes underlying these common neurodegenerative diseases. FTLD has three clinical subtypes: fontotemporal dementia (FTD), progressive nonfluent aphasia (PNFA), and semantic dementia (SD). We have identified a large pedigree from the Southern Italy region segregating FTLD in 11 patients over three generation with a dominant inheritance pattern. The family includes members with different clinical phenotypes manifesting neuropsychological and behavioral heterogeneity. These comprise a behavioral variant of frontotemporal dementia (bvFTD) with apathetic or disinhibited behavioral syndrome and executive, attentional and memory impairment. Parkinsonism and epilepsy are also common during the illness but rarely at onset. We screened the 5 remaining patients who were still alive from the 11 identified members of the pedigree for MAPT and PGRN genes. Specifically, we sequenced all exons, exon-intron boundaries and 5′ and 3′ regulatory regions in search for mutations in the exon areas. We determined that three out of the five patients revealed the RS34424835 variation on PGRN exon 5, which could be associated with disease severity and early onset expression. This suggests that PGRN may be a possible modifier of the course of the disease in FTLD patients.

1986W

Proteomic analysis of Shank3 over-expression in mouse neuroblastoma cultures showed differentially expressed proteins involved in glycolysis, cytoskeleton,biosynthetic and cell cycle processes,and ubiquitin-proteasome. *N. Zhong[1,2,3,4], W. Ju[1], W. Yan[2,3], X-L. Zhao[2,3], E.C. Jenkins[1], W.T. Brown[1], Y. Wang[4], J.H. Zhou[2,3].* 1) Dept. Human Genetics, New York State Institute for Basic Research in DD, Staten Island, NY; 2) Peking University Center of Medical Genetics, Beijing, China; 3) Dept. Medical Genetics, Peking University Health Science Center, Beijing, China; 4) Shanghai Municipal Children's Hospital Affiliated to Shanghai Jiaotong University, Shanghai, China.

Shank proteins assemble glutamate receptors with their intracellular signaling apparatus and cytoskeleton at the postsynaptic density. Shank expression is sufficient to induce functional dendritic spine synapses in aspiny neurons. In our study, overexpression of Shank3 in neuroblastoma cell cultures (Neuro-2A) followed by two-dimensional difference gel electrophoresis (2D-DIGE) analysis provided further understanding on the mechanism of Shank3 in forming spines and facilitating synaptogenesis. In the Shank3 overexpression cell model, proteins related to glycolysis that is a metabolic process to produce energy like ATP, such as Pkm2, Pgk1, Eno1, Pgam1 and Ldha were found to be up-regulated. Cytoskeleton protein γ-actin, tubulin alpha-1B and microtubule organizing center Cct5 were also found to be up-regulated, which further confirms that Shank3 can recruit cytoskeleton and microtubule protein to increase the number of spines. Another up-regulated protein was Phgdh that is involved in the L-serine biosynthetic and cell cycle processes. A down-regulated protein, Uchl1, was found, which is related to the ubiquitin-proteasome system. Our proteomic study of Shank3 over-expression in mouse Neuro-2A cells has shed new light on its function in spinogenesis and synaptogenesis.

1987T

Transcription Factor Profile of a Single Nucleotide Polymorphism located -224 A/G in Neuropeptide Receptor Y2 (NPY2R) Gene in Predisposition to Hypertension. *E. Albino[1], K. Sugimoto[2], T. Katsuya[2], A. Deng[3], J. Dutil[1].* 1) Department of Biochemistry, Ponce School of Medicine and Health Sciences, Ponce, PR; 2) Department of Geriatric Medicine, Osaka University Graduate School of Medicine, Suita, Osaka, Japan; 3) Molecular Genetics Laboratory, University of Montreal Research Center CHUM, Montreal, QC, Canada.

Essential hypertension (HTN) is a multifactorial disease under the complex regulation of genetic and environmental factors. Previous work in animal models and human populations identified Neuropeptide Y receptor 2 (NPY2R) as a candidate gene for HTN. In two independent Japanese populations, the GG genotype single nucleotide polymorphism (SNP) located 224 bp upstream of the transcription start site was associated with an increased risk of HTN. The aim of this study was to assess the effect of the A to G substitution in the promoter of NPY2R on the expression levels of this gene and to predict the transcription factors (TF) binding sites for each sequence using TESS and the ENCODE project. Human Embryonic Kidney (HEK) 293 cells were transfected with pGL4.10 DNA constructs containing the promoter of NPY2R with either an A or G nucleotide at the -224 position upstream of the luciferase reporter gene. Interestingly, the luciferase activity induced by the NPY2R promoter with the G allele in position -224 was reduced by 34% compared to the promoter containing an A in that position. Differences in the level of relative Luciferase activity between A and G constructs were statistically significant ($p < 0.01$). When sequences were analyzed using TESS, the two sequences containing either an A or a G were predicted to bind different transcription factors. The ENCODE project proposes CTCF as a functional element that binds to the -224 SNP in the NPY2R gene. The data provide strong evidence for a functional role of NPY2R in genetic predisposition to HTN.

1988F

Functional assessment of CRY2 Type II Diabetes variants using a targeted circadian Luciferase assay. *J. Lane[1, 2, 3], A. Tare[1,3], C. Cowan[4], R. Saxena[1,2,3].* 1) Center for Human Genetic Research, Massachusetts General Hospital, Boston, MA; 2) Department of Anesthesia, Massachusetts General Hospital, Boston, MA, 02114, USA; 3) The Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA; 4) Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, Massachusetts 02138, USA.

Genome-wide association studies revealed a link between Type II Diabetes and variation in circadian genes. SNPs in CRY2, a core component of the circadian clock, are associated with increased fasting glucose levels and risk of Type II Diabetes. The means by which variation in CRY2 influences diabetes risk is currently unknown. In this study we utilize genome editing to create the first targeted human circadian Luciferase reporter. Mediated by the targeting specificity of TALENs, we introduce a Luciferase cassette into the cellular genome in the stop codon of PER2. We then utilized the circadian Luciferase cell lines to monitor changes to the cellular phase, amplitude, and period induced by CRY2 variants. We further utilized genome editing to assess the functional consequence of CRY2 variation in an isogenic cell line. Homologous recombination was used to create cell line homozygous for the major allele, minor allele, or heterozygous at causal CRY2 SNPs. Moving forward we will monitor the Luciferase output of CRY2 variant cell lines over 6 days to assess changes to circadian phase, period, or amplitude.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1989W

Modification of Endothelial Progenitor Cell Recruitment by Tyrosinase in Oxygen Induced Retinopathy. *B.E. O'Bryhim[1,2], R.S. White[1], R.C.A. Symons[1,2].* 1) Department of Ophthalmology, University of Kansas Medical Center, Kansas City, KS; 2) Molecular and Integrative Physiology, University of Kansas Medical Center, Kansas City, KS.

Retinopathy of prematurity (ROP) is a common cause of blindness in children. In addition to gestational age at delivery, birth weight, and overuse of oxygen in the nursery, genetic factors are thought to play a role in disease development. Similarly, inbred strains of mice revascularize at different rates after hyperoxic vascular damage in the murine oxygen-induced retinopathy (OIR) model of ROP. A mapping cross between albino BALB/cByJ mice, which revascularize more rapidly, and pigmented C57BL/6ByJ mice, which revascularize more slowly, identified a quantitative trait locus (QTL) on chromosome 7 (LOD = 8.9). This study also found that albino F_2 mice are more resistant to OIR than their pigmented littermates. Albinism in the cross is caused by homozygous loss of function of tyrosinase, which is located within the identified QTL. Thus we sought to determine whether tyrosinase may be one of the quantitative trait genes (QTGs) underlying the QTL, and whether control of recruitment of endothelial progenitor cells (EPCs) into the retina may be part of the mechanism by which tyrosinase influences avascular area. The *B6(Cg)-Tyr$^{c\text{-}2J}$*/J ("*c-2J*") mouse strain is derived from the C57BL/6J strain with a spontaneous loss of function mutation in tyrosinase that has been bred to homozygosity. Pups from both strains were exposed to the OIR model and their retinal avascular area was measured immediately upon return to normoxia on P12, or four days later on P16. Tyrosinase-deficient *c-2J* mice were found to revascularize more rapidly after return to normoxia than do C57BL/6J mice ($p=2.5x10^{-10}$). *c-2J* mice were also found to recruit more AC133$^+$ endothelial progenitor cells (EPCs) to perivascular areas in the retina ($p < 0.01$), which was associated with greater numbers of CD45$^-$/AC133$^+$/CD34$^+$ EPCs both in the circulation ($p = 0.02$) and in the bone marrow ($p = 0.02$). As these strains are genetically identical except for a loss of function mutation in tyrosinase, the data confirm the role of tyrosinase as a QTG in OIR revascularization and suggest that tyrosinase activity modifies angiogenic repair mechanisms, including EPC recruitment.

1990T

Understanding The Impact of Body Mass Index and Associated SNPs on Serum Metabolites. *J. Kumar[1], R. Karlsson[1], J. Prince[1], M. Hong[1], C. Broeckling[2], J. Prenni[2,3], E. Ingelsson[1], F. Wiklund[1].* 1) Dept of Medical Epidemiology and Biostatistics, Karolinska Institute, Stockholm, Sweden; 2) Proteomics and Metabolomics Facility, Colorado State University, Fort Collins, USA; 3) Department of Biochemistry and Molecular Biology, Colorado State University, Fort Collins, USA.

Introduction: Obesity is a serious public health problem of increasing prevalence causing various disorders such as type 2 diabetes and cardiovascular diseases. Body mass index (BMI) is used as a surrogate marker to assess obesity. The present study integrating metabolomics and genomics data was performed to assess the impact of BMI and associated single nucleotide polymorphisms (SNPs) identified in another GWAS on various serum metabolites. Materials and Methods: Two case-control studies consisting of prostate cancer patients (diagnosed between year 2001 and 2003) and controls (all men) were recruited utilizing Swedish Cancer Register and Swedish Population Register, respectively. The first study had 275 cases and 182 controls (discovery group), while the second had 514 cases only (replication group). Ultra performance liquid chromatography coupled with mass spectrometry was employed for metabolic profiling in these individuals. Linear regression analysis was performed for metabolite peaks testing their association with BMI and adjusting for potential confounding factors like seasonal variation, storage, handling and fasting time, age and prostate cancer status. Significant metabolites were further studied for their association with BMI-associated SNPs. Results: Metabolite concentrations were log-transformed to normalize their distribution. 203 metabolite peaks were found to be significantly associated with BMI in the discovery stage (P<8.15x10-6, correcting for multiple testing), and of those, eighteen could be robustly replicated in the replication sample (P<2.4x10-4). Eighteen peaks corresponded to nine unique metabolites and four were putatively identified from these nine. The four peaks represented monoacylglycerol, diacylglycerol and two phosphocholine (34 and 36C). A total of 21 SNPs available (out of 32 from Speliotes EK et al Nat Genet 2010;42(11):937-48) associated with BMI or related traits were tested for their association with four putatively identified peaks. No SNP reached statistical significance after adjusting for multiple correction. Conclusion: In conclusion, four identified fatty acids and five unknown metabolites in an untargeted metabolomics experiment were found to be significantly associated with BMI. BMI associated genetic variants did not show association with these serum metabolites. Further analysis is underway to explore the impact of BMI on ratios between various metabolites to dissect obesity-related metabolic pathways.

1991F

Biological processes of human development and genome-wide linkage disequilibrium. *I. Culminskaya, A. Kulminski, A. Yashin.* Duke University, Durham, NC.

Recently (Age, 2012, PMID: 22282054) we documented two clusters of extensive intra- and inter-chromosomal linkage disequilibrium (LD) in the Framingham Heart Study families comprised of parents and their children. Both clusters were observed on two independent arrays, the Gene-Chip Human Mapping 500K (250K Sty and 250K Nsp) and the 50K Human Gene Focused arrays. The LD has unlikely been generated by stochasticity, population or family structure, or mis-genotyping. Alleles of SNPs in strong inter-chromosomal LD were inherited as an intact unit which was associated with complex heritable phenotype of premature death. The analyses suggested involvement of a large number of genes for SNPs in LD in fundamental biological processes in an organism. In this presentation we report on significant enrichment of genes for SNPs in genome-wide intra- and inter-chromosomal LD in Gene Ontology (GO) biological processes. Annotation of two (partly overlapping) clusters of SNPs in LD (r^2=0.02-1.0) with minor allele frequency 0.009-0.499 results in two reference sets of 1697 and 4637 genes. These gene sets have been categorized into biological processes using ArrayTrack and GoMiner classification tools. The evidences on enrichment have been validated by comparison of the results for the reference sets with those for randomly generated sets of the same number of genes. The analyses using the Fisher's exact test and the False Discovery Rate (FDR) approach reveal highly significant enrichment of GO terms linked to developmental process. The strongest associations have been observed for nervous system development/axonogenesis/axon guidance ($p<10^{-4}$, $FDR<10^{-4}$ for each category) for both sets and for cardiovascular system development for the largest set ($p\leq10^{-4}$, $FDR\leq0.007$). Both gene sets showed highly significant enrichment of biological processes which are tightly linked with neural development including cell differentiation, morphogenesis, adhesion, projection organization, signaling, chemotaxis ($p<10^{-4}$, $FDR<10^{-4}$ for each category). The results provide evidences on significant role of functional interactions and major biological processes of human development in generation of genome-wide LD.

1992W

The Type 1 Diabetes gene *CLEC16A* functions in NK cells to restrain secretory functions including cytokine release and cytotoxicity. *M. Bakay, R. Pandey, A. Mohamed-Hadley, P. Banerjee, H. Hakonasron.* Center for Applied Genomics, Children's Hosp Philadelphia, Philadelphia, PA.

Type 1 Diabetes (T1D) is a multifactorial childhood disease with a strong genetic component. Several GWAS have shown association of CLEC16A with T1D and other autoimmune diseases. CLEC16A has homology domains that identify it a C-type lectin, however the protein has no known function yet. The highest levels of CLEC16A expression were identified in NK cells, which are required for development of T1D. Therefore, defining the role of CLEC16A in NK cells may provide insight into the pathogenesis of T1D. We reported previously that protective alleles of CLEC16A [A/A] are associated with higher levels of mRNA. Thus, we hypothesize that CLEC16A functions in NK cells to restrain cytotoxicity after activation.

To address the role of CLEC16A in T1D we studied consequences of knockdown and overexpression of this protein in NK cell lines. Using a retroviral expression system, we created NK cell lines stably over-expressing the canonical form of CLEC16A with a GFP expression reporter. Expression was validated by Western blot analysis. Cytotoxicity assay revealed decreased killing in YTS-CLEC16A overexpression cell line. Optimized protocols using CLEC16A siRNA mediated knockdown enabled a 70% reduction in CLEC16A protein levels in NK cells and also resulted in a 35% increase in cytotoxicity compared to cells receiving control siRNA. Search for association partners of CLEC16A by immunoprecipitation (IP) pull down using anti-CLEC16A antibody (AB) with following protein sequencing revealed alpha-Centractin (ACTR1A) - an actin-related protein. ACTR1A encodes a 43kD subunit of dynactin, a macromolecular complex. Dynactin binds to both microtubules and cytoplasmic dynein and is involved in a diverse array of cellular functions, including ER-to-Golgi transport, the centripetal movement of lysosomes and endosomes. Preliminary microscopy data shows CLEC16A co-localization with tubulin. CLEC16A association with alpha-Centractin suggests it may impact on the process of translocation of the microtubule-organizing center (MTOC) and focused secretion of effector molecules to the synapse in NK cells. Our results indicate that CLEC16A may serve a role in restraining NK cell function. Studies are in progress to validate the proposed mechanism. Our improved understanding of this novel T1D-linked gene and the protein it codes may lead to new therapeutic interventions aimed at T1D prevention.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1993T

Influence of Ethnicity with Type 2 Diabetes in Association of UCP2 -866G/A, PGC1α (Gly 482 Ser) and SIRT1 -1400T/C Polymorphisms in North Indian Punjabi Population Groups. *A. Bhanwer[1], N. Kaul[1,2], R.N.K. Bamezai[2].* 1) HUMAN GENETICS, GURU NANAK DEV UNIVERSITY, AMRITSAR, PUNJAB, India; 2) National Centre of Applied Human Genetics, School of Life Science, Jawaharlal Nehru University, New Delhi 110067, India.

Type 2 Diabetes (T2D) involves complex disorder involving defects in insulin secretion and insulin action pathways. Several studies have reported the role of UCP2 -866G/A, PGC1α (Gly482Ser) and SIRT1 -1400T/C polymorphisms in development of T2D involving insulin secretary pathway. The association of above mentioned polymorphisms was investigated in 859 Type 2 diabetics and 954 normal healthy control subjects from Punjab. India exhibits different ethnic population structure comprising various caste groups existing from ages which have mostly remained endogamous even today. Thus to assess the impact of ethnicity, independent association of these polymorphisms was carried out in Brahmins, Backward Castes (BCs), Jat Sikhs, Khatris, Rajputs, Banias and Scheduled Castes (SCs). Binary logistic regression was applied to test the association after adjusting for age, sex and BMI. In total population UCP2-866AA [p=0.0001, OR-1.46 (1.2-1.78)] and PGC1α AA [p=0.000002, OR-1.46 (1.20-1.78)] significantly provided risk. Caste based stratification revealed UCP2 -866 A/ AA provided risk only in Banias [p=0.000004, OR- 4.7×10-6, OR- 4 (2-5.9)], BCs [p=0.01, OR- 2.2 (1.1-4.2)] and SCs [p=0.0001, OR- 1.9 (1.16-3.2)] while PGC1α A/AA provided risk in BCs [p=0.0003, OR- 3.3 (1.7-5.5)], Jat Sikhs [p=0.0003, OR- 1.7 (1.20-2.4)] and Khatris [p=0.03, 1.51 (1.02- 2.27)]. SIRT1 -1400T/C polymorphism does not seem to provide risk either in the total population or in any of the caste groups except Khatris [p=0.004, OR- 2.4 (1.3-4.5)] of Punjab. The present investigation concludes that differential pattern of association of polymorphisms is observed for different caste groups, suggesting the putative role of ethnicity. Thus, for risk calculation and proper medical intervention, knowledge of the ethnicity and nature of variation in risk factors needs serious attention.

1994F

Allelic variation in the protein stability of HLA shapes genetic association of HLA with type 1 diabetes. *H. Miyadera[1], J. Ohashi[2], K. Tokunaga[1].* 1) Department of Human Genetics, The University of Tokyo, Tokyo, Japan; 2) Molecular and Genetic Epidemiology, University of Tsukuba, Japan.

Among many disease-susceptibility loci, human leukocyte antigens (HLA) provide the most strong genetic risk for autoimmune and variety of immune disorders, but underlying mechanisms have not been fully explained by allelic variations of HLA in antigen presentation. We hypothesized that functional variation of HLA other than peptide binding spectrum might contribute to genetic association of HLA with autoimmune diseases. In the present study, we analyzed the stability of HLA class II (HLA-DQ) proteins and their contribution to genetic risk for type 1 diabetes (T1D). Through a newly devised protein stability assay, we measured the stability of DQ proteins that are encoded by major DQ alleles, including 10 alleles of DQA1 and 14 alleles of DQB1, that generate 140 different DQA1-DQB1 heterodimer combinations in total. We demonstrate over 100-fold of stability differences among DQ allele products. Polymorphic residues that regulate the intrinsic stability of DQ protein have been also identified through mutagenesis and by utilizing protein crystal data of DQ. Using case-control data for T1D of European and East Asian populations, we identified that DQ protein stability contributes to the genetic risk/protection for T1D. These findings indicate that the stability of HLA protein determines genetic risk for T1D. Our study provides a novel mechanistic basis of HLA associations with autoimmune diseases.

1995W

Functional characterization of gene regulatory elements associated with epilepsy. *R.Y. Birnbaum[1,2], Y. Zhang[3], C. Wei[3], N. Ahituv[1,2].* 1) Bioengineering & Therapeutic Sci, UCSF, San Francisco, CA; 2) Institute for Human Genetics, University of California San Francisco, San Francisco, CA, USA; 3) DOE Joint Genome Institute, Walnut Creek, CA, USA.

Epilepsy is a complex and heterogeneous disease making it difficult to precisely diagnose and provide effective treatments. Infantile spasms (IS) is an uncommon epilepsy syndrome that typically begins in infancy and is associated with ventral forebrain development and forebrain synapse function. Since a major cause of complex diseases, such as epilepsy, could be mutations in gene regulatory elements, we set out to identify these elements in the mouse forebrain at embryonic day 16.5 (E16.5) which could be associated with IS. Using chromatin immuneprecipitation followed by deep sequencing (ChIP-Seq) with enhancer marks (H3K27ac, p300, RNA-Poll2), we identified active enhancer candidates in the mouse E16.5 forebrain. In addition, using chromatin interaction analysis followed by paired-end tag sequencing (ChIA-PET) on mouse E16.5 forebrain we are determining the physical interactions of these enhancer candidates and IS-associated genes. Several enhancer candidates are then tested for their enhancer activity using both zebrafish and mouse transgenic enhancer assays. Our results provide a novel dataset of neuronal enhancers that could control the spatiotemporal expression of IS-associated genes. These sequences pose as great candidates for mutation analyses in IS patients, a screen which we are currently carrying out. Combined, this work will shed light on neuronal gene regulation in general and identify novel genomic regions that could be involved in epilepsy pathogenesis and brain development.

1996T

Differential Behavior of Splice Isoforms of the Asthma Susceptibility Gene *DENND1B*. *M.E. March, P.M.A. Sleiman, C. Hou, J. Bradfield, C.E. Kim, E.C. Frackleton, J.T. Glessner, H. Hakonarson.* Children's Hospital of Philadelphia, Philadelphia, PA.

Asthma is a chronic inflammatory condition of the lungs, characterized by excessive responsiveness of the lungs to stimuli, in the forms of infections, allergens, and environmental irritants. Currently, 22.9 million Americans suffer from asthma, and the prevalence has increased dramatically since 1980. Asthma is the leading chronic illness in U.S. children, with 6.8 million affected in 2006. Twin studies have shown that there is a genetic element to asthma susceptibility (with heritability of the condition estimated at between 0.36 and 0.77), but only a fraction of the heritability is explained with previously identified loci. In a genome wide association screen (GWAS), we identified *DENND1B* as a novel asthma susceptibility locus. At least two splicing isoforms of *DENND1B* exist, and we have attempted to examine the relative activities of these isoforms. Overexpression of the short or long isoform of DENND1B in 293T cells had opposite effects on the response of those cells to TNFα stimulation, with the long isoform enhancing the response and the short isoform diminishing it. Confocal microscopy with GFP-tagged versions of the DENND1B isoforms have revealed that the long isoform displays some localization to the plasma membrane of expressing cells, while the short isoform appears entirely cytoplasmic. Biochemical cell fractionation experiments show similar localization patterns. The DENND1B proteins possess a domain (the conserved DENN domain) that functions as a guanonucleotide exchange factor for the small GTPase Rab35. The DENN domain is identical in the long and short isoforms, however we find that the short isoform more readily interacts with Rab35, particularly when the Rab35 is expressed in a mammalian cell system. We are currently determining which of the splice isoforms is expressed in different cell types and to what relative levels, and if splicing is altered in individuals with either asthma-susceptible or -resistant genotypes.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

1997F

Targeted resequencing identifies secretion-defective variants of decoy receptor 3 in pediatric-onset inflammatory bowel disease. *C. J. Cardinale[1], S. Panossian[1], F. Wang[1], E. C. Frackelton[1], C. E. Kim[1], F. D. Mentch[1], R. M. Chiavacci[1], K. E. Kachelries[2,3], R. Pandey[1], S. F. A. Grant[1,3], R. N. Baldassano[2,3], H. Hakonarson[1,3].* 1) Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, PA; 2) Division of Gastroenterology, Hepatology, and Nutrition, Children's Hospital of Philadelphia, Philadelphia, PA; 3) Department of Pediatrics, Perelman School of Medicine at the University of Pennsylvania.

Tumor necrosis factor receptor superfamily *(TNFRSF)6B* encodes a secreted protein, decoy receptor (DcR)3, that binds to and neutralizes pro-inflammatory cytokines of the tumor necrosis factor superfamily, namely TL1A, LIGHT, and Fas ligand. Genome-wide association studies have implicated common variation in the *TNFRSF6B* locus in inflammatory bowel disease, particularly of the pediatric Crohn's type. We sought to enhance the evidence of DcR3′s role in pediatric IBD by identifying rare missense mutations with functional significance through Sanger sequencing in 592 cases. The variants uncovered in the cases went on to be assessed functionally for their ability to be secreted from 293T cells and to bind to their ligands. Our pediatric IBD cohort harbored 12 missense variants at the *TNFRSF6B* locus, with six affecting secretion of DcR3 from cultured cells. The load of missense variation in the TNFRSF6B gene was significantly enriched in our Caucasian Crohn's subclass (2%) than in publically-available control (0.34%) exomes (p less than 4.2×10^{-4}; odds ratio = 5.9;). Furthermore, the 3 missense variants that are seen in both cases and controls are all secretors, pointing to the fact that the majority of the variants exclusive to the cases are non-secretors, revealing a non-random pattern of mutation. As personalized medicine targeting IBD emerges, sequencing or genotyping of variants at the TNFRSF6B locus should be included. Development of therapeutics targeting this pathway is a priority for future research.

1998W

A population-based immunologic phenotype study for the identification of high risk individuals to childhood infectious diseases in Vietnam. *R. Miyahara[1], L.M. Yoshida[1], H. Nakamura[1], K. Takahashi[1], X.M. Trinh[2], D.A. Dang[3], N.H. Tran[4], K. Ariyoshi[1], M. Yasunami[1].* 1) Institute of Tropical Medicine, Nagasaki University, Nagasaki, Japan; 2) Pasteur Institute Nha Trang, Nha Trang, Vietnam; 3) National Institute of Hygiene and Epidemiology (NIHE), Hanoi, Vietnam; 4) Pasteur Institute Ho Chi Minh City, Ho Chi Minh City, Vietnam.

The susceptibility to certain infectious disease is attributable to augmented responsiveness to pathogen-derived stimulus in part. Accordingly, properties of host defense mechanism evaluated by *ex vivo* cellular response to stimuli mimicking microbial pathogens would be eligible to infer the individual predisposition to infectious diseases. A population-based study of such immune-related phenotype analysis was conducted. A cohort of 1,999 infants who were born at Khanh Hoa Provincial Hospital, a principal medical facility in Nha Trang, central Vietnam, was set from May 2009 to May 2010 to observe the incidence of severe infectious diseases. In the present study, the registered children were invited to medical check in the month they became two year old. Venous blood samples were collected upon their guardians' informed consent. A simple procedure of small-scale whole blood culture was established to obtain a reliable result under the situation where the amount of samples and manpower are limited; a fresh whole blood was mixed with cell culture medium at 1:9 ratio and incubated at 37 degree in the presence of LPS (a ligand for TLR4), Pam3CSK4 (a ligand for TLR2-TLR1 heterodimer) or other chemical stimulants of immune-competent cells. Twenty hours later, the cells were collected for RNA preparation. The levels of induced mRNA expression of cytokine genes such as IL8 and those of internal standards such as GAPDH were quantified by real time RT-PCR assay. Among 124 children examined in May and June 2011, 118 gave complete quantification data set. 11, 17, 2 and 11 of 118 children had a history of admission to the hospital because of respiratory infection, gastrointestinal infection, dengue virus infection and undefined fever, respectively, and the rest 77 were free of severe infectious disease in their first two years of life. Induction of IL8 mRNA expression by LPS (p=0.0075) and Pam3CSK4 (p=0.0095) was higher in children with a history of respiratory infection in comparison to those without history of admission by infectious disease, mainly because of lower basal expression level of IL8 in children with a history of respiratory infection (p=0.0069). Receiver-operator curve analysis indicated that the basal lower IL8 expression was a good predictor for admission by respiratory infection in this cohort (AUC=0.71). The results suggested that impaired regulation of inflammatory cytokine gene predisposes children to severe infectious diseases.

1999T

A birth-cohort study for the identification of genetic risk for childhood infectious diseases in Vietnam. *M. Yasunami[1], L.M. Yoshida[1], R. Miyahara[1], H. Nakamura[1], K. Takahashi[1], H. Moriuchi[2], D.A. Dang[3], N.H. Tran[4], K. Ariyoshi[1].* 1) Clinical Medicine, Inst Trop Med, Nagasaki Univ, Nagasaki, Japan; 2) Medical Virology, Grad Sch of Biomedical Sciences, Nagasaki Univ, Nagasaki, Japan; 3) National Institute of Hygiene and Epidemiology (NIHE), Hanoi, Vietnam; 4) Pasteur Institute Ho Chi Minh City, Ho Chi Minh City, Vietnam.

It has been generally accepted that individual susceptibility to certain infectious diseases are genetically determined in part. A large amount of information about association between common genetic variations and incidence and/or severity of infectious disease has been accumulated mainly as results of retrospective case-control studies. We conducted a prospective study to evaluate the predictive power of common genetic variation in candidate genes for infections/infectious diseases by employing all the newborn babies in a single core medical facility in urban-rural mixed area in Vietnam. A total of 1999 full-term newborns without apparent abnormality at birth, who were from 16 communes of Nha Trang city and adjacent suburban area which cover approximately one third of total population of the province, were enrolled from the beginning of May 2009 to the end of May 2010. Umbilical cord blood was collected to obtain plasma samples for testing congenital infections and nucleated cells for genomic DNA. All the children were provided with standard national immunization program composed with DPT+HBV and BCG after birth. Occurrence of severe illness was captured at the core hospital. Upon the admission to the hospital, the presence of certain viral and bacterial genome in the pharyngeal swab was examined for the patients with respiratory infection. All the participants were invited to medical check in their two year-old birth month and venous blood was collected after obtaining their guardians' informed consent. HLA and other 750 selected candidate genetic variations were genotyped by using genomic DNA isolated from cord blood. Venous blood collected at two year-medical check was used for blood cell count, plasma separation for antibody tests for HBV and other microbial pathogens, and blood cell phenotype analysis. Frequencies of HLA-A, -B and -DRB1 alleles of 1985 children of the cohort whose DNA was available were similar to those of healthy individuals of northern and southern Vietnam reported by Hoa (2007) and Nguyen (2008), respectively. During two years of observation, 33% of children had a history of severe disease which required hospitalization once or more, and incidence of severe respiratory infection was 86.9/1000 person;lbyears The effect of HLA and other polymorphisms on the incidence of severe respiratory infection and other infectious diseases was evaluated.

2000F

Variant genotypes of PKR1 and PKR2 in patients of recurrent pregnancy loss. *M. Su, L. Wu, P. Kuo.* National Cheng-Kung University Hospital, Tainan City, Taiwan.

Recurrent pregnancy loss (RPL) is a multi-factorial disorder and up to 50% of cases remain undetermined causes after detailed clinical examination. Angiogenesis plays a critical role in early gestation and endocrine gland-derived VEGF (EG-VEGF) is a tissue-specific angiogenesis-associated gene. The role of EG-VEGF and its two receptor genes (PKR1, PKR2) in human early pregnancy was believed to have a direct effect on both endothelial and trophoblastic cells and are likely to play important roles in placentation. We previously found gene polymorphisms of PKR1, PKR2 were significantly associated with human RPL using tag SNP analysis. We now direct sequenced these genes in 100 RPL patients and 100 normal controls, trying to fine map the variation sites that interfere with early pregnancy and further functional validate in vitro studies. We found allele and genotype frequencies of PKR1(I379V) and PKR2(V331M) were significantly higher in the normal controls and may play protective roles in RPL ($p<0.05$). Both variants induced nonsynonymous change of amino acids and located in the intracellular C-terminal domains of G protein-coupled receptors. We further demonstrated PKR1(I379V)- and PKR2(V331M)-overexpressed cell had altered intracellular calcium influx and significantly higher ability of cell invasiveness in both HEK293 and JAR (trophoblast) cell lines. We therefore concluded that PKR1(I379V) and PKR2(V331M) may play protective roles in preventing RPL by altering intracellular calcium signaling and enhancing cell invasion ability.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2001W

Post GWAS analysis of the *BCL11A* intronic region to define its role in regulating HbF levels. *F. Anedda[1], S. Sanna[1], I. Asunis[1], G. Usala[1], F. Danjou[2], L. Perseu[1], A. Cabriolu[1], C.A. Caria[1], L. Porcu[1], M.G. Marini[1], M.F. Marongiu[1], C. Sidore[3, 4], R. Berutti[3, 5], M. Pala[3], A. Angius[1, 5], F. Busonero[1], A. Maschio[1, 4], S. Satta[2], F. Demartis[2], L. Maccioni[2], A. Meloni[1], R. Nagaraja[6], G. Abecasis[4], D. Schlessinger[6], M.S. Ristaldi[1], R. Galanello[2], P. Moi[2], F. Cucca[1, 3], S. Sanna[1], M. Uda[1].* 1) Istituto di Ricerca Genetica e Biomedica (IRGB), Consiglio Nazionale delle Ricerche (CNR), Cittadella Universitaria di Monserrato, Monserrato, Cagliari, Italy; 2) Clinica Pediatrica 2a, Dipartimento di Scienze Biomediche e Biotecnologie - Università di Cagliari, Ospedale Regionale Microcitemie ASL8, Cagliari, Italy; 3) Dipartimento di Scienze Biomediche, Università di Sassari, Sassari, Italy; 4) Center for Statistical Genetics, Department of Biostatistics, University of Michigan, Ann Arbor, Michigan, USA; 5) Center for Advanced Studies, Research and Development in Sardinia (CRS4), AGCT Program, Parco Scientifico e tecnologico della Sardegna, Pula, Italy; 6) Laboratory of Genetics, National Institute on Aging, Baltimore, Maryland, USA.

BCL11A transcription factor is one of the main genetic modifiers of both β-thalassemia and sickle cell anemia phenotypes. It functions as a developmental stage-specific repressor of fetal hemoglobin (HbF) expression, controlling globin switching, both in human and mouse. Several genome-wide association studies (GWAS) have identified a region of about 15 Kb in intron 2 of the *BCL11A* gene containing all variants associated with HbF levels as well as clinical severity of both diseases. Recently we performed a GWAS on 2343 individuals from the SardiNIA project, integrating, by imputation procedure, DNA microarray data and low pass whole-genome sequencing of 826 unrelated individuals. Our analysis revealed that two independent variants in the region, representing the strongest haplotypic signal, fully account for the association observed at this locus. Given that intron 2 shows chromatin signatures of functional genomic elements with a potential regulatory activity, we carried out functional analysis to assess the specific impact of associated variants in gene action. In particular we evaluated whether these SNPs might be themselves the causative variants at this locus or if their surrounding sequence has regulatory features. Furthermore, we conducted expression studies on erythroid cell lines (K562 and HEL) as well as BFU-E cells, isolated at different days of differentiation from 10 patients with β^0-thalassemia, a Mendelian disorder whose phenotype is ameliorated by these two HbF associated variants. To estimate differential expression of *BCL11A* transcripts, we selected a group of 5 patients with a milder phenotype carrying the high HbF levels haplotype, and the remaining 5 with a severe phenotype carrying the low HbF levels haplotype. Preliminary results revealed significant differential expression levels of the *BCL11A* main isoform between the two groups ($p = 0.006$), in BFU-E at day 9 of differentiation. Overall, our results suggest that genetic variants at intron 2 of *BCL11A* play a crucial role in modulating its expression and in turn HbF levels with consequent amelioration of β-thalassemia severity.

2002T

Additional Patients and an Association Study support a Role of SOX9 in CD-ACD-PRS Phenotypic Continuum and in CPO. *M. Quentric[1,2], L. Desmyter[1], M. Ghassibe[1], W. Courtens[3], S. Kivirikko[4], H. Antoine-Poirel[3], G. Ameye[3], B. Bayet[5], G. Francois[5], R. Vanwijck[5], O. Boute[6], P. Pellerin[7], M. Rubini[8], M. Vikkula[1,3].* 1) Laboratory of Human Molecular Genetics, De Duve Institute, Université catholique de Louvain, Brussels, Belgium; 2) Inserm UMRS872 Team 5, Centre de Recherches des Cordeliers, Université Paris Descartes, Paris, France; 3) Center for Human Genetics, Cliniques universitaires St Luc, Université catholique de Louvain, Brussels, Belgium; 4) Department of Clinical Genetics, Helsinki University Central Hospital, Helsinki, Finland; 5) Centre Labiopalatin, Division of Plastic Surgery, Cliniques universitaires St Luc, Brussels, Belgium; 6) Centre de génétique, Centre hospitalier universitaire de Lille, Lille, France; 7) Service de chirurgie plastique et reconstructive, Centre hospitalier universitaire de Lille, Lille, France; 8) Sez. Genetica Medica, Dip. Medicina Sperimentale e Diagnostica, Università di Ferrara, Ferrara, Italy.

Sox9 has an essential role in chondrogenesis. Nonsense mutations and deletions of Sox9 suggested haploinsufficiency to underly campomelic dysplasia (CD), a rare autosomal dominant disease characterized by campomelia, skeletal defects and Pierre Robin sequence (PRS). Translocations in the 350kb region upstream or downstream of SOX9 were reported in less severe CD patients. Some translocations and deletions further upstream have been identified in patients with acampomelic campomelic dysplasia (ACD) and PRS, suggesting that these three syndromes form a continuum of phenotypes. We report the identification of a translocation 600kb upstream of SOX9 in a patient with classical features of ACD. In parallel, we identified a deletion in a PRS family that overlaps a deletion reported in a patient with the same phenotype. These data and the literature lead us to test whether or not there is a genetic association between the SOX9 locus and cleft palate (CP) and/or PRS. We used three SNPs in a cohort of case-parent trios analyzed using TDT. While two SNPs, tagging a conserved region upstream of SOX9, were not associated, significant over-transmission of an intragenic SNP was detected in the combined CP+PRS cohort ($p=0.026$) and PRS cohort ($p=0.042$), with a relative risk of 0.54 under a dominant model. The association has been replicated in an independent Italian CP-cohort. Our data are in agreement with the hypothesis that removal or disruption of cis-regulatory elements upstream of SOX9 can lead to phenotypes of gradual severity. We show for the first time an association of SOX9 with CP and PRS.

2003F

The impact of amino acid polymorphisms in four genes of the tumor necrosis factor (TNF) and interferon-gamma (IFNG) pathways on BCG-triggered production of TNF. *T. Yu[1], L. de Léséleuc[2], A. Cobat[1], L. Simkin[1], G.F. Black[4], K. Stanley[4], P. van Helden[4], L. Abel[3], A. Alcaïs[3], E.G. Hoal[4], E. Schurr[1].* 1) McGill Centre for the Study of Host Resistance & Dept. of Human Genetics, McGill University, Montreal, Quebec, Canada; 2) Institute for Biological Sciences, National Research Council, Ottawa, Canada; 3) Université de Paris Descartes, Necker Medical School, Paris, France; 4) Faculty of Health Sciences, Stellenbosch University, Tygerberg, South Africa.

Tuberculosis (TB) is an infectious disease caused by *Mycobacterium tuberculosis* for which host genetic factors play an important role. Numerous cytokines participate in the immune pathway of TB, including tumor necrosis factor (TNF) and interferon gamma (IFNG). Prior genetic analyses detected four non-synonymous amino acid changes in the following four genes located in IFNG and/or TNF pathway that were significantly associated with BCG (*M. bovis* bacille Calmette-Guérin)-triggered TNF production in South African families: M594I (*STAT2*), N248S (*TLR1*), N75Y (*DNAJA3*) and H712Q (*NFKB1*). Our main research objective is to provide functional validation of these non-synonymous polymorphisms that are associated with TNF production. For this, we first assessed whether the corresponding human genes were modulators of TNF secretion in various *in-vitro* cell assays. The possible allelic influence on biological activity of the corresponding proteins was determined next. Previous study showed that N248S is in complete linkage disequilibrium (LD) with I602S in *TLR1*, which regulates NF-κB signaling and affects the innate immune response to lipopeptides (Hawn *et al*., 2007). This strongly suggests that the association of N248S with TNF production is mediated by I602S. While *DNAJA3* may affect TNF release, we did not observe any allelic impact of N75Y. Preliminary data indicates that H712Q (*NFKB1*) affects p50 conversion, supporting its potential function in NF-κB signaling. The possible allelic impact of H712Q (*NFKB1*) and M594I (*STAT2*) on TNF release is currently under investigation. In conclusion, our study helps to understand the underlying relationships among specific amino acid changes, their biological functions and TNF secretion.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2004W

Functional variants of *NFKBIE* and *RTKN2* genes are associated with rheumatoid arthritis susceptibility in Japanese. *K. Myouzen[1], Y. Kochi[1,2], Y. Okada[1,2,3], C. Terao[4,5], A. Suzuki[1], K. Ikari[6], T. Tsunoda[7], A. Takahashi[3], M. Kubo[8], A. Taniguchi[6], F. Matsuda[4,9,10], K. Ohmura[5], S. Momohara[6], T. Mimori[5], H. Yamanaka[6], N. Kamatani[11], R. Yamada[12], Y. Nakamura[13], K. Yamamoto[1,2].* 1) Laboratory for Autoimmune Disease, Center for Genomic Medicine (CGM),RIKEN, Tokyo, Japan; 2) Department of Allergy and Rheumatology, Graduate School of Medicine, the University of Tokyo, Tokyo, Japan; 3) Laboratory for Statistical Analysis, CGM, RIKEN, Yokohama, Japan; 4) Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan; 5) Department of Rheumatology and Clinical immunology, Graduate School of Medicine, Kyoto University, Kyoto, Japan; 6) Institute of Rheumatology, Tokyo Women's Medical University, Tokyo, Japan; 7) Laboratory for Medical Informatics, CGM, RIKEN, Yokohama, Japan; 8) Laboratory for Genotyping Development, CGM, RIKEN, Yokohama, Japan; 9) CREST program, Japan Science and Technology Agency, Kawaguchi, Saitama, Japan; 10) Institut National de la Sante et de la Recherche Medicale (INSERM) Unite U852, Kyoto University Graduate School of Medicine, Kyoto, Japan; 11) Laboratory for International Alliance, CGM, RIKEN, Yokohama, Japan; 12) Unit of statistical Genetics, Center for Genomic Medicine Graduate School of Medicine Kyoto University, Kyoto, Japan; 13) Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, University of Tokyo, Tokyo, Japan.

[Purpose] The aims of study are 1) to confirm the association signals of our previous genome-wide association study (GWAS) for rheumatoid arthritis (RA) in Japanese population and 2) to search for disease causing variants in the candidate loci. [Methods] We selected 31 loci that showed moderate associations in the GWAS (2,303 cases and 3,380 controls, $5\times10^{-8}<P<5\times10^{-5}$) and performed replication studies of association using additional Japanese individuals (5,604 cases and 31,982 controls). To identify disease-causing nsSNPs, we sequenced the coding region of candidate genes in RA patients (n=48) and investigated their effects on protein function by NF-κB reporter assays. We next performed allele-specific transcript quantification (ASTQ) analysis by using TaqMan assays to investigate allelic imbalance of expression. In addition, we evaluated regulatory potential of SNPs in the region by *in silico* analyses using genomic and epigenomic databases as well as by *in vitro* analyses using gel-shift assays and reporter assays. [Results] We identified variants in two gene loci that were associated with RA susceptibility (rs2233434 in *NFKBIE*, OR=1.20, $P=1.3\times10^{-15}$; rs3125734 in *RTKN2*, OR= 1.20, $P=4.6\times10^{-9}$). We identified disease-associated nsSNPs (rs2233434 and rs2233433 in *NFKBIE*, rs3125734 and rs61850830 in *RTKN2*), which comprised a single common risk haplotype in each locus. The risk haplotype of *NFKBIE* showed higher NF-κB activity than non-risk haplotype, but no difference was detected in *RTKN2*. In the ASTQ analysis, the risk alleles showed 1.2-fold lower amounts of transcripts in *NFKBIE* and 1.5-fold higher amounts of transcripts in *RTKN2*, suggesting that both gene loci were eQTL. 14 SNPs in *NFKBIE* and 10 SNPs in *RTKN2* were selected as having regulatory potential by the *in silico* analysis, and 9 SNPs showed allelic differences in the gel-shift assays. In the reporter assays, the risk allele of rs2233424 in *NFKBIE* showed enhanced repressor activity, while the risk alleles of rs12248974 and rs61852964 in *RTKN2* displayed more enhancing activity than the non-risk alleles. [Conclusions] SNPs in *NFKBIE* and *RTKN2* genes were significantly associated with RA susceptibility. Through functional analysis, we identified candidate causal SNPs in *NFKBIE* (rs2233434, rs2233433 and rs2233424) and *RTKN2* (rs12248974 and rs61852964). The risk alleles of these SNPs were implicated in the enhancement of NF-κB activity, indicating the importance of this pathway in the etiology.

2005T

Resequencing *ITGAM* reveals two functionally deleterious rare variants in Systemic Lupus Erythematosus cases. *A.L. Roberts[1], E.R.A. Thomas[2], S. Bhosle[2], L. Game[2], O. Obraztsova[1], T.J. Aitman[2], T.J. Vyse[1], B. Rhodes[1].* 1) Dept. of Genetics & Molecular Medicine, King's College London, London, United Kingdom; 2) MRC Clinical Sciences Centre, Imperial College London, London, United Kingdom.

Systemic lupus erythematosus (SLE) is an autoimmune disease with a strong genetic basis. Despite notable gene discoveries in genome-wide association studies there remains a large amount of “missing heritability”, that may be accounted for by rare variants (MAF<1%) which these techniques are underpowered to detect. A genetically associated non-synonymous variant (R77H) in *ITGAM* (CD11b) impairs complement receptor 3 (CR3; CD11b/CD18) function, including the phagocytosis of iC3b-opsonised sheep erythrocytes ($sRBC_{iC3b}$). We screened *ITGAM* in an SLE cohort to identify functionally deleterious rare variants. 24Kb of the *ITGAM* gene, including all 30 exons, was PCR amplified in 75 SLE patients and 95 controls. PCR products were pooled and tagged with unique DNA barcodes using the Parallel Tagged Sequencing protocol and sequenced on the Roche 454 Titanium high-throughput sequencing platform. An average coverage of 105x (896,206 reads) for cases and 35x (542,867 reads) for controls was achieved. Identified variants were confirmed by Sanger sequencing. Two missense mutations were identified - one rare (F941V) and one private (G1145V) - in one patient sample each. Neither variant is found in the 1000 genomes database; F941V is described in 1/6707 chromosomes of European ancestry in the Exome Variant Server database but G1145V is not found. The variants were introduced into pcDNA3.1-*ITGAM* vectors using site-directed mutagenesis. Transfected COS-7 cells were challenged with $sRBC_{iC3b}$, using an identical technique to our published approach for R77H variant studies. We measured the association index (AI = mean number of engaged $sRBC_{iC3b}$/100 COS cells) and phagocytic index (PI = mean number of phagocytosed $sRBC_{iC3b}$/100 COS cells). Comparisons were made by paired t-test. We observed no difference in transfection efficiency (measured by surface expression of CD11b). Neither variant impaired iC3b binding (AI F941V P=0.61; AI G1145V P=0.74). However, both variants significantly impaired phagocytosis (PI F941V P=0.007; PI G1145V P=0.031). Through resequencing a candidate gene, we have identified two rare variants in SLE patients. We observed an under-functioning affect, of similar magnitude to the SLE-associated R77H, on phagocytosis of RBC_{iC3b}. Our work adds confirmatory evidence to the functional importance of CR3 in SLE pathogenesis, and provides a further example of a gene which harbours both rare and common variation predisposing to the same disease.

2006F

***DcR3* Mutations in Systemic Lupus Erythematosus Patients Lead to Enhanced Lymphocyte Proliferation.** *K. Suphapeetiporn[1,2], C. Chokdeemeeboon[2,3], P. Ammarinthnukrowh[2,3], S. Tongkobpetch[2,3], C. Srichomtong[1,2], T. Deekajorndech[1], P. Rianthavorn[1], P. Kingwattanakul[1], Y. Avihingsanon[4], V. Hoven[5], S. Edwards[6], N. Hirankarn[7], V. Shotelersuk[1,2].* 1) Department of Pediatrics, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand; 2) Center of Excellence for Medical Genetics, Department of Pediatrics, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand; 3) Interdepartment of Biomedical Sciences, Faculty of Graduate School, Chulalongkorn University, Bangkok, Thailand; 4) Department of Medicine, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand; 5) Department of Chemistry, Faculty of Science, Chulalongkorn University, Bangkok, Thailand; 6) Institute of Integrative Biology, University of Liverpool, Liverpool, UK; 7) Department of Microbiology, Faculty of Science, Chulalongkorn University, Bangkok, Thailand.

Systemic lupus erythematosus (SLE) is a multifactorial disorder. Although several of its underlying genes have been identified, many more remain to be discovered. Previous studies suggested that death decoy receptor 3 (DcR3) might play a role in pathogenesis of adult SLE. Here, we found that the serum DcR3 levels of juvenile SLE patients in active state (440.2 ± 169.3 pg/ml) were significantly higher than those in inactive state (181.7± 42.72 pg/ml; p = 0.003) and unaffected controls (65.52 ± 20.66 pg/ml; p = 0.001). This indicated an association between an increased serum DcR3 level and juvenile SLE. We then performed PCR-sequencing of the entire coding regions of *DcR3* in 100 juvenile and 100 adult unrelated SLE patients and found two heterozygous missense mutations (p.T56I and p.H122Y) in two patients. Using surface plasmon resonance assay, we found that the p.T56I mutant had a 2.4 fold greater, whereas the p.H122Y had 0.7 fold lower affinity to FasL than the wild-type DcR3. Although both mutations did not change the activation-induced apoptosis, they statistically significantly increased lymphocyte proliferation more than that of the wild-type. In summary,serum levels of DcR3 are elevated in juvenile SLE patients with active disease. These data demonstrate, for the first time, that polymorphisms in *DcR3* might play an etiologic role in human SLE pathogenesis via regulation of T cell function.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2007W

A *CD14* promoter polymorphism is implicated in tuberculosis susceptibility in a South African population. *M. Möller[1], C. Wagman[1], M. Daya[1], C. Kok[1], L. van der Merwe[2, 3], P.D. van Helden[1], E.G. Hoal[1].* 1) Biomedical Sciences, Stellenbosch University, Cape Town, Western Cape, South Africa; 2) Biostatistics Unit, Medical Research Council, Tygerberg, South Africa; 3) Department of Statistics, University of Western Cape, South Africa.

Tuberculosis is a multi-factorial disease dependent not only on the infectious agent *Mycobacterium tuberculosis*, but also affected by the environment and host genetic factors. The rs2569190 (C-159T) promoter polymorphism in the LPS-binding cell surface molecule CD14 gene (*CD14*) has been reported to be a risk factor for pulmonary tuberculosis in Mexico and Korea and is known to regulate promoter activity of *CD14*. We investigated the role of this polymorphism in tuberculosis susceptibility in a South African population. The variant was genotyped in 398 tuberculosis patients and 394 control individuals in a population-based case-control association study. Individuals carrying the T allele were protected against developing tuberculosis (p = 0.019, OR = 0.77, 95% CI 0.62 - 0.96). We investigated the functional significance of this polymorphism in gene expression experiments using qPCR and reporter gene assays and showed that the TT genotype of rs2569190 results in higher expression of CD14. Our results demonstrate that an association exists between a *CD14* polymorphism and tuberculosis and also affirmed its effect on gene expression.

2008T

Effects of IL9 and IL9 receptor(IL9R)gene polymorphisms on Allergic Rhinitis in Iranian females. *F. Fatahi, H. Khazraee, K. Ghatreh, M. Hashemzadeh.* Biochemistry and Genetic Center, Shahrekord University Of Medical Sciences, Shahrekord, Chaharmahal Va Bakhtiari, Iran.

Background and aims: The development of allergic rhinitis entails a complex interaction between genetic predisposition and environmental exposure to different factors that allergens are the most important. Responding molecules are; chemokines and their receptors, interleukins and their receptors, eosinophil peroxidase and leukotrienes, among others.The interleukin-9 cytokine and its reseptor are emerging as critical players in allergic diseases. This study investigated the association between single-nucleotide polymorphisms(SNP) in IL9 gene (rs2069885,)and IL9R (rs731476) and Rhinitis-related traits among the females in Chaharmahal va Bakhtiari province. Methods: DNA was extracted using standard phenol-chloroform method. The screening of mentioned polymorphisms were performed using PCR-RFLP procedure. A case- control association study was performed(rhinitis group;n=195 and control group;n=190). Chi- square test was performed to compare proportions of subjects with different clinical features among subjects with different genotypes.(All statisticall analyses were performed using SPSS). Result: There was significant association between rs731476 ;IL9R and allergic rhinitis (p=0.004) but no association between rs2069885;IL9 and cited disease between females in cases and controls goups in Chaharmahal va Bakhtiari province was found (p=0.991). Conclusions: Our data indicated that the IL9R may play an important role in the inflammatory response and promoting allergic rhinitis in females and rs 2069885 ;1L9 may not have a role in femals with rhinitis in Chaharmahal va Bakhtiari province. Keywords: Allergic Rhinitis ;polymorphism;Interleukin-9 and IL9R;Association study.

2009F

TRIB1 is involved in the susceptibility of non-alcoholic fatty liver disease. S. *Iwamoto, Y. Ishizuka, Y. Kitamura, S. Makishima, S. Boonvisut, K. Nakayama.* Human Genetics, Center for Molecular Medicine, Jichi Medical University, Shimotsuke, Tochigi, Japan.

Tribbles homolog 1 (TRIB1) is the gene associated robustly with plasma triglyceride (TG) and LDL cholesterol level. Moreover, association with myocardial infarction is also repeatedly shown. The hepatic over expression reduced plasma TG and LDL in mice. Conversely, those in TRB1 knockout mice were increased. The Trib1 over expression impaired VLDL production in liver through down regulation of lipogenesis. This study aims to assess the involvement of TRIB1 gene in the development of hepatic steatosis, which is a recently nominated risk factor of cardiovascular diseases. MATERIALS and METHODS: Japanese general population admitted to Hospital of Jichi Medical University for health checkup, 3013 (1623 males and 1390 females), was enrolled in this study with written informed consent as the first genome panel. Hepatic steatosis was diagnosed by the finding of diffuse ultrasonographic bright liver and hepatorenal contrast. The second panel was extracted from large scale Japanese genome bank (J. Med. Genet. 2009;46;370-374) on the condition of the attachment of ultrasonographic data. Genotyping of TRIB1 SNPs was performed by using TaqMan Genotyping Assay Systems and an ABI PRISM 7900 HT Sequence Detection System. The effects of copy numbers of minor alleles of each SNP on hepatic steatosis were assessed by logistic regression models. Age, sex, body mass index (BMI), status of alcohol consumption, and medication for diabetes mellitus were included in multiple linear regression models as covariates. The functional involvement of Trib1 in hepatic lipid storage was assessed by TRIB1 cDNA or shRNA transduction into mice liver via adenovirus vector system. Hepatic TG strage and plasma TG were measured at day4 and day8 after viral administration. RESULTS and DISCUSSION: The Japanese major allele of re17321515, G, correlated with higher plasma TG level (P= 4.8E-4 (panel 1), P=4.6E-6 (panel 2)) and with the ultrasonographic diagnosis of hepatic steatosis (P=0.001 (panel 1), P=0.019 (panel 2), P=3.87E-5 (combined)). The mice enforced the over expression of Trib1 in liver reduced the plasma TG level and the hepatic TG storage. The knockdown vector increased the mice plasma and hepatic TG. These results indicated that the expression level of TRIB1 is involved in the development of non-alcoholic fatty liver disease.

2010W

Alleles of a Rapidly-Evolving ETV6 Binding Site in *DCDC2* Confer Risk of Reading and Language Impairment. *N.R. Powers[1], J.D. Eicher[1], F. Butter[3], L.L. Miller[4], S.M. Ring[4], M. Mann[3], J.R. Gruen[1,2].* 1) Genetics Dept, Yale University, New Haven, CT; 2) Pediatrics Dept, Yale University, New Haven, CT; 3) Proteomics and Signal Transduction Dept, Max Planck Institute of Biochemistry, Martinsried, Germany; 4) School of Social and Community Medicine, University of Bristol, Bristol, UK.

Learning disabilities (LDs) are disorders involving a specific deficit in a particular mode of learning, such as reading, writing, math, or speech. The most common LDs involve language=mthe National Institute of Child Health and Development (NICHD) estimates that 15-20% of Americans have a language-based LD, and that the vast majority of these have reading disability (RD), also known as dyslexia. RD is a specific deficiency in ability to process written language. Another common LD, language impairment or LI, is a specific deficiency in ability to process spoken language. Comorbidity between RD and LI is common. Both disorders are substantially genetic, and their inheritance is typical of a complex trait. We here report two risk haplotypes in the RD-associated gene *DCDC2* that associate strongly with an RD and an LI endophenotype ($p = 6.068 \times 10^{-5}$ and $p = 3.98 \times 10^{-4}$, respectively). The haplotype block containing these haplotypes is almost directly adjacent to BV677278, a rapidly-evolving, highly polymorphic, compound short tandem repeat (STR) that we previously showed to associate with RD. Each of the two risk haplotypes is strongly linked to an allele of BV677278; 92% of RD risk haplotype carriers also carry BV677278 allele 5, and 78% of LI risk haplotype carriers also carry BV677278 allele 6. We also showed previously that BV677278 binds a nuclear protein with high specificity, and that it can modulate reporter gene expression from the *DCDC2* promoter in an allele-specific manner. Using stable isotope labeling of amino acids in cell culture (SILAC) based quantitative mass spectrometry, we have now identified the protein as the transcription factor ETV6. We confirmed the ETV6-BV677278 interaction *in vivo* via chromatin immunoprecipitation (ChIP). These data suggest that BV677278 is a regulatory element that influences risk of reading and language impairment. Intriguingly, the two *DCDC2* risk haplotypes interact synergistically with a previously described RD risk haplotype encompassing the 5′ and upstream region of *KIAA0319*, another RD-associated gene within 200kB of *DCDC2* on the chromosome, to adversely affect mean performance on several reading, language, and cognitive measures. These data lend support to the hypothesis that regulatory elements may be largely responsible for the variation seen in complex traits, and to the `phantom heritability' hypothesis, which explains missing heritability as resulting from non-additive gene-gene interactions.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2011T

Polymorphisms in phosphatidylethanolamine N-methyltransferase (PEMT) gene associates with obesity and gene expression. *S.K. Das, K.A Langberg, A.K. Mondal, N.K. Sharma.* Internal Medicine-Endocrinology, Wake Forest School of Medicine, Winston-Salem, NC.

Recent studies in animal models have shown a mechanistic link between the alteration of endoplasmic reticulum (ER) phospholipid composition and activation of ER stress in the development of insulin resistance in obesity. Expression of genes involved in phosphatidylcholine (PC) synthesis (*PCYT1A* and *PEMT*) was upregulated, while phosphatidylethanolamine (PE) synthesizing gene (*PCYT2*) was downregulated, leading to high PC/PE ratio in the livers of obese mice. Expression of *PEMT* is increased in adipose of high fat fed mice and is involved in lipid droplet biogenesis during adipocyte differentiation. To understand the role of genes involved in phospholipid biosynthesis in human obesity we have analyzed their expression in adipose and muscle from 170 individuals who spanned a broad range of insulin sensitivity (S_I) and body mass index (BMI). We have investigated the association of SNPs in these genes with obesity and glycaemic traits in large GWA studies. We found positive correlation of % fat mass with expression of *PEMT* (r= 0.46, p= 9.57E-09), *PCYT1A* (r= 0.36, p= 1.09E-05) and *PTDSS2* (r= 0.29, p= 4.72E-04), and negative correlation of *PCYT2* (r= -0.30, p= 4.15E-04) in adipose but not in muscle. Expression of *PEMT* (r = -0.33, p= 1.37E-04) and *PCYT2* (r= 0.34, p= 7.15E-05) in adipose was also correlated with S_I, but the correlation was reduced after adjustment for BMI. SNPs in these genes showed no significant association with glycaemic traits in Caucasian subjects from the MAGIC consortium. Interestingly nine SNPs in *PEMT* gene showed association (p<0.001) with Waist-to-hip ratio (WHR) in a meta-analysis of 77158 Caucasian subjects from the GIANT consortium. The strongest association of WHR was observed with intronic SNP rs4646404 (p= 3.1E-06). To evaluate the role of obesity associated SNPs in regulating the expression of *PEMT* gene in adipose we genotyped four tag-SNPs (including two nsSNPs) in 154 subjects. The SNP rs4646343 in intron1 showed the strongest association with *PEMT* expression (P_{res} = 3.21E-04). *PEMT* expression was higher in obese insulin resistant subjects and allele G of rs4646343 was associated with higher WHR (P= 2.47E-05) in Caucasians and with higher expression of *PEMT*. Significant positive correlation of BMI and % fat mass was observed with expression of *PEMT* in adipocytes but not in the stromal vascular cells of adipose tissue. Thus our study indicates an important function of *PEMT* gene and its regulatory polymorphisms in obesity.

2012F

Identification of a *cis* regulatory variant that binds a transcriptional repressor complex including PDX1 at the *JAZF1* type 2 diabetes locus. *M.P Fogarty[1], T.M Panhuis[1,2], S. Vadlamudi[1], M.L. Buchkovich[1], K.L. Mohlke[1].* 1) Dept Genetics, Univ North Carolina, Chapel Hill, NC.Department of Genetics, University of North Carolina, Chapel Hill, NC; 2) Department of Zoology, Ohio Wesleyan University, Delaware, OH.

Genome-wide association studies (GWAS) have identified at least fifty genome-wide significant loci associated with type 2 diabetes (T2D). For many of these loci, association signals are localized to non-protein-coding intronic and intergenic regions, and the underlying functional variants and target genes remain unknown. For the T2D association signal in *JAZF1* intron 1, we hypothesized that the underlying risk variants(s) have a *cis* regulatory effect in islets or other T2D-relevant cell types. We prioritized variants for functional follow up studies of transcriptional activity using maps of experimentally-predicted open chromatin regions. All SNPs in strong linkage disequilibrium ($r^2 \geq .8$, 1000 Genomes CEU) with the GWAS index SNP rs849134 and also present in an islet FAIRE peak or DNase peak were tested for evidence of differential transcriptional activity in INS-1-derived 832/13 and MIN6 insulinoma cells. Twelve regions containing T2D-associated variants were tested for enhancer activity, including five located in regions of islet open chromatin. Three regions exhibited enhancer activity, and of these only rs1635852 displayed allelic differences in enhancer activity. The rs1635852 T2D risk allele T showed less transcriptional activity than the non-risk allele C (832/13 P = 7.8 x 10^{-5}, MIN6 P = 1.0 x 10^{-5}). We assessed whether alleles of rs1635852 differentially bind nuclear proteins using electromobility shift assays. The T2D risk allele showed increased binding to protein complexes, suggesting that it functions as part of a transcriptional repressor complex. We applied DNA affinity capture and mass spectrometry to identify factors in the complex and determined that the T allele preferentially binds the pancreatic master regulator PDX1. These data suggest that maps of open chromatin are useful to guide identification of variants with allelic differences in regulatory activity at T2D loci and that rs1635852 in *JAZF1* intron 1 is part of a *cis* regulatory complex that affects transcriptional activity through binding of a protein complex that includes PDX1.

2013W

Investigation of mtDNA mutations in non-syndromic hearing loss patients in Fars province, Iran, using PCR- RFLP procedure. *S. Heydari Sodjani[1], M. Montazer Zohori[2], E. Farokhi[3], A. Shirmardi[4], G. Banitalebi[3], S. Reisi[3], M. Abolhasani[3], M. Akbari[2], M. Hashemzadeh[3].* 1) Department of Genetics, shahrekord university, Shahrekord, chaharmahal va bakhtiyari, Iran; 2) Department of Medical Genetics, School of Medical Sciences, Tarbiat Modares University, Tehran, I. R. Iran; 3) Cellular and Molecular Research Center, Shrekord University of Medical Sciences, Shrekord, I. R. Iran; 4) Welfare Organization of Chaharmahal va Bakhtiari Province, Shrekord, I. R. Iran.

Abstract: Background: Hearing loss is a sensorineural disorder that affecting 1 in 500 newborns. It happens due to genetic or environmental causes, or both. More than 60% of cases are inherited and 80% of hereditary cases are non-syndrome with autosomal recessive inheritance. In the present study we investigated the frequency of mtDNA A1555G, A3243 and A7445G mutations in patients in Fars province. Materials and Methods: Seventy two non-syndromic hearing impaired subjects were studied. DNA was extracted using standard phenol-chloroform method. The screening of the mitochondrial gene mutations were performed by PCR-RFLP procedure. Finally, the possible mutations were confirmed by direct sequencing. Results: None of the A1555G, A3243G and A7445G mutations were detected in this study. However, destroying a MTTL1 restriction site for the investigation of A3243G mutation, revealed a G3316A allelic variant in 1.4% of deaf subjects studied. Coclusion: Our data indicated that the mitochondrial A1555G, A3243G and A7445G mutations have any role in deaf subjects studied.

2014T

Genetic and environmental predictors of serum vitamin D3 levels in African Americans. *R.A. Kittles[1,2,3], K. Batai[2], E. Shah[2], M. Ruden[2], J. Newsome[2], S. Agate[2], A. Murphy[4], HY. Chen[3].* 1) Dept Med, M/C 767, Univ Illinois at Chicago, Chicago, IL; 2) Institute of Human Genetics, University of Illinois at Chicago, Chicago, IL; 3) Division of Epidemiology and Biostatistics, School of Public Health, University of Illinois at Chicago, Chicago, IL; 4) Department of Urology, Feinberg School of Medicine, Northwestern University, Chicago, IL.

Vitamin D deficiency is more common among African Americans (AAs) than European Americans (EAs). Vitamin D has been shown to inhibit growth and induce apoptosis in several cancers. Also epidemiologic evidence has linked vitamin D status to cardiovascular disease, diabetes, and autoimmune disease. Several GWAS studies in EAs have found vitamin D metabolic pathway gene variants associated with serum 25(OH)D level, but few of these SNPs have been replicated in AAs. It has been difficult to replicate theses SNPs mainly because serum 25(OH)D levels are influenced by skin color, genetic ancestry, age, BMI, and environmental factors such as sunlight exposure and diet. Also studies fail to look at the interaction of these genetic and environmental factors. Here, we test if vitamin D pathway gene SNPs, genetic ancestry, sunlight (UV) exposure, skin color, season, geography, and education were associated with serum 25(OH)D concentration and/or vitamin D deficiency in AAs. Thirty-seven variants in *GC, DHCR7/NADSYN1, VDR, CYP2R1, CYP27A1, CYP27B1, CYP3A4, and CYP24A1*, including 19 GWAS identified variants were genotyped in 1,000 AAs from Washington, D.C. and Chicago. We performed linear and logistic regression analyses including genetic ancestry, age, skin color as measured using reflectance (M-index), season of blood draw, vitamin D intake, education, and geographic region. Thirteen SNPs (10 of which were GWAS identified; four in *GC*, four in *CYP2R1*, and two in *DHCR7/NADSYN1*) were associated with serum 25(OH)D levels and/or vitamin D deficiency. The *CYP2R1* SNP, rs2060793 revealed the strongest association (P<0.001) with serum 25(OH)D levels. Three SNPs, including the *VDR* SNP rs11568820 (*Cdx2*), were also significantly associated in our full model. Significant interactions were observed between several genetic and environmental factors. Among darker pigmented AAs (M-index >58.6), rs12800438 in *DHCR7/NADSYN1*, had the strongest association (P=0.004) with the risk of vitamin D deficiency (<20 ng/ml) (OR=0.40; CI 0.22-0.75). Our study replicated the findings from previous GWAS in EAs and found that many of the same variants were associated with serum 25(OH)D and/or vitamin D deficiency in AAs if skin color, vitamin D intake, and UV exposure were included in the model. Our findings provide insights on the biological and environmental modifiers of serum vitamin D3 and will help guide future studies on the role of vitamin D in high risk populations such as AAs.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2015F

Vitamin D receptor ChIP-seq in human CD4+ T-cells and association with multiple sclerosis. *A. Handel[1,2], G. Disanto[1,2], G. Giovannoni[2], G. Ebers[1], S. Ramagopalan[1,2].* 1) Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom; 2) Blizard Institute of Cell and Molecular Science, Barts and The London School of Medicine and Dentistry, London, United Kingdom.

Vitamin D insufficiency has been associated with risk of multiple sclerosis (MS) in prospective case-control studies. We recently demonstrated an association between MS genetic risk loci and vitamin D receptor (VDR) binding in a lymphoblastoid cell line. However, no studies have examined the relationship between MS risk loci, chromatin architecture, gene expression and VDR binding in primary immune cells. We extracted CD4+ T-cells from peripheral blood in nine healthy Caucasian individuals by MACS. We studied genome-wide VDR binding using chromatin immunoprecipitation and massively parallel DNA sequencing (ChIP-seq). RNA-seq was undertaken on a subset of these samples. Vitamin D levels for each sample were determined by LC/MS. VDR binding peaks were determined by ZInba. MS risk SNPs were obtained from the Wellcome Trust Case Control Corsortium (Sawcer et al. Nature 2012). Overlap with MS risk intervals (100kb either side of each SNP), DNase I hypersensitivity sites (from UCSC) and RNA-seq peaks was performed in the Genomic Hyperbrowser. The number of VDR binding sites was far higher in the 5 individuals with 25-hydroxyvitamin D levels >75 nmol/L than in the 4 individuals with levels <75 nmol/L. There was significant enrichment for VDR binding near MS risk loci, DNase I hypersensitivity peaks and RNA-seq expression peaks. There was significant overlap between VDR binding in primary CD4+ T-cells and our previously published lymphoblastoid cell line VDR ChIP-seq. There is good a priori evidence for an epidemiological association between vitamin D levels and MS susceptibility. Our VDR ChIP-seq data support a potential role for vitamin D in gene-environment interactions determining MS risk. VDR-binding is associated with underlying chromatin architecture. Importantly the association between VDR binding sites and MS risk loci is observed both in cell lines and in primary CD4+ T-cells.

2016W

Gene-Gene Interaction in disease association for systemic lupus erythematosus in Asian populations. *J. Yang[1], Y. Zhang[1], X. Zhang[2], Y. Lau[1], W. Yang[1].* 1) University of Hong Kong, Pokfulam, Hong Kong; 2) Anhui Medical University, Hefei, China.

Systemic lupus erythematosus (SLE) is characterized as an autoimmune disorder with unclear etiology. Genome-wide association study (GWAS) has been proved to be a powerful approach for uncovering association between genetic variation and disease risk. Single locus analysis is widely used in most GWAS to date. However, all the confirmed susceptibility genes still only explain a small fraction of disease heritability. Gene-gene interaction may play a role in disease association, but has so far been missing in the bigger picture of connection between genetic factors and complex diseases. In this study, genome-wide gene-gene interaction was analyzed based on two GWAS data sets from two different regions in China, namely Hong Kong and Anhui. Three statistical methods installed in PLINK and MDR were used to calculate the genetic epistasis using logistic regression, machine learning and information gain theory. First, logistic regression was used for preliminary genome-wide selection of pairwise variant-variant interactions in the two independent GWAS datasets. A number of SNP pairs showed statistically significant P value on genetic interaction. These SNP pairs were kept for further validation if one or both of which are located in a region with annotated biological function. Selected SNP pairs were then analyzed by two MDR methods. By machine learning method, new variables were constructed to classify high risk and low risk genotype combinations. Further, information gain theory was used to detect whether the information between two SNPs were redundant. A number of variant pairs were found to have significant epistatic interaction in disease association in both Hong Kong and Anhui GWAS, such as interaction between SNPs in MTHFR and CXCR4. Interactions between established loci were also confirmed by our data, such as that between BANK1 and BLK. These findings are currently being replicated by a larger data set in independent cohorts.

2017T

Prevalence of diabetes-associated gene variants and its association with blood glucose levels in the Algarve population, Portugal. *M. Barreto da Silva[1], V. Gaio[1], A. Fernandes[2], F. Mendonça[3], F. Orta Gomes[3], A. Beleza[2], A. Gil[1], M. Bourbon[1], A.M. Vicente[1], C.M. Dias[1].* 1) Instituto Nacional de Saúde Doutor Ricardo Jorge, Lisboa, Portugal; 2) Laboratório de Saúde Pública Dra. Laura Ayres, Faro, Portugal; 3) Administração Regional de saúde do Algarve, Faro, Portugal.

The global rise in incidence of type 2 (T2D) has been called a pandemic, constituting a major public health concern. Although environmental factors play a substantial role in the etiology of T2D, genetic susceptibility has been established as a key component in T2D risk. Given the absence of studies regarding the prevalence of T2D associated variants in the Portuguese population, our aim was to determine the prevalence of disease-associated variants and determine its relative contribution to this phenotype. For this purpose, we have recruited 221 individuals (93 males and 128 females), between 26-91 years old (mean age 57.1), who were enrolled in the Health Centre of S. Brás de Alportel (Algarve). For each participant, we have measured total glucose levels and collected DNA. In addition, each participant has answered an exhaustive questionnaire including socio-demographic information, health history and lifestyle. We have selected and analysed three of the most significant loci previously reported to be associated with T2D in Caucasian populations (*TCF7L2* rs7903146, *PARPG* rs1801282 and *FTO* rs9939609) and performed an association analysis between glucose levels in this population and the selected gene variants. The mean total population glucose level was 103.85±35.3 g/dl. We found a significant difference in the mean glucose levels between males (mean = 111.5±51.3 g/dl) and females (mean = 98.4±17.6 g/dl) (Mann-Whitney test $P < 0.001$). The relative allele frequencies of the genotyped variants have been established. Genotype distribution for all investigated SNPs was in Hardy-Weinberg equilibrium. We found a marginal association between glucose levels and genotypes at the *TCF7L2* locus (Mann-Whitney test $P = 0.045$) in females but not in males, with carriers of the T allele displaying higher levels of blood glucose than homozygous for the A allele. This difference is also observed in males, although not reaching significance. No association was found between glucose levels and the other genotyped variants. These results suggest that the pathophysiology of the disease may be different between males and females, or that environmental factors are influencing this trait in males. We are currently investigating the later hypothesis by increasing our sample size and by analysing lifestyle information provided by the participants in order to evaluate gene-environment interactions influencing glucose levels in the Portuguese population.

2018F

Genome-wide methylation analysis of DNA from offspring exposed to a diabetic intrauterine environment. *M. del Rosario, R.L. Hanson, V. Ossowsky, W.C. Knowler, C. Bogardus, L.J. Baier.* Phoenix Epidemiology and Clinical Research Branch, NIDDK, NIH, Phoenix, AZ.

The offspring of mothers who had diabetes during pregnancy have high risk of developing type 2 diabetes and obesity. We explored the possibility of DNA methylation as a mechanism through which epigenetic signals are passed on to the offspring. We studied DNA methylation in peripheral blood leukocytes from 14 non-diabetic Pima Indians who were offspring of diabetic mothers (ODM) and 14 offspring of nondiabetic mothers (ONDM). These two groups were matched for age (mean±SD=16.1 ± 5.4), sex, age of mother, and fraction of Pima ethnicity to reduce confounding from other factors. To find regions that are differentially methylated between the two groups, a MeDIP-chip assay was performed on the Affymetrix Human Tiling 2.0R Array. Data were analyzed using the model based analysis of tiling arrays (MAT) algorithm implemented in the Partek Genomic Suite using the mean of the corrected probe on 600 bp window to asses methylation (quantified by the MAT score). Differences in MAT score between ODM and ONDM groups were assessed by ANOVA. Initially we restricted the analysis to 50,726 regions that overlap with known promoter regions (between -2 kbp and +500 bp of Refseq Bld 36 transcription start sites), resulting in analysis of 50,726 regions across 35,830 different promoters. We analyzed 80 permutations to generate the expected distribution of p-values under the null hypothesis for determination of the empirical False Discovery Rate. Based on the one-sided Kolmogorov-Smirnov test, there is a significant enrichment of low p-values in the observed data compared with the expected distribution if there was no differential methylation between the two groups ($p<0.001$). 54% of the differentially methylated regions were hypomethylated in ODM. A total of 5,975 promoters that were differentially methylated (nominal $p<0.005$, FDR~90%) were subjected to KEGG pathway analysis using DAVID. The top pathway results were Type II Diabetes ($p<0.003$) and maturity onset diabetes of the young (MODY) ($p<0.005$). These preliminary findings support the hypothesis that epigenetic dysregulation of genes known to be involved in glucose metabolism mediate the increased risk for diabetes in ODM.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2019W

HNF1α and ABCA1 genes polymorphisms in gestational diabetes mellitus. *E. Zamarron-Licona[1], M.C. Martinez-Lopez[1], R. Diaz-Martinez[1, 2]*. 1) DIAGNOSTICO MOLECULAR, UNIVERSIDAD JUAREZ AUTONOMA DE TABASCO, VILLAHERMOSA, TABASCO, Mexico; 2) Especialidad en Genética de la Universidad de Guadalajara.

Gestational Diabetes Mellitus (GDM) is defined as any degree of glucose intolerance with onset or first recognition during pregnancy. Hepatocyte nuclear factor-1 homeobox A (HNF1A) single-nucleotide polymorphisms (I27L, S487N) are associated with monogenic diabetes with early-onset. ATP binding cassette A1 (ABCA1) single-nucleotide polymorphisms (R230C) have been associated with Tangier's disease and familial high-density lipoprotein deficiency. This study aims to investigate the relationship of HNF1A and ABCA1 SNPs with gestational diabetes mellitus, as well as to characterize them in contrast with non-diabetics and family history of type 2 Diabetes Mellitus. 91(age at onset > 40 y.o.; diabetes in 3 contiguous generations) along with 58 non-diabetic and family history patients and 60 normoglycemic controls (N group) have been evaluated for gestational diabetes mellitus and SNPs of HNF1A and ABCA1. HNF1α (L27L) GDM group 43%, N group 26% and T2DM family history group 24%. (I27I) GDM group 21%, N group 32% and T2DM family history group 24%. (I27L) GDM group 36%, N group 42% and T2DM family history group 52%. p: 0.027695 (N487N) GDM group 27%, N group 7% % and T2DM family history group 2%. (S487S) GDM group 41%, N group 18% and T2DM family history group 22%. (S487N) GDM group 32%, N group 75% % and T2DM family history group 76%. p: 0.025427 ABCA1 (R230C) GDM group 18%, N group 23% and T2DM family history group 22%. (C230C), GDM group 82%, N group 77% and T2DM family history group 78%. (R230R) Not found p: 0.097375. HNF1α polymorphisms I27L and S487N are associated to gestational diabetes mellitus. In spite of R230C variant of ABCA1 appeared in a higher frequency for that described for Mexican population, It does not contribute to gestational diabetes mellitus. 1.Villarreal-Molina, M. T., et al(2008) Diabetes 57, 509-513 2.Bjorkhaug, L., et al J. Clin. Endocr. Metab. 88: 920-931, 2003. 3.Fernando MA Giuffrida,et al. Cardiovascular Diabetology 2009, 8:28.

2020T

External sources of vitamin D modify the effects of the GC and CYP2R1 genes on 25-hydroxyvitamin D concentrations: CAREDS. *C. D. Engelman[1], K. J. Meyers[2], S. K. Iyengar[3], Z. Liu[2], C. Karki[2], R. P. Igo, Jr.[3], B. Truitt[3], J. Robinson[4], G. E. Sarto[5], R. Wallace[4], L. Tinker[6], E. LeBlanc[7], Y. Song[8], J. A. Mares[2], A. E. Millen[9]*. 1) Population Health Sciences, University of Wisconsin, Madison, WI; 2) Ophtalmology and Visual Sciences, University of Wisconsin, Madison, WI; 3) Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, OH; 4) Epidemiology, University of Iowa, Iowa City, Iowa; 5) Obstetrics and Gynecology, University of Wisconsin, Madison, WI; 6) Fred Hutchinson Cancer Research Center, Seattle, WA; 7) Kaiser Permanente Center for Health Research, Portland, OR; 8) Division of Preventive Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA; 9) Social and Preventive Medicine, University at Buffalo, Buffalo, NY.

Vitamin D deficiency (defined by the blood concentration of 25-hydroxyvitamin D [25(OH)D]) has been associated with many adverse health outcomes. Genetic and non-genetic factors account for variation in [25(OH)D], but the role of interactions between these factors is unknown. To assess this, we examined 1,204 women of European descent from the Carotenoids in Age-Related Eye Disease Study (CAREDS), an ancillary study of the Women's Health Initiative Observational Study. Twenty-nine SNPs in 4 genes, *GC*, *CYP2R1*, *DHCR7* and *CYP24A1*, from recent meta-analyses of genome-wide association studies of [25(OH)D], were genotyped. Associations between these SNPs and [25(OH)D] were tested using generalized linear regression under an additive genetic model adjusted for age, blood draw month and ancestry. Results were stratified by season of blood draw and, separately, vitamin D intake for 6 SNPs showing a significant association with [25(OH)D] at the $p<0.01$ level. Two non-synonymous SNPs in *GC* and 4 SNPs in *CYP2R1* were strongly associated with [25(OH)D] in individuals whose blood was drawn in summer ($p\leq0.0001$ for the most significant SNP in each gene) but not winter months ($p>0.05$ for all SNPs) and, independently, in individuals with vitamin D intake ≥400 ($p<0.0001$) but not <400 IU/day ($p\geq0.04$). Only one SNP in each of these genes was independently associated with [25(OH)D]: rs4588 in *GC* and rs2060793 in *CYP2R1*. A genetic risk score summing the number of risk alleles at these two SNPs was significantly associated with [25(OH)D] ($p<0.0001$). The analysis was, again, stratified by high and low external sources of vitamin D. Women with all 4 risk alleles (from the genetic risk score) had low mean [25(OH)D], regardless of summer/winter season or high/low vitamin D intake (46.5-47.6 nmol/L for all 4 groups); however, women with no risk alleles and either summer season or high vitamin D intake (≥400 IU/day) attained a mean [25(OH)D] of 71.4 and 71.9 nmol/L, respectively, compared to only 53.8 and 55.3 nmol/L, respectively, for women with no risk alleles and either a winter blood draw or low vitamin D intake (<400 IU/day), indicating a stronger genetic effect in the presence of external sources of vitamin D. This effect modification has important implications for the design of discovery and replication genetic studies for all health outcomes and for public health recommendations and clinical practice guidelines regarding achievement of adequate vitamin D status.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2021F

Mapping variation in the response to vitamin D in peripheral blood. *C. Jeong, J.C. Maranville, S. Baxter, A. Di Rienzo.* Department of Human Genetics, University of Chicago, Chicago, IL., USA.

Vitamin D (VD) functions as a critical regulator of the immune system and is likely to play a role in a variety of diseases such as asthma, autoimmune diseases, and cancers. In addition, because of the known inter-ethnic differences in circulating vitamin D levels, it has been proposed that variation in the vitamin D pathway underlies inter-ethnic differences in disease risk. The main mechanism of VD action is through activation of the vitamin D receptor (VDR), which in turn acts a transcription factor to induce transcriptional changes at its target genes. This initiates a cascade that, in the immune system, results in suppression of antigen-driven inflammatory reactions, such as proliferation of lymphocytes. To elucidate the genetic bases of inter-individual variation in VD response, we measured changes in gene expression levels and in the suppression of cell proliferation in cultures of peripheral blood mononuclear cells (PBMCs) from 88 healthy African American donors. Each PBMC sample was treated, in paired aliquots, with vehicle control or with VD. In a genome-wide association mapping of the suppression of cell proliferation by VD, we found two QTLs, one on chromosome 5 (p=9.10×10-8, Bonferroni-corrected p=0.045), and the other on chromosome 18 (p=1.06×10-7, Bonferroni-corrected p=0.053). Gene expression data suggest that these associations are mediated by transcriptional responses because: 1) association between suppression of cell proliferation and log2 fold change of gene expression was found at 857 genes (FDR=0.10), 2) a large proportion of these genes showed association between their log2 fold change and either of the two proliferation QTLs (650 of 857 genes with FDR = 0.05). In addition, we used a Bayesian method to identify polymorphisms that interact with VD treatment to influence transcriptional response. This method detected 5 high-confidence interaction eQTLs (posterior p > 0.7), which show an effect of genotype on expression only in the VD-treated expression, but in the vehicle control. Together, these findings provide new information on the genetic bases of variation in the physiological response to VD treatment, with important implications for the risk to common diseases and their inter-ethnic disparities as well as for the assessment of response in VD supplementation trials.

2022W

The Interaction between Adolescent Parental Knowledge and Genetic Risk for Alcohol Dependence Predicts Adult Alcohol Dependence. *J.L. Meyers[1], J.E. Salvatore[1], R.J. Rose[5], J. Kaprio[2, 3, 4], D.M. Dick[1].* 1) Human Genetics, Virginia Commonwealth University, Richmond, VA; 2) Institute for Molecular Medicine Finland FIMM, University of Helsinki, Finland; 3) Department of Public Health, Hjelt Institute, University of Helsinki, Finland; 4) Department of Mental Health and Substance Abuse Services, National Institute for Health and Welfare, Helsinki, Finland; 5) Indiana University, Bloomington, Indiana.

Background. Previous studies demonstrate that parental knowledge moderates latent genetic influences on adolescent externalizing behavior and alcohol use (Dick et al., 2007, Latendresse et al., 2010) as well as specific genetic predispositions, such as CHRM2, to predict adolescent externalizing behavior (Dick et al., 2009). Little is known however, about the longitudinal effects of the parental knowledge in moderating genetic risk for alcohol problems from adolescence into adulthood.**Methods.**This study examines whether parental knowledge in adolescence continues to moderate genetic influences on alcohol use in young adulthood. We approached this question using data from a longitudinal, population based twin sample, Finntwin12 (Kaprio et al., 1999). We first conducted twin analyses to examine whether parental knowledge (measured at age 14) moderated genetic and environmental influences on alcohol dependence symptoms at age 22. We then created genetic risk sum scores (Yang et al., 2009) using GWAS data available on the twins (scores were comprised of all SNPs associated at p<0.01 with DSM-IV Alcohol Dependence symptoms). Next, we examined the interaction between this aggregate measure of risk genes and parental knowledge, and its effect on age 22 alcohol dependence symptoms. **Results.** The twin analyses indicated that parental knowledge significantly moderates the genetic influences on alcohol dependence symptoms at age 22 (χ2=10.31, p<0.0001). The genotypic analyses indicated that the interaction between genetic risk sum scores and parental knowledge significantly predicted alcohol dependence symptoms at age 22 (β=0.308, p<0.001). **Conclusion.** Converging evidence from two analytic methods suggests that parental knowledge in adolescence has an enduring moderating influence on genetic predispositions to alcohol use disorders in young adulthood. Parental knowledge may be an important proxy for some stable aspect of the individual's environment from adolescence into early adulthood, or may scaffold the adolescent's burgeoning behavioral regulation skills. There is a need for future research to elucidate the depth and limitations of the lasting effects of this aspect of adolescent parenting throughout development.

2023T

Association study of FUT2 (rs601338) with celiac disease and inflammatory bowel disease in the Finnish population. *A.S. Parmar[1], N. Alakulppi[2], P. Paavola-Sakki[3], K. Kurppa[7], L. Halme[4], M. Färkkilä[3], U. Turunen[3], M. Lappalainen[1], K. Kontula[5], K. Kaukinen[6], M. Mäki[7], K. Lindfors[7], J. Partanen[2], P. Sistonen[2], J. Mättö[2], P. Wacklin[2], P. Saavalainen[1], E. Einarsdottir[1].* 1) Research Program for Molecular Medicine and Department of Medical Genetics, Haartman Institute, University of Helsinki, Helsinki, Finland; 2) Finnish Red Cross Blood Service, Helsinki, Finland; 3) Department of Gastroenterology, Helsinki University Hospital, Helsinki, Finland; 4) Department of Surgery, Helsinki University Hospital, Helsinki, Finland; 5) Department of Medicine, Helsinki University Central Hospital, Helsinki, Finland; 6) Department of Gastroenterology and Alimentary Tract Surgery, Tampere University Hospital and School of Medicine, University of Tampere, Tampere, Finland; 7) Pediatric Research Center, University of Tampere and Tampere University Hospital, Tampere, Finland.

Homozygosity for a non-sense mutation in the fucosyltransferase2 (FUT2) gene (rs601338G>A) leads to the absence of secreted Lewis blood group antigens and consequently ABH blood groups in body fluids. FUT2 non-secretor status has previously been associated with risk of type-1 diabetes and Crohn's disease (CD). We studied the association of rs601338-FUT2 with celiac disease (CelD) and inflammatory bowel disease (IBD) in the Finnish population. rs601338 was genotyped in CelD (n=909), dermatitis herpetiformis (DH) (n=116), ulcerative colitis (UC) (n=496) and CD (n=280) patients and healthy controls (n=2738). CelD showed significant genotypic (P=0.0074, OR:1.28) and recessive (P=0.015, OR:1.28) association with the rs601338-AA genotype. This was also found in the combined CelD+DH dataset (genotype association: P=0.0060, OR:1.28; recessive association: P<0.011, OR:1.28). The A allele of rs601338 showed nominal association with dominant protection from UC (P=0.044, OR:0.82) and UC+CD (P= 0.035, OR: 0.84). The frequency of non-secretors in controls, CelD, DH, UC and CD datasets was 14.7%, 18%, 18.1%, 14.3% and16.1% respectively. No association was evident in the DH or CD datasets alone, possibly due to lack of power. FUT2 non-secretor status is associated with CelD susceptibility and FUT2 secretor-status may also play a role in UC in the Finnish population.

2024F

Identifying novel gene-environment interactions for HDL-cholesterol. *L. Parnell, K. Richardson, C.Q. Lai, J.M. Ordovas.* Nutrition & Genomics, JM-USDA Human Nutrition Research Center on Aging, Boston, MA.

Many genetic variants associating with clinical measures of metabolic diseases have been uncovered by genome-wide association studies (GWAS). While many of those associations replicate, many fail to validate in other populations, likely due to genetic and environmental differences. However, a substantial number of gene-environment (GxE) interactions have been described from approaches examining specific genes where environment equals diet, physical activity (PA), tobacco/alcohol use and sleep. Mining our database of over 500 GxE interactions involving obesity and blood lipids from disparate research reports, we built a gene network for interactions involving PA. This network of interacting genes and proteins was then supplemented with exercise-based gene expression data, HDL metabolism pathway membership, and main effect associations for HDL from GWAS all to prioritize candidate genes to test for novel GxE interactions for HDL-cholesterol and physical activity. Further refinement of the list of candidate HDL-PA interaction genes was done by measuring the distance across the network from the candidate gene to a known HDL-PA gene as identified in our GxE database. This resulted in 43 top candidate genes, which have been selected for analysis in a deeply phenotyped population. While analysis of the GxE interactions is ongoing, initial results suggest this is a viable approach to identify novel GxE interactions.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2025W

Genetically-determined differences in arsenic metabolism efficiency influence risk for arsenic-induced skin lesions: A Mendelian Randomization Study. *B. Pierce, L. Tong, M. Argos, F. Jasmine, M. Kibriya, H. Ahsan.* Health Studies, Univ Chicago, Chicago, IL.

Arsenic is a class-I human carcinogen, and arsenic contamination of drinking water is a serious public health issue in many countries, increasing risk for a wide array of diseases, including cancers of the bladder, kidney, lung, and skin, as well as neurological and cardiovascular diseases. There is substantial inter-individual variation in arsenic metabolism efficiency, as some individuals are able to methylate, and thus excrete, arsenic more efficiently than others. Consumed arsenic enters the blood as inorganic arsenic (iAs) and undergoes a series of methylation steps to produce monomethylarsonic acid (MMA) and then dimethylarsinic acid (DMA). Using data on urinary arsenic metabolite concentrations of iAs, MMA, and DMA and genome-wide single nucleotide polymorphisms (SNPs) for 2,060 arsenic-exposed Bangladeshi individuals, we show genome-wide significant associations for urinary concentrations of both MMA% ($P=2x10^{-11}$) and DMA% ($P=2x10^{-14}$) in a large linkage disequilibrium block containing the AS3MT gene (arsenite methyltransferase; 10q24.32), with five genetic variants showing evidence of independent association. Although the association signals for MMA% and DMA% appear to be driven by different SNPs, haplotype analysis of these five variants suggests that MMA and DMA are to a large degree influenced by the same underlying variants. In a Mendelian randomization analysis of 2,073 skin lesion cases (the classical sign of arsenic toxicity) and 3,267 controls, we used two-stage regression methods to assess the effect of genetically-determined differences in arsenic metabolism efficiency on skin lesion risk. Results from both SNP-based and haplotype-based Mendelian randomization analysis support the hypothesis that iAs and MMA are more toxic than DMA (i.e., increased methylation efficiency decreases risk for arsenic toxicity, presumably because DMA is more readily excreted in urine and expelled from the body). For example SNP-based MR analyses produce causal ORs of 1.25 (CI: 1.17-1.33) for a 1-unit change in iAs%, 1.15 (CI: 1.09-1.21) for a 1-unit change in MMA%, and 0.89 (CI: 0.87-0.92) for a 1-unit change in DMA%). Our results support a causal role for arsenic methylation efficiency in arsenic toxicity. Knowledge of these biological processes could be used to develop intervention and pharmacological strategies aimed at preventing large numbers of arsenic-related deaths in arsenic-exposed populations.

2026T

The Ubiquitin Proteasome System: Gene-Environment Interactions in Parkinson's Disease. *S. Rhodes[1], B. Ritz[1,2].* 1) Dept Epidemiology, UCLA Sch Public Hlth, Los Angeles, CA; 2) Dept Neurology, Geffen Medical School, UCLA, Los Angeles, CA.

The underlying causes of Idiopathic Parkinson's disease (PD) are likely complex interactions between environmental exposures and genetic variants that confer a predisposition towards damage resulting from the relevant environmental exposure. Pesticide exposure or its proxy (e.g. rural living, well-water consumption, farming occupation) are regularly observed in association with PD. Animal studies have demonstrated that multiple different types of pesticides are toxic to dopamine neurons and that the mechanisms of toxicity tend to fall in one or the other known biologic pathways implicated in PD etiology including mitochondrial dysfunction, oxidative stress, and ubiquitin proteasome system (UPS) disruption. We investigated in the population-based Parkinson's Environment and Genes (PEG) study possible gene-environment (GxE) interactions between five genes hypothesized to have a role in UPS or mis-folded protein trafficking and five pesticides determined to inhibit the UPS in an SK-N-MC GFP^u model assessing 26S activity in living cells. The PEG study is unique in that pesticide exposure was assessed using state-mandated pesticide use reports, land use maps, and residential and occupational histories from each subject, thereby minimizing subject recall bias with regards to pesticide exposure. 354 cases and 543 controls were genotyped for single nucleotide polymorphisms (SNPs) in PARK2, UCHL1, SKP1, UBA1, and UBA6; pesticides considered included benomyl, endosulfan, ziram, dieldrin, and rotenone. Four SNPs, one in each gene except UCHL1, were suggestive ($p<0.10$) of a 30-40% change in PD risk per variant allele and were considered for GxE interaction with pesticides. Two SNPs, rs2284312 in SKP1 and rs354872 in UBA6, were associated ($p<0.05$) with a 2- to 2.5-fold increase in PD risk under the additive genetic model but only in the presence of one or more UPS-inhibiting pesticides. These SNPs were not associated ($p>0.15$) with any increase in PD risk in the absence of pesticide exposure. rs9365292 in PARK2 was suggestive of a GxE interaction. These findings are consistent with the multi-factorial etiology of PD and are suggestive of a greater need to integrate genetic investigations with epidemiologic studies so as to better understand the combined impact of genes and environment.

2027F

GENETIC TRADE-OFFS MAY EXPLAIN SOME PARADOXES OF GENETICS OF HUMAN LONGEVITY. *S. Ukraintseva, K. Arbeev, A. Kulminski, I. Akushevich, D. Wu, A. Yashin.* Duke University, Durham, NC.

Chronic diseases, such as CVDs, cancer, diabetes and AD, are major contributors to mortality in old age. It might therefore be expected that genetic factors which increase risks of such diseases will negatively affect person's survival. However, several studies suggested that the presence of "risk alleles" in individual genomes does not necessary compromise longevity, and that the proportion of such alleles among centenarians is often similar to that in younger controls (e.g., Beekman et al. 2010; Sebastiani et al. 2011). It has also been shown that the frequency of some "risk alleles" may change non-monotonically with age, e.g., decrease from youth to the age 80, and increase afterwards (Yashin et al. 1999; Bergman et al. 2007). Few explanations of this phenomenon have been suggested (Yashin et al. 1999, 2001; Bergman et al. 2007), while more comprehensive investigation of its mechanism is required and critically important for understanding the relationships between genetic factors influencing health and longevity. Here we propose that the apparent paradox of accumulating "harmful" alleles in long-living individuals can plausibly be explained by "genetic trade-offs". The latter in our case includes the situations when a genetic variant: (i) increases the risk of a disease (e.g. cancer) in middle/old age, but decreases that risk at the oldest old ages; (ii) increases the risk of one disease (e.g. CHD) at middle/old age, but decreases the risk of another life-threatening condition (e.g., internal bleeding) at older ages; (iii) increases the risk of a disease (e.g. cancer), but slows/postpones some physiological aging changes, so it becomes advantageous for person's survival at the very old age, when senescent state itself becomes leading cause of death. We provide evidence from our recent research and work of others in support of this explanation and conclude that the paradoxical accumulation of "risk alleles" in long-living individuals may in fact be a common consequence of a normal change in phenotypic effects of genes with age in constantly changing environment in aging human body, as well as be a result of antagonistic pleiotropic influence of some genes on different health traits.

2028W

Interaction analysis of exogenous estrogen in age-related macular degeneration (AMD): new gene-based 2-degree-of-freedom (2df) joint test finds joint effects within the VEGF signaling pathway. *MD. Courtenay[1,2], WH. Cade[1], PL. Whitehead[1], SG. Schwartz[3], JL. Kovach[3], G. Wang[1], A. Agarwal[4], JL. Haines[5], MA. Pericak-Vance[1,2], WK. Scott[1,2].* 1) Human Genetics and Genomics, Hussman Institute for Human Genomics, Miami, FL; 2) Dr. John T. Macdonald Foundation Department of Human Genetics, University of Miami Miller School of Medicine, Miami, FL; 3) Bascom Palmer Eye Institute, University of Miami Miller School of Medicine, Miami, FL; 4) Ophthalmology, Vanderbilt University, Nashville, TN; 5) Center for Human Genetics Research, Vanderbilt University, Nashville, TN.

AMD is the leading cause of irreversible vision loss in developed countries. Its multifactorial etiology includes both genetic and environmental risk factors. Overexpression of VEGF promotes angiogenesis, making it a key influence on the choroidal neovascularization (CNV) found in advanced AMD. Anti-VEGF therapy is the primary treatment for CNV. VEGF expression can be modulated by exposures to exogenous estrogen, and women who take exogenous estrogen in the form of hormone replacement therapy (HRT) have reduced risk of AMD. This study tested the association of 74 genes in the VEGF pathway for association with AMD, accounting for genetic effects and HRT use using a new gene-based test of genetic and environmental factors. SNPs (N=1119) were selected from a previous AMD GWAS that included 312 Caucasian women with AMD (AREDS Grade 5 CNV) and 179 unaffected women controls (AREDS Grade 1). History of ever taking HRT was collected with a self-administered questionnaire. The gene-based test was conducted as an extension of the set-test implemented in PLINK, using the chi-square statistic at each SNP derived from Kraft's 2 degree of freedom (df) joint test of genetic and environmental factors. Using a function written in R, the 2 df joint test utilized a likelihood ratio test computed from a "full" logistic regression model with SNP and HRT main effects, a pairwise interaction term, age, and smoking status, and a "restricted" model excluding the SNP and interaction terms. Gene-based chi-square test statistics were calculated as the mean of all independent ($r2<0.8$) joint test statistics. The analysis was then repeated 10,000 times, permuting phenotype labels to generate an empirical p-value. The same set-test was conducted without incorporating HRT. Although no results reached Bonferroni corrected significance ($p<3.4x10\text{-}4$) for either analysis, nominally significant results were obtained at five genes: NFATC1 (joint test p=0.0036), VEGFA (p=0.0043), MAPKAPP3 (p=0.0072), PLA2G6 (p=0.0288), and RAF1 (p=0.032). Only the test incorporating HRT detected the effect at PLA2G6, (main effects gene-based set test p=0.1239). PLA2G6 plays a role in arachidonic acid metabolism which leads to inflammation, a feature of AMD pathogenesis. These results illustrate the utility of incorporating environmental exposures in tests of genetic association in complex diseases, at the SNP, gene, or pathway level.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2029T

Heritability and host SNP associations of microbial species in the human gut. *J.K. Goodrich[1,2], R. Blekhman[1,3], J.T. Bell[4], T.D. Spector[4], A.G. Clark[1], R.E. Ley[1,2].* 1) Dept of Molecular Biology & Genetics, Cornell University, Ithaca, NY, USA; 2) Dept of Microbiology, Cornell University, Ithaca, NY, USA; 3) Dept of Neurology & Neuroscience, Weill Cornell Medical College, New York, NY USA; 4) Dept of Twin Research & Genetic Epidemiology, King's College London, U.K.

The composition of the human gut microbiome is shaped by several factors including environment, diet, and host disease status. Variation in the human genome would also be expected to impact many aspects of the gut environment that in turn should impact microbiome composition, but this aspect is under-explored. In order to determine the degree to which species composition of the microbiome is heritable, and to identify associations of specific regions of the genome with particular attributes of the human gut microbiome, we characterized the gut microbial communities of 165 dizygotic (DZ) twin pairs and 118 monozygotic (MZ) twin pairs from the United Kingdom Adult Twin Registry (TwinsUK). Each individual provided a single stool sample. The environmental stability of the gut microbiota over time was assessed through repeat sampling of a subset of the individuals. The 16S rRNA gene sequences from microbes in the stool samples were sequenced in sufficient abundance to produce a phylogenetic representation of the microbiota species composition. We determined the degree to which the microbial tree is heritable using the classical ACE twin model, which partitions the total variance of a phenotype into additive genetic effects (A), common environment (C), and unique environment (E). We identify specific regions of the tree with moderate heritability such as members of the families Ruminococcaceae, Lachnospiraceae, and Catabacteriaceae, while other regions have little to no heritability. An assessment of the twin-pair correlations of microbial abundances throughout the phylogenetic tree, showed on average that MZ twin pair correlations were significantly greater than the DZ twin pair correlations, further indicating heritability of the microbiota. To determine specific regions of the genome influencing microbial composition we performed an association analysis of attributes of the microbiota including diversity measures and relative abundance of shared bacterial taxa. The genes near loci that are significantly associated with the composition of the gut microbiome are enriched in pathways involved in host immune response. This analysis opens a window into the unexplored world of host-genetic control over the microbiome, begging further investigation into specific genes and pathways and exploration of the interaction of host-genetics and the microbiome in the context of disease.

2030F

SNPs for BMI demonstrate consistent results through the lifecourse. *N. Heard-Costa[1,2], L.D. Atwood[1,2], C. Jaquish[3], C.T. Liu[4], L.A. Cupples[1,4], C.S. Fox[1,5].* 1) Division of Intramural Research, National Heart, Lung and Blood Institute, Framingham Heart Study, Framingham, MA; 2) Department of Neurology, Boston University School of Medicine, Boston, MA; 3) Division of Prevention and Population Sciences, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD; 4) Department of Biostatistics, Boston University School of Public Health, Boston, MA; 5) Division of Endocrinology, Metabolism, and Diabetes, Department of Medicine, Harvard Medical School, Boston MA.

Introduction: Body mass index (BMI) has a substantial genetic component, and 32 common genetic variants have been reported in association with BMI. However, few studies have investigated whether the associations of these variants with BMI are consistent throughout the lifecourse. In this study we take advantage of multiple observations available for participants in the Framingham Heart Study (FHS) to assess whether the genetic association of each of the 32 SNPs with BMI changes with age. Methods: We analyzed the association of BMI with 32 confirmed SNPs (Speliotes, 2010) in 8379 FHS participants. We grouped participants into 15 distinct 5 year intervals, beginning at age 15 and ending with those ≥ 85 years of age. For each participant, a single observation was chosen in each interval; if multiple observations were available, the one with the minimum age was retained for those under 50 and the one with maximum age was retained for those 50 and above. This resulted in a total of 33048 observations, with an average of 2203 observations per age interval. To account for familial relatedness, we used pedigree-based mixed effects models that allow for repeated measure analysis; the model included an age interaction term. To account for multiple testing, we considered a p-value < 0.0016 to be significant. Results: Mean BMI levels increased from 23.1 kg/m^2 at age 15, peaked at 27.6 kg/m^2 in the sixth decade of life, and declined after age 65 to 25.8 kg/m^2. Of the 32 SNPs tested, we generally observed consistent associations with BMI by age. For example, for SNP rs4836133, in ZNF608, the average effect size was 0.019 ± 0.086. Only 1 SNP, rs10150322 at NRXN3, revealed a significant interaction with age (p value 1.203 x 10^{-4}) with stronger effects observed in older individuals, with estimated mean BMI difference per copy of the risk allele ranging from 0.01 kg/m^2 at age 20 years to 0.51 kg/m^2 at age 70. Conclusion: These findings highlight how the majority of SNPs known to be associated with BMI show remarkable robustness of association through the lifecourse.

2031W

Examining the roles of diet, age, and sex on the composition of the human fecal microbiome. *ER. Davenport[1], O. Mizrahi-Man[1], LB. Barreiro[2], C. Ober[1], Y. Gilad[1].* 1) Department of Human Genetics, University of Chicago, Chicago, IL; 2) Sainte Justine Hospital Research Centre, Department of Pediatrics, Faculty of Medicine, University of Montreal, Montreal, QC, Canada.

The composition of the human microbiome is known to be influenced by many environmental factors, particularly diet. Host genetic variation also likely plays a role in shaping the composition of the microbiome. However, this has not yet been demonstrated in humans and it is important to minimize the strong influence of inter-individual dietary differences when attempting to identify the potentially small, multifactorial genetic effects. In this study, we sought to understand the relative importance of environmental and genetic factors on the composition of the gut microbiome by examining the bacterial communities of feces from members of an isolated population, the Hutterites of South Dakota. This communal-living, farming population prepares and eats all meals in a common dining room. Although their diet differs seasonally with respect to the proportion of fresh fruits and vegetables consumed, it is remarkably uniform among individuals within seasons. This allowed us to investigate the roles of both genetics (by examining inter-individual genetic variation within season, diets controlled) and the environment (by examining the same individuals between seasons, with diets differing) in shaping microbiome composition. To this end, we collected fecal samples from 138 adult Hutterites (51 men, 87 women; ages 16-76 years) in winter and summer months (33 men and 55 women sampled at both times). DNA was extracted from feces in duplicate, the 16S rRNA V4 region was amplified, each sample was sequenced to high depth (mean: 2,721,718 ± 1,451,405 reads per replicate) using Illumina HiSeq technology;, and reads were classified using mothur's classify.seq function. We observed high correlation between replicates (median spearman r=0.85-0.95, depending on taxonomic rank). As a first step, we examined age and sex effects on microbial diversity, controlling for season. Microbial diversity decreased with age (F-test; p=0.018). Additionally, in contrast to reports of mice and Rhesus macaques, there were no significant differences in Shannon diversity between males (n=48; mean= 2.27) and females (n=75; mean=2.28) (t-test; p=0.94). Our next steps include analyzing genome-wide SNP information to identify associations with microbiome abundance and diversity measures, in order to assess the role of genetics in maintaining microbiome composition.

2032T

Gene-to-gene interaction in the development of gout. *A. Taniguchi, W. Urano, N. Ichikawa, H. Yamanaka.* Institute of Rheumatology, Tokyo Women's Medical University, Shinjuku-ku, Tokyo, Japan.

Background: More than ten genetic loci have been reported to be associated with serum levels of uric acid (SUA). Most of those are polymorphisms in genes of renal transporters or genes associated with metabolic syndrome (MetS). Gout is a disease based on hyperuricemia and some of those genetic factors are expected to be associated with gout. In gout, renal excretion of uric acid is impaired and it is well known that more than half of gout patients have MetS. Some of renal transporters are responsible of renal reabsorption or excretion of urate. Furthermore, previous reports indicate the relationship between impaired renal urate excretion and metabolic syndrome. Therefore, it is important to investigate gene-to-gene interactions between genetic factors associated with SUA or gout. Purpose: To investigate gene-to-gene interactions between genetic factors associated with gout in Japanese. Method: 153 patients with gout and 534 Japanese male controls were subjects of the study. We first investigated the association of gout with genetic factors that have been reported the association with SUA by GWAS. Then, we investigated gene-to-gene interaction. Gene-to-gene interactions between pairwise combinations of SNPs were tested using logistic regression. For each pair of genetic variant, two models were developed. The model (A) was developed based on the assumption of interaction, while the model (B) developed without it. Model (A): g(P) = SNP1 risk allele (+) + SNP2 risk allele (+) + {SNP1 risk allele (+) x SNP2 risk allele (+) }; Model (B): g (P) = SNP1 risk allele (+) + SNP2 risk allele (+). A comparison between model (A) and (B) were performed using χ^2 test. Bonferroni correction for multiple testing was carried out where appropriate. Results: 12 polymorphisms in genes such as *PDZK1, GCKR, ABCG2, LRRC16A, LRP2, SLC17A1, SLC2A9, SLC16A9, SLC22A11* and *SLC22A12* were selected for this study. Polymorphisms in *GCKR, ABCG2, LRRC16A, LRP2, SLC17A1, SLC2A9* and *SLC22A12* showed significant association with gout. Then, we evaluated gene-to-gene interaction for each pair of those loci. The interaction between *SLC16A9* and *ABCG2* was shown with the p-value of 0.018. However, Bonferroni correction removed the significance. We did not detect significant gene-to-gene interaction between genetic polymorphisms of renal transporters. Conclusion: In gout, the role of gene-to-gene interaction between renal transporters or between renal transporters and MetS might be small.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2033F

The genetic basis of preeclampsia in populations adapted to high altitude. *P. Ortiz-Tello[1], K. Sandoval Mendoza[1], V. Villanueva Dávalos[2], A. Moreno Estrada[1], J. Manzaneda[4], F. Manzaneda[4], A. McKenney[5], L. Enriquez Lencinas[2,3], C. Bustamante[1], J. Baker[1].* 1) Dept Genetics, Stanford Univ Sch Med, Stanford, CA; 2) Dept of Obstetrics and Gynecology, Hospital Regional “Manuel Núñez Butrón”, Puno, Peru; 3) Dept of Pathology, Hospital Regional “Manuel Núñez Butrón”, Puno, Peru; 4) Universidad Nacional del Altiplano, Puno, Peru; 5) Dept Pathology, Stanford Univ Sch Med, Stanford, CA.

Preeclampsia - a disorder of pregnancy characterized by the sudden onset of maternal hypertension and proteinuria - complicates 3-8% of all pregnancies and is the largest killer of pregnant women worldwide. Preeclampsia also poses a risk to newborns, particularly in areas of the world where resources are limited. The heterogeneity of the disorder has posed a challenge in our understanding the molecular basis of the preeclampsia. High altitude populations provide an untapped resource to understand the genetics underlying preeclampsia due to the relative homogenous population living under strong selective forces. Moreover, at high altitudes, pregnant women have a 2-3 fold higher rate of preeclampsia than women at low altitudes. Importantly, it has been known for hundreds of years, and recently documented, that Native populations living in the Andes are far more equipped for pregnancy at high altitudes than their European counterparts. This is similarly true for the Native Tibetans compared to the recently immigrated Han Chinese. Thousands of years of genetic adaptation to the low oxygen environment at high altitudes may have conferred protection to these individuals against preeclampsia, although the mechanisms for this protection have not been elucidated. In order to investigate the genetic basis of preeclampsia at high altitude, we have embarked in an international collaboration with the Regional Hospital “Manuel Núñez Butrón” and the National University of the Altiplano (UNA) in Puno, Peru to sample affected individuals living at nearly 4,000 m.a.s.l. We have consented and collected maternal, paternal, and cord blood for genotyping and sequencing, as well as associated sections of placenta cut and embedded for histopathology on more than 400 trios (mother, father, baby), as well as pertinent medical records. We are applying admixture mapping, as well as methods of detecting positive selection, in combination with RNAseq data, histopathology, and quantifiable phenotype data, in order to identify new loci correlated with preeclampsia and shed light on better understanding human physiopathology.

2034W

The effect of race, sex, age, socioeconomic, behavioral factors, and genetics on dental caries in different tooth surfaces. *D. Lewis[1], E. Feingold[1,2], Z. Zeng[1,2], J. Shaffer[1], X. Wang[3], R. Weyant[4], D. McNeil[5], M. Marazita[1,3,6].* 1) Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA, USA; 2) Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA, USA; 3) Center for Craniofacial & Dental Genetics, Department of Oral Biology, School of Dental Medicine, Pittsburgh, PA, USA; 4) Department of Dental Public Health and Information Management, School of Dental Medicine, University of Pittsburgh, Pittsburgh, PA, USA; 5) Dental Practice and Rural Health, West Virginia University, Morgantown, WV, USA; 6) Clinical and Translational Science, School of Medicine, University of Pittsburgh, Pittsburgh, PA, USA.

Dental caries (cavities) is a multifactorial disease that involves many interacting variables to promote its development. It constitutes a major public health problem that begins in childhood. The prevalence of dental caries in adults from underserved and minority populations in the U.S can be over 90%. Risk factors associated with dental caries include environmental, behavioral, socioeconomic, and genetic factors. Many studies have demonstrated caries varying by race, gender and socioeconomic status (SES) and several genetic risk factors have been established as well. In this study, we ask whether these predictive factors have different effects on caries in different tooth surfaces (smooth surfaces vs. pit and fissure surfaces).

Families were recruited by the Center for Oral Health Research in Appalachia (COHRA) for a collection of demographic data, biological samples, and clinical assessment of oral health. There were a total of 1,129 Black and White adults examined in this study. We used exploratory statistics and regression analyses to examine the relationships among various demographic, SES, genetic and behavioral variables.

We confirmed associations between a number of known risk factors and caries, including sex, race, age and tooth brushing frequency. We also found that these risk factors are different for different tooth surfaces. Regression modeling was used to reach conclusions not only about which predictors are important and the size of their effects but, also the independence of predictors simultaneously affecting different tooth surfaces. Supported by; DE018903 and DE014899.

2035T

Telomere length in human blood cells and the prediction of survival. *J. Deelen[1,2], M. Beekman[1,2], V. Codd[3,4], H.E.D. Suchiman[1], A.J.M. de Craen[5], N.J. Samani[3,4], J.J. Houwing-Duistermaat[6], P.E. Slagboom[1,2].* 1) Section of Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands; 2) Netherlands Consortium for Healthy Ageing, Leiden University Medical Center, Leiden, The Netherlands; 3) Department of Cardiovascular Sciences, University of Leicester, Glenfield Hospital, Leicester, UK; 4) NIHR Leicester Cardiovascular Biomedical Research Unit, Glenfield Hospital, Leicester, UK; 5) Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, The Netherlands; 6) Section of Medical Statistics, Leiden University Medical Center, Leiden, The Netherlands.

Telomere length (TL) in human leukocytes declines with age and several studies have shown that decreased TL is associated with increased mortality, although not in all studied populations. Previous studies had a relatively small sample size and follow-up in larger populations is warranted. In addition, it is not clear whether telomere shortening measured in blood cells marks merely a history of cell division induced by infections or whether there is a causal contribution of length reduction to mortality. We measured TL in peripheral blood of 870 nonagenarian siblings (mean age 93 years), 1,580 of their offspring (mean age 59 years) and 725 spouses thereof (mean age 59 years.) from the Leiden Longevity Study. These participants have been followed up for vital status during 7.56 years on average. Mean TL was measured as a ratio (T/S) of telomere repeat length (T) to the copy number (S) of the single-copy gene *36B4*. There was no difference in TL between the offspring and spouses ($P = 0.927$), so we could analyse them as one group. Survival analysis, using Cox regression, in the nonagenarians showed that longer telomeres are associated with better survival into very old age (HR = 0.66, $P = 0.028$). This effect is even more pronounced in the offspring and their spouses (HR = 0.24, $P = 0.001$). To determine whether the association could be based on TL marking a history of infections, we determined the association between TL and markers for immune response (white blood cell counts, hsCRP and incidence of CMV infection) and a marker for cell replication (IGF1/IGFBP3) in the offspring and spouses. We conclude that decreased TL is significantly associated ($P < 0.05$) with decreased neutrophil counts, basophil counts and IGF1/IGFBP3 ratio, increased lymphocyte counts and a higher incidence of CMV infection. When the survival analysis of TL among the nonagenarians was adjusted for the mentioned parameters of cell replication and immune response, we found that TL still showed an independent association with survival (HR = 0.60, $P = 0.007$). We conclude that TL predicts survival in very old and middle age and that this effect is independent from markers for immune response and cell replication. This observation is currently being replicated in independent cohorts in which we will also determine the impact of known genetic determinants of TL on survival.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2036F

Early life infection associated with shorter adult blood telomere lengths in Cebu, the Philippines. *D.T.A. Eisenberg[1,2,3], J. Borja[4], M.G. Hayes[2,5,6], C.W. Kuzawa[2,3].* 1) Department of Anthropology, University of Washington, Seattle, WA; 2) Department of Anthropology, Northwestern University, Evanston, Illinois, 60208, USA; 3) Cells 2 Society: the Center for Social Disparities and Health, Institute for Policy Research, Northwestern University, Evanston, IL 60208; 4) Office of Population Studies Foundation, University of San Carlos, Cebu City, Philippines; 5) Division of Endocrinology, Metabolism and Molecular Medicine, Department of Medicine, Northwestern University Medical School, Chicago, Illinois, 60611, USA; 6) Center for Genetic Medicine, Northwestern University, Chicago, Illinois, 60611.

Telomeres are repetitive DNA sequences found at the ends of chromosomes. Telomere lengths (TL) shorten with age in proliferating human tissues due to the `end-replication problem' and oxidative stress. This shortening is implicated in senescence, with previous work suggesting that shorter TL impairs immune and cardiovascular function and results in increased mortality. Repeat activation of cellular immune responses, each involving extensive clonal expansion, is necessary to defend against rapidly reproducing and evolving pathogens. Because cellular division shortens telomeres and shortened telomeres are related to increased infections and subsequent mortality, ecologies in which immune activation are more extensive might also limit later immune function. Studies examining associations between infectious diseases and TL are primarily limited to clinical contexts in developed countries. To our knowledge, no study has evaluated the effect of infectious diseases on TL within the context of a non-clinical, low-income country nor examined the influence of early life infections on TL. To test the hypothesis that increased infections early in life cause shortened TL, TL from young men and women from Cebu, Philippines were measured (N = 1,639) from DNA extracts of blood using a qPCR method. Mothers were asked every two months for the first two years of the child's life how many diarrheal episodes the child had over the past week. Consistent with expectations, increased incidence of early life diarrheal episodes was associated with shorter TL (P = 0.04). On closer examination the association was most marked for diarrheal episodes which occurred when the individuals were between 8 and 12 months old=ma time when probably due to weaning and increased interaction with the environment diarrheal incidence is greatest. In order to consider the biological importance of this association we compared it to the age related decline in blood TL in 36 -69 year old Filipino women measured in our laboratory. The effect of having one more diarrheal episode in the past week before a survey in the second half year of life is equivalent to approximately three years worth of decrease in TL in adulthood. These results imply that infections early in life might predispose toward later life development of senescence and chronic disease mediated via TL shortening. This work was supported by NSF and Wenner Gren Foundation grants. DNA extracts generously provided by Karen Mohlke.

2037W

Characterization and generalization of HFE rs1800562 genotype-phenotype relationships in the diverse National Health and Nutrition Examination Surveys (NHANES) as part of the Epidemiologic Architecture for Genes Linked to Environment (EAGLE). *K. Brown-Gentry, B. Niloufar, J. Hailing, B. McClellan, J. Boston, C. Sutcliffe, H. Dilks, D. Crawford.* Ctr Human Gen, Vanderbilt Univ, Nashville, TN.

Hemochromatosis (HFE) rs1800562 (C282Y) is known for its strong association with hemochromatosis, a disease of excess iron storage from diet, and iron status biomarkers in European (EA) populations. In genome-wide association studies (GWAS) in EA descent populations, HFE rs1800562 has also been associated with red blood cell traits, cardiovascular risk factors, and markers of metabolic syndrome. Despite characterization of this common variant in EA descent populations, little is known about its frequency and genetic associations in diverse populations. To fill these gaps in knowledge, the Epidemiologic Architecture for Genes Linked to Environment (EAGLE) genotyped HFE rs1800562 in 14,998 participants from the National Health and Nutrition Examination Surveys (NHANES) III and 1999-2002. NHANES, a nationally representative sample of the United States is conducted by the National Center for Health Statistics at the Centers for Disease Control and Prevention. Among 6,634 non-Hispanic whites (NHW), 3,458 non-Hispanic blacks (NHB), and 3,950 Mexican Americans (MA), the minor allele was observed at a frequency of 6.28%, 1.88%, and 1.68%, respectively. Tests of association were performed using linear regression assuming an additive genetic model for the following dependent variables, stratified by self-reported race/ethnicity and adjusted for sex and age: serum transferrin saturation (%), serum iron (ug/dL), ln(serum ferritin)(ng/mL), mean cell volume (MCV; fL), and fasting plasma glucose (mg/dL). As expected, among NHW, serum iron (β=13.3), serum transferrin saturation (β=5.9), and MCV (β=1.2) were strongly associated with rs1800562 (p<2.0x10-7). Also associated in NHW was ln(serum ferritin) (β=0.12 and p=3.1x10-4). Among both NHB and MA, rs1800562 was associated with serum iron (p<0.01) and serum transferring saturation (p<3.0x10-4) with similar genetic effect sizes compared with NHW. MCV was associated with rs1800562 in MA (β= 2.3; p=1.5x10-7), but MCV was not associated in NHB possibly due to power. Neither glucose nor ln(serum ferritin) was associated with rs1800562 in NHB nor MA. Despite generalizing known associations between rs1800562 and iron trait markers to diverse populations, we failed to generalize all GWAS-identified genotype-phenotype relationships described in EA descent populations highlighting the need for further genetic association studies for HFE in large non-EA populations.

2038T

Analysis of Established Type 2 Diabetes Associated Variants in American Indians. *R. Hanson, S. Kobes, W. Knower, C. Bogardus, L. Baier.* Diabetes Epidemiology Clin Res, NIDDK, Phoenix, AZ.

Variants at several genetic loci have recently been reproducibly associated with type 2 diabetes mellitus (T2DM) in European and Asian populations. There is little information, however, on the relationship of these variants with T2DM in American Indian populations, many of which are at high risk for the disease. We genotyped a “sentinel” single nucleotide polymorphism (SNP) at 39 of these loci in 3401 Pima Indians (47% with diabetes) from a population-based study; acute insulin secretion was measured by an intravenous glucose tolerance test in 287 individuals. Generalized estimating equations, which account for familial relationships, were used to calculate the odds ratio (OR) for the association of T2DM with each risk allele, as determined by previous studies. Heterogeneity between Pimas and Europeans was assessed by comparison of the OR in Pimas with that published for Europeans. A multiallelic risk score, constructed by summing the number of risk alleles over all loci, was also analyzed. Stouffer's method for combining p-values was used to construct a summary test of heterogeneity across all loci. In analysis of the multiallelic score, the number of risk alleles was associated with higher risk of diabetes (OR=1.05 per risk allele, $p=8.3\times10^{-5}$); the multiallelic score was also strongly associated with reduced acute insulin secretion (r=-0.32, $p=3.1\times10^{-8}$). Nominally significant (p<0.05) associations with diabetes were observed for SNPs in *GCKR* (OR=1.38), *KCNQ1* (OR=1.32), *FTO* (OR=1.24), *CDKAL1* (OR=1.15) and *ZBED3* (OR=1.15). The summary test for heterogeneity across all 39 loci indicated that the effects in Pimas were weaker than in Europeans ($p=3.0\times10^{-4}$). Nominally significant heterogeneity between Pimas and Europeans was observed for SNPs in *IRS1*, *ADAMTS9*, *PTPRD*, *TCF7L2* and *DUSP9*. When allele frequencies were compared with the HapMap CEU population, differences with genome-wide significance ($p<5\times10^{-8}$) were seen for 27 SNPs; however at 14 of these the risk allele was at lower frequency in Pimas. These analyses suggest that in general type 2 diabetes risk loci identified in other populations also influence risk of T2DM in Pima Indians, and that risk is mediated in part through an effect on acute insulin secretion. However, the effects on T2DM risk are generally weaker in Pimas than in Europeans, and systematic differences in allele frequencies at these variants do not appear to account for differences in population risk of T2DM.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2039F

Describing the contribution of Multiple Sclerosis genetic risk factors in Sardinian Multiple Sclerosis cases. *A. Hadjixenofontos[1], L. Foco[2], V. Bakthavachalam[3], P-A. Gourraud[3], A. Ticca[5], P. Bitti[6], R. Pastorino[2], L. Bernardinelli[2,4], J.L. McCauley[1].* 1) John P. Hussman Institute for Human Genomics, University of Miami, Miami, FL; 2) Department of Public Health, Neurosciences, Experimental and Forensic Medicine, Medical Statistics and Epidemiology Unit, University of Pavia, 27100, Pavia, Italy; 3) Department of Neurology, School of Medicine, University of California at San Francisco, San Francisco, CA, USA; 4) Statistical Laboratory, Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WA, UK; 5) Divisione di Neurologia, Ospedale S. Francesco, Nuoro, Italy; 6) Immunohaematology and Blood Transfusion Department, Ospedale S. Francesco, Nuoro, Italy.

Multiple Sclerosis (MS) is an autoimmune, demyelinating disease that constitutes the leading cause of neurological disability in young adults. Prevalence rates follow a latitude gradient with the lowest prevalence in countries closest to the equator. Sardinia, in Insular Italy defies this gradient with a prevalence range of 150-200 per 100,000 people. The involvement of genetic risk factors has been established through twin, adoption and migration studies. Dissections of the genetic risk factor landscape have focused on populations with Northern European ancestry. The geographical isolation of Sardinia presents a unique opportunity to study a relatively homogeneous population; however, the underlying genetic risk factors for MS in Sardinia remain largely unknown. Clinical expression studies suggest overlapping risk factors across these populations, and population genetics coupled with Sardinian population history suggest that fewer variants act at the population level. To qualify the involvement of Northern European-identified variants we assess the effectiveness of the MS genetic burden (MSGB) calculated on the basis of established MS susceptibility variants to differentiate between Sardinian MS cases and controls. Using genotypes from a custom Illumina genotyping chip dubbed the “Immunochip” we built a logistic regression model with disease status as the outcome, and MSGB, and sex as predictors/covariates. An increase in the Northern European derived MSGB was not associated with MS case status in our preliminary analysis of cases from 16 Sardinian families, and 94 unrelated Sardinian controls (OR=1.136, 95% CI 0.857-1.493, P > 0.05). This suggests that perhaps a separate set of genetic risk factors, are driving MS in Sardinia. A more conclusive analysis is currently underway. This analysis will include additional cases from each family, using Generalized Estimating Equations to correct for correlations between relatives, and additional controls from the SardiNIA Medical Resequencing Project (NIH). Furthermore, we will assess the extent to which the MSGB can differentiate between Northern European and Sardinian MS cases.

2040W

Heritability and Linkage Analysis of the Health Aging Index (HAI) in the Long Life Family Study. *M.M. Barmada[1], J.L. Sanders[2], A.M. Matteini[3], K. Christensen[4], R. Mayeux[5], T. Perls[6], I. Borecki[7], Q. Zhang[7], A.B. Newman[2], Long Life Family Study.* 1) Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA; 2) Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA; 3) Division of Geriatric Medicine and Gerontology, School of Medicine, Johns Hopkins University, Baltimore, MD; 4) Department of Epidemiology, University of Southern Denmark, Odense, Denmark; 5) Department of Neurology, College of Physician and Surgeons, Columbia University, New York, NY; 6) Department of Medicine, School of Medicine, Boston University, Boston, MA; 7) Division of Statistical Genomics, School of Medicine, Washington University in St. Louis, St. Louis, MO.

A goal of gerontology research is to find longevity assurance genes that affect many specific conditions leading to age-related chronic disease and death, providing clues to an underlying aging process. We constructed a Healthy Aging Index (HAI) and determined its association with death and heritability to validate it as a longevity phenotype in the Long Life Family Study (LLFS). LLFS is a two-generation, family-based cohort designed to study sibships of exceptionally healthy older individuals and their adult offspring. LLFS comprises a total of 480 families with information measured on 3770 individuals; 1294 in the older generation (mean age = 89 yr), and 2476 in the offspring generation (mean age = 61 yr). The HAI includes systolic blood pressure, pulmonary vital capacity, serum creatinine, fasting glucose, and Modified Mini Mental Status Exam score, each scored as 0, 1 or 2 using approximate tertiles. Scores were summed from 0 (healthy) to 10 (unhealthy) using either equal weighting, or with mortality-optimized component weights. Using maximum likelihood methods, heritability of this trait was estimated in LLFS (overall age-adjusted h^2 = 0.285, P=8.2e-11; probands h^2 = 0.369, P=9e-7; offspring h^2 = 0.248, P=0.00078). We next performed variance components linkage analyses using multipoint identity by descent estimates derived from three sets of SNPs. Significant linkage (LOD > 3.0) was detected on chromosome 8 and suggestive linkage (LOD > 2.5) was observed on chromosome 19. These findings support further evaluation of the genetic and environmental factors contributing to this phenotype of healthy aging in the elderly.

2041T

Genetic determinants of Age-related Macular Degeneration in Diverse Populations: the Population Architecture using Genomics and Epidemiology (PAGE) Study. *N. Restrepo[1], K. Spencer[3], R. Goodloe[1], T. Garrett[4], G. Heiss[4], P. Buzkova[5], N. Jorgensen[5], R. Jensen[5], T. Matise[6], B. Klein[7], R. Klein[7], W. Tien Yin[8], B. Cornes[8], E. Shyong Tai[9], M. Ritchie[10], J. Haines[1], D. Crawford[1,2].* 1) Center for Human Genetics Research, Vanderbilt University, Nashville, TN; 2) Vanderbilt Department of Molecular Physiology & Biophysics; 3) Department of Biology, Heidelberg; 4) Department of Epidemiology, University of North Carolina, Chapel Hill; 5) Department of Biostatistics, University of Washington, Seattle; 6) Department of Genetics, Rutgers University, Piscataway; 7) Office of Population Genomics, National Human Genome Research Institute, national Institutes of Health, Bethesda, MD; 8) Department of Ophthalmology and Visual Sciences, University of Wisconsin, Madison, WI; 9) Singapore Eye Research Institute, Singapore national Eye Centre, Singapore; 10) PennState Department of Biochemistry and Molecular Biology.

Age-related Macular Degeneration (AMD) is the third leading cause of visual impairment worldwide. Unlike many other common, complex disorders, substantial progress has been made in identifying susceptibility variants for AMD. The most widely replicated loci are Complement Factor H (CFH) Y402H and Age-related Maculopathy Susceptibility-2 (ARMS2), both of which were discovered in populations of mainly European-descent. To date, little data exist for these AMD-associated variants in diverse populations. A major goal of the Population Architecture using Genomics and Epidemiology(PAGE) study is to describe the underlying genetic architecture of common, complex diseases such as AMD across diverse populations. The PAGE study consists of several epidemiologic collections including the National Health and Nutrition Examination Surveys (NHANES), the Atherosclerosis Risk in Communities Study (ARIC), and the Cardiovascular Health Study (CHS). Also included in this study are the Singapore Prospective Study Programme (SP2) and the Singapore Malay Eye Study (SiMES). Targeted genotyping was performed for AMD-associated SNPs in CFH, CFI, ARMS2, VEGF, and lipid trait loci. All cases of AMD were >60 years of age and presented with early or late AMD as determined by fundus photography. A total of 830, 95, 107, and 21 cases and 5,710, 1,172, 863, and 206 controls from European Americans, African Americans, Malays, and Chinese, respectively, were available for study. We performed a meta-analysis following logistic regression assuming an additive genetic model performed at each study site and adjusted for age, sex, body mass index, smoking status, and HDL cholesterol. In European Americans, rs1061170 (CFH) and rs10490924 (ARMS2) replicated at p=3.7x10-10 (OR= 1.61, 95% CI= 1.85and 1.41) and p=1.8x10-5 (OR = 1.49, 95% CI= 1.78 and 1.25), respectively. None of the CFH or ARMS2 SNPs tested generalized to African Americans (p>0.05). Interestingly, HDL associated variant CETP rs1800775 (OR=0.55; 95% CI=0.86 and 0.35; p=0.009) associated with AMD in African Americans but not European Americans or Asians suggesting a potential risk modifier in lipid pathways to AMD in this population. Further studies are needed to determine if lack of generalization in major CFH and ARMS2 variants is due to statistical power or differences in linkage disequilibrium and allelic distribution across these diverse populations.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2042F

Genetic variation and age at natural menarche and menopause in African American women from the Population Architecture using Genomics and Epidemiology (PAGE) Study. *J. Malinowski[1,14], K.L. Spencer[12,14], C.L. Carty[5], N. Franceschini[4], L. Fernández-Rhodes[4], A. Young[5], I. Cheng[7], M.D. Ritchie[1,13], C.A. Haiman[6], L. Wilkens[7], C. Wu[5], T.C. Matise[2], C.S. Carlson[5], K. Brennan[8], A. Park[9], A. Rajkovic[10], L.A. Hindorff[11], S. Buyske[2,3,14], D.C. Crawford[1,14].* 1) Center for Human Genetics Research, Vanderbilt University, Nashville, TN; 2) Department of Genetics, Rutgers University, Piscataway, NJ; 3) Department of Statistics, Rutgers University, Piscataway, NJ; 4) Department of Epidemiology, University of North Carolina, Chapel Hill, NC; 5) Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA; 6) Keck School of Medicine, University of Southern California, Los Angeles, CA; 7) University of Hawaii Cancer Center, University of Hawaii, Honolulu, HI; 8) Department of Obstetrics and Gynecology, University of California Los Angeles, Los Angeles, CA; 9) Department of Obstetrics and Gynecology, School of Medicine, Georgetown University, Washington D.C; 10) Magee-Womens Research Institute, Department of Obstetrics, Gynecology and Reproductive Sciences, University of Pittsburgh, Pittsburgh, PA; 11) Office of Population Genomics, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD; 12) Department of Biology and Environmental Science, Heidelberg University, Tiffin, OH; 13) Biochemistry and Molecular Biology, Penn State University, University Park, MD; 14) The authors contributed equally to the work.

Age at menarche (AM) and age at natural menopause (ANM) define the boundaries of the reproductive lifespan in women. The timing of these events has been associated with various diseases, including cancer and cardiovascular disease. Genome-wide association studies have identified several genetic variants associated with either AM or ANM in populations of largely European or Asian descent women. The extent to which these associations generalize to diverse populations remains unknown. Therefore, we sought to replicate previously reported AM and ANM findings and also performed a large scale association study to identify novel AM and ANM variants using Metabochip data (161,098 SNPs) genotyped in 4,159 and 1,860 African American women, respectively, in the Women's Health Initiative (WHI) and Atherosclerosis Risk in Communities (ARIC) studies, as part of the Population Architecture using Genomics and Epidemiology (PAGE) Study. We were able to generalize only one previously identified variant each for AM and ANM to our African American cohort, rs1361108/*CENPW* and rs897798/*BRSK1*, respectively. Overall, generalization of the majority of previously-identified variants for AM and ANM, including *LIN28B* and *MCM8*, was not observed in this African American sample, though these loci were not well covered in the array. In our discovery analysis, we identified three loci associated with ANM that reached significance after accounting for multiple testing (*LDLR* rs189596789, $p=5x10^{-08}$; *KCNQ1* rs79972789, $p=1.9x10^{-07}$; *COL4A3BP* rs181686584, $p=2.9x10^{-07}$) as well as ninety-two additional SNPs suggestively associated with ANM, $p<1x10^{-4}$. Our most significant AM association was upstream of *RSF1*, a gene implicated in ovarian and breast cancers (rs11604207, $p=1.6x10^{-06}$). We identified suggestive AM and ANM associations in or near genes implicated in obesity, type 2 diabetes, cancer, or cardiovascular traits (*HEXIM2* rs59622946, *LPL* rs4922116, *PPP2R3A* rs11922097, *CYP4F22* rs1273516), which, in part, likely reflects the Metabochip focus on cardiometabolic loci. While most associations were unique to either AM or ANM, we also identified novel associations in genes common to both: *PHACTR1* (rs117124693(ANM) and rs73725617(AM)) and *ARHGAP42* (rs11224401(ANM) and rs11224447(AM)). The lack of generalization coupled with the potentially novel associations identified here emphasize the need for additional genetic discovery efforts for AM and ANM in diverse populations.

2043W

African Ancestry Is a Risk Factor for Asthma and High Total IgE Levels in African Admixed Populations. *C.I. Vergara[1], T. Murray[2], N. Rafaels[1], R. Lewis[1], M. Campbell[1], C. Foster[1], L. Gao[1], M. Faruque[3], R. Riccio Oliveira[4], E. Carvalho[4], M.I. Araujo[4], A.A. Cruz[5], H. Watson[6], D. Mercado[7], J. Knight-Madden[8], I. Ruczinski[9], G. Dunston[3], J. Ford[2], L. Caraballo[7], T. Beaty[2], R.A. Mathias[1], K.C. Barnes[1].* 1) Division of Allergy and Clinical Immunology, Johns Hopkins University, Baltimore, MD; 2) Department of Epidemiology, Bloomberg School of Public Health, JHU, Baltimore,MD 21225; 3) National Genome Center at Howard University, Washington, D.C; 4) Servico de Imunologia, Hospital Universitario Professor Edgard Santos, Salvador, Bahia, Brazil; 5) ProAR-Nucleo de Excelencia em Asma, Federal University of Bahia and CNPq, Salvador, Bahia, Brazil; 6) Faculty of Medicine, Univesity of the West Indies, Cave Hill Campus, Barbados; 7) Institute for Immunological Research, University of Cartagena, Cartagena, Colombia; 8) Tropical Medicine Research Institute, The University of West Indies, Mona Kingston 7, Jamaica, West Indies; 9) Johns Hopkins School of Public Health, Biostatistics, Baltimore, Maryland.

Characterization of genetic admixture of populations in the Americas and the Caribbean is of interest for anthropological, epidemiological and historical reasons. Asthma has a higher prevalence and is more severe in populations with a high African component. Association of African ancestry with asthma has been demonstrated. We estimated admixture proportions of samples from 6 trihybrid populations of African descent and determined the relationship between African ancestry and asthma and total serum IgE levels (tIgE). We genotyped 237 ancestry informative markers in asthmatics and non-asthmatic controls from Barbados (190/277), Jamaica (177/529), Brazil (40/220), Colombia (823/941), African Americans from New York (207/171) and African Americans from Baltimore/Washington, D.C. (625/757). We estimated individual ancestries and evaluated genetic stratification using STRUCTURE and principal component analysis. Association of African ancestry and asthma and tIgE was evaluated by regression analysis. Mean ± SD African ancestry ranged from 0.76 ± 0.10 among Barbadians to 0.34 ± 0.14 in Colombians. The European component varied from 0.14 ± 0.05 among Jamaicans and Barbadians to 0.25 ± 0.07 among Colombians. African ancestry was associated with risk for asthma in Colombians (OR: 4.3; 95% CI: 2.1-8.6; P = 0.0001)), Brazilians (OR= 136.5; 95% CI: 5.4-3406; P = 0.003) and African Americans of New York (OR: 4.7; 95% CI: 1.07-21.04, P=0.040). African ancestry was also associated with higher tIgE levels among Colombians (β=1.3, 95% CI: 1.1-2.6, P = 0.003), Barbadians (β= 3.8, 95% CI: 1.1-13.1; P = 0.03) and Brazilians (β = 1.6, 95% CI: 1.1-2.6; P = 0.03). Our findings indicate that African ancestry can account for, at least in part, the association between asthma and its associated trait, tIgE levels.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2044T

Validation of the HELIC population isolate collections as cohorts for complex trait association mapping. *E. Zeggini[1], I. Tachmazidou[1], A. Farmaki[2], L. Southam[1,3], N.W. Rayner[1,3], K. Daoutidou[2], A. Kolb-Kokocinski[1], K. Panoutsopoulou[1], E. Tsafantakis[4], M. Karaleftheri[5], G. Dedoussis[2].* 1) Wellcome Trust Sanger Institute, Hinxton, United Kingdom; 2) Harokopio University, Athens, Greece; 3) Wellcome Trust Centre for Human Genetics, Oxford, UK; 4) Anogia Medical Centre, Greece; 5) Echinos Medical Centre, Greece.

Over the last five years, genome-wide association studies (GWAS) have been successful in identifying multiple robustly replicating complex trait genetic loci. The field is now shifting focus from common to low frequency and rare sequence variation. Population isolates can enhance power to detect association at these sites, as rare variants may have risen in frequency and linkage disequilibrium tends to be extended. We have collected samples from two isolated populations in Greece under the auspices of the HELIC (HELlenic Isolated Cohorts) study (HELIC-Pomak and MANOLIS cohorts), both accompanied by information on a wide array of anthropometric, cardiometabolic, biochemical, haematological and diet-related traits. We recently established these cohorts as genetic isolates based on measures of extended haplotype sharing derived from GWAS data and have whole-genome sequenced 250 individuals from MANOLIS at 6x to enable imputation and subsequent association testing. To date, we have carried out a GWAS for glycaemic traits and compared results at established trait loci (Dupuis et al, Nat Genet, 2010) to validate the utility of these cohorts in complex trait mapping. 747 and 1118 individuals were typed on the Illumina OmniExpress platform for HELIC-Pomak and MANOLIS respectively. Fasting glucose (total n=1566) and fasting insulin (total n=1759) trait values were inverse-normal transformed and residualized using age, age^2 and the kinship matrix as covariates. A fixed-effects meta-analysis across ~665,000 SNPs was performed across the two cohorts. Of the 18 established glucose-associated variants, 14 were present in our analysis and all had the same direction of effect as the published data (binomial p=6.1x10^{-5}). The strongest association was observed for the *MTNR1B* SNP rs10830963 (p=8.7x10^{-5}), followed by rs560887 at the *G6PC2* locus (p=1.1x10^{-4}). Of the 2 established insulin loci, one (rs780094 at the *GCKR* locus) went in the same direction and showed evidence for association with the trait (p=4.2x10^{-3}). These data provide validation of the HELIC-Pomak and MANOLIS cohorts for use in complex trait association mapping. Phasing of the 250 sequenced individuals and imputation into the full analysis cohorts will enable assessment of low frequency and rare variant associations with quantitative traits of cardiometabolic relevance.

2045F

Evidence of Natural Selection at Crohn's GWAS Loci Based on Population Genetic Variation. *B.M.P.M. Bowen, J.H. Cho.* Department of Genetics, Yale University, New Haven, CT.

Crohn's Disease (CD) is an inflammatory bowel disease (IBD) affecting 150 per 100,000 Americans on average; however, the prevalence of CD among American subpopulations varies with ancestry. In particular, the prevalence of CD is four to seven-fold higher in Ashkenazi Jews (AJs) than in non-Jewish individuals of European descent (NJs). Using available genotype data from a previous genome-wide association study (GWAS), as well as data from the recently completed Immunochip IBD association study (IC), we found an enrichment of derived risk alleles in AJs and an enrichment of derived protective alleles in NJs. Further, for shared risk alleles, the risk allele frequency was higher in AJs. For shared protective alleles, the protective allele frequency was lower in AJs. We examined whether the allele frequency differences between AJs and NJs could be explained in part by natural selection. We examined population genetic patterns in AJ and NJ CD cases. AJ and NJ individuals were identified by self-report and ancestry was confirmed by principal components analysis. Admixed individuals were omitted from subsequent analyses. Each individual's GWAS data was phased using BEAGLE and was used to calculate extended haplotype homozygosity (EHH). We identified a number of genes for which cores of EHH differed between the two populations. We also observed more extensive EHH in AJs. We noted an enrichment of genes from the tumor necrosis factor super family (TNFSF) and its receptors in cores of EHH in AJs compared to NJs including *TNFSF8*, *TNFRSF8*, and *TNFRSF18*, among others. We then looked for evidence of coevolution for receptor-ligand pairs based on the correlation of receptor and ligand allele frequencies using a merged dataset of the Human Genome Diversity Panel and HapMap populations as well as our own sample populations, totaling a sample of 65 populations covering 7 geographic regions around the world. We found an inverse correlation between the frequency of the CD-associated *TNFSF18* and *TNFRSF18* alleles, as well as the *TNFRSF8* allele and the CD-associated *TNFSF8* allele, with the former driven largely by a selective sweep at *TNFSF18* in East Asians. These data support that natural selection has acted on CD loci in multiple populations and that such events may explain population-specific differences in disease prevalence.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2046W

Genetic Associations with Serologic Autoimmunity in a Large Multi-Ancestral Systemic Lupus Erythematosus Cohort. *S.N. Kariuki[1], B.S. Franek[1], A.A. Kumar[1], M. Kumabe[1], K.M. Kaufman[2], J.M. Anaya[3], M.E. Alarcón-Riquelme[2,4], S.C. Bae[5], E.E. Brown[6], B.I. Freedman[7], G.S. Gilkeson[8], C.O. Jacob[9], J.A. James[2], R.P. Kimberly[6], J. Martin[10], J.T. Merrill[2], B. Pons-Estel[11], B.P. Tsao[12], T.J. Vyse[13], C.D. Langefeld[7], J.B. Harley[14], K.L. Moser[2], P.M. Gaffney[2], A.D. Skol[1], T.B. Niewold[1].* 1) Gwen Knapp Ctr Lupus & Immunology Res, University of Chicago, Chicago, IL; 2) Arthritis and Clinical Immunology Research Program, Oklahoma Medical Research Foundation, Oklahoma City, OK; 3) Center for Autoimmune Diseases, Universidad del Rosario-Corporacion para Investigaciones Biologicas, Bogota, Colombia; 4) Centro de Genómica e Investigación Oncológica, Pfizer-Universidad de Granada-Junta de Andalucía (GENYO), Granada, Spain; 5) Clinical Research Center for Rheumatoid Arthritis, Hanyang University Hospital for Rheumatic Diseases, Seoul, South Korea; 6) Department of Medicine, University of Alabama at Birmingham, Birmingham, AL; 7) Wake Forest School of Medicine, Winston-Salem, NC; 8) Division of Rheumatology and Immunology, Medical University of South Carolina, Charleston, SC; 9) Department of Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA; 10) Instituto de Parasitologia y Biomedicina Lopez-Neyra (CSIC), Granada, Spain; 11) Sanatorio Parque, Rosario, Argentina; 12) Division of Rheumatology, Department of Medicine, David Geffen School of Medicine University of California Los Angeles, Los Angeles, CA; 13) Divisions of Genetics and Molecular Medicine and Immunology, King's College London, London, United Kingdom; 14) Cincinnati Children's Hospital Medical Center and the US Department of Veterans Affairs Medical Center, Cincinnati, OH.

Systemic lupus erythematosus (SLE) is a highly heterogeneous disorder characterized by production of autoantibodies directed at particular self-antigens in the cell nucleus. These autoantibodies are an important molecular sub-phenotype of SLE and are directly pathogenic. Prevalence of these autoantibodies varies by ancestral background, and genetic variation may play a role in these differences. We followed up a previous case-only GWAS which identified loci that were also associated with particular autoantibodies in SLE patients. We genotyped 43 tag SNPs from 10 regions identified in the above GWAS in a multi-ancestral SLE cohort of over 8000 patients split between African-American, European, Hispanic, and Asian ancestral backgrounds. The MAPK4 gene was associated with autoantibodies across multiple ethnic groups: rs2051308 was strongly associated with anti-dsDNA in African-Americans (OR=0.784, $p=5.49x10^{-4}$), rs7242442 was significantly associated with anti-RNP in Hispanic-Americans (OR= 1.578, $p= 3.18x10^{-4}$) while rs8097986 was strongly associated with anti-Ro (OR= 0.8248, $p= 4.14x10^{-3}$) in European-Americans. We also detected associations which were specific to one ancestral background: in African-Americans, rs310230 SNP in JAK1 was significantly associated with both anti-dsDNA and anti-RNP antibodies (OR=1.63, $p=6.20x10^{-05}$); in Asians, rs12202764 in the PRL gene was significantly associated with anti-La (OR= 4.681, $p= 1.53x10^{-4}$). We demonstrate one locus which is associated with diverse autoantibody specificities across multiple ancestral backgrounds, and other loci that are uniquely associated with a particular autoantibody profile in a specific ancestral background. The various SNP-autoantibody associations identified in this study highlight important molecular pathways that predispose to the characteristic serologic autoimmunity which contributes to the pathogenesis of SLE.

2047T

HLA DRB1 amino acid position 11 is highly associated with rheumatoid arthritis in African-Americans. *R.J. Reynolds[1], S. Raychaudhuri[2], P.I.W. de Bakker[2,3], X. Jia[2], M.I. Danila[1], A.F. Ahmed[4], L.W. Moreland[5], R. Brasington[6], L.F. Callahan[7], E.A. Smith[8], R.M. Plenge[2], S.L. Bridges Jr.[1], CLEAR Investigators.* 1) Department of Medicine, Division of Clinical Immunology and Rheumatology University of Alabama at Birmingham, Birmingham, AL; 2) Division of Genetics, Brigham & Women's Hospital, Harvard Medical School, Boston, MA; 3) Department of Medical Genetics, University Medical Center Utrecht, Utrecht, The Netherlands; 4) University of Alabama School of Medicine, Birmingham, AL; 5) Department of Medicine, Division of Rheumatology and Clinical Immunology, University of Pittsburgh, Pittsburgh, PA; 6) Washington University School of Medicine, St. Louis, MO; 7) University of North Carolina, Chapel Hill, NC; 8) Department of Medicine, Division of Rheumatology and Immunology, Medical University of South Carolina, Charleston, SC.

Although more than 30 non-major histocompatibility complex risk loci for rheumatoid arthritis (RA) have been identified in populations of European descent, of the known variants, 75% of the genetic variance for RA is confined to three human leukocyte antigen (HLA) loci. Compared to peoples of European descent, African-Americans have a lower prevalence of RA, the highest risk HLA DRB1 classical alleles (e.g. *04:01, *04:04) are found at lower frequency, and the effect size of the high risk alleles with RA is lower. Here we test the hypothesis that the source of the risk among HLA DRB1 alleles with RA in African-Americans is amino acid positions 11, 71, and 74, as shown in populations of European descent. Using Allele SEQR HLA-DRB1 (Abbot), we sequenced 665 African American autoantibody-positive RA cases and 976 healthy controls and assigned four digit alleles using Assign SBT v 3.5 (Conexio Genomics). Alleles were converted to binary variables indicating the presence/ absence of a specific amino acid residue at a given position. A series of linear models were fit to assess whether the joint effect of amino acid residues within each position was a significant source of variation for the presence of RA. Of 30 amino acid positions examined, amino acid position 11 (P=7.8E−23) was most highly associated with RA followed by position 13 (P=2.7E−19), 71 (P = 2.2E−04), 70 (1.9E−03) and 74 (P= 3.5E−03). After conditioning on position 11, the effects at the other positions were diminished, suggesting that in this sample amino acid position 11 was the primary contributor to the association with HLA DRB1. Within position 11 the multivariate odds ratios (+, − 95% CI) of association were: Val 3.3 (2.4, 4.5); Asp 2.4 (1.6, 3.5); Leu 1.5 (1.1, 2.1); Gly 1.4 (0.98, 1.9); Ser 1.0 (0.80, 1.3), with Pro as the referent. Two of the three high risk residues, Val and Leu, were also reported to be high risk residues in populations of European descent (Val 3.8, Leu 1.3). The odds ratio for Asp (2.4) was substantially higher for African-Americans than for Europeans (Asp 1.0). The finding that amino acid position 11 is most strongly associated with RA in African-Americans validates the association for this position described in populations of European descent. However, substantial differences are noted among amino acid residues within position 11, which can partly be attributed to differences in the frequencies of the presence of the residues between the ethnic groups.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

† Trainee Award Finalist

2048F

***GCKR* Polymorphism and Insulin Resistance in Americans of Mexican origin.** *Y. Lu[1], H.Q. Qu[1], Q. Li[2], S.P. Fisher-Hoch[1], J.B. McCormick[1].* 1) The University of Texas School of Public Health, Brownsville, TX; 2) The McGill University Health Center (Montreal Children's Hospital), Montreal, QC, Canada.

Our previous study of the community based Cameron County Hispanic Cohort (CCHC) consisting of Mexican Americans showed that elevated homeostasis model assessment-estimated insulin resistance (HOMA-IR) is significantly more frequent in our adult and adolescent Mexican American population than in European populations(Qu et al., 2011, Rentfro et al., 2011). Using a set of ancestral informative markers (AIM) comprised of 103 single nucleotide polymorphisms (SNP) we then conducted a genetic ancestral analysis in which we demonstrated that elevated HOMA-IR levels are highly correlated with Amerindian ancestry, highlighting ethnic effects in this admixed Mexican American population(Qu et al., 2012). We are now investigating possible molecular mechanisms associated with this ethnic effect. An interesting candidate genetic mechanism is suggested by recent genome-wide association studies (GWAS), i.e. a SNP rs780094 of the glucokinase regulatory protein gene (*GCKR*) that is significantly associated with HOMA-IR (Dupuis et al., 2010). GCKR is an inhibitor of glucokinase (GCK) in liver and pancreatic islets, while GCK catalyzes the first step of glycogen synthesis and glycolysis, and serves as a metabolic glucose sensor in liver and pancreatic islets and thus plays a key role in glucose metabolism. Significant differences in rs780094 frequencies occur among different human populations according to the HapMap phase 3 study (http://hapmap.ncbi.nlm.nih.gov/), which therefore suggested to us as one possible mechanism mediating our reported ancestral effect on HOMA-IR. To test this hypothesis we genotyped rs780094 in 1551 CCHC participants who were also genotyped for the 103 AIMs. We found that, rs780094 is highly significantly associated with HOMA-IR levels (β=0.088, P=6.43E−05) after correction for population structure, body mass index (BMI), age, and gender, and also associated with Amerindian ancestry (β=0.081, P=1.48E−03). This DNA marker alone explains 6.14% of the ancestral effect on HOMA-IR levels in Mexican Americans. We are currently pursuing this approach to systematically identify other potential molecular mechanisms involved in the ancestral effect of HOMA-IR in people of Mexican origin.

2049W

Characterization of clinical characteristics and genetic burden of Multiple Sclerosis in Hispanics. *I. Konidari[1], A. Hadjixenofontos[1], C.P. Manrique[1], A.H. Beecham[1], P.L. Whitehead[1], V. Bakthavachalam[2], P-A. Gourraud[2], M.A. Pericak-Vance[1], L. Tornes[3], M. Ortega[3], K.W. Rammohan[3], S. Delgado[3], J.L. McCauley[1].* 1) John P. Human Genomics, University of Miami - Miller School of Medicine, Miami, FL; 2) Department of Neurology, School of Medicine, University of California at San Francisco, San Francisco, CA, USA; 3) Multiple Sclerosis Division, Department of Neurology, Miller School of Medicine, University of Miami, Miami, FL, USA.

Multiple Sclerosis (MS) is a demyelinating disease with autoimmune etiology and variable expression across populations. The influence of race and ethnicity is suggested by the differences in disease presentation among African Americans, Asians, Hispanics and non-Hispanic Whites (NHW). Hispanic populations were considered to be less susceptible to MS, however, recent studies suggest that the incidence and prevalence of MS in this admixed population is increasing. Studies of Hispanics in the US with primarily Mexican origin note a preponderance of opticospinal MS at first presentation, with a younger age at onset, and a progression of disability that is similar to that of NHW. The primary purpose of our study was to describe the expression of MS in a sample of Hispanics from the South Florida community with primarily Caribbean origin. Chi-squared tests were used to identify statistically significant differences in sex, disease type, symptoms at initial presentation, site of initial presentation, Expanded Disability Status Scale score, and interferon response. Age at onset, age at diagnosis, diagnostic lag, disease duration, MS genetic burden score (MSGB) and genetic burden due to HLA-DRB1*1501 were tested for differences by race/ethnicity using t-tests. Since the MSGB is designed to capture the genetic load attributed to variants identified in populations of European ancestry, it is expected that an admixed population will be characterized by a partly-overlapping set of genetic variants and/or environmental influences. In our studies, we observe a significantly younger age at onset (P=0.016) and age at diagnosis (P=0.016) in Hispanics (n=251) when compared to NHW (n=247). Interestingly, we also observe a significantly lower MSGB (P<0.001) in Hispanics as compared to NHW. When we excluded HLA-DRB1*1501 from the analysis, this difference was no longer significant (P=0.126). Our study highlights the need for further studies of MS risk factors among Hispanics with the goal of understanding the underlying causes of MS in this population in hope of improved disease management and treatment of MS in all populations.

2050T

Investigating a role for functional variation in HIV-1 control through exome sequencing. *P.J. McLaren[1, 2], P. Shea[3], I. Bartha[1, 2], K.V. Shianna[3], D.B. Goldstein[3], J. Fellay[1, 2].* 1) School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne, Lausanne, VD, Switzerland; 2) Institute of Microbiology, University Hospital Lausanne, Lausanne, VD, Switzerland; 3) Center for Human Genome Variation, Duke University School of Medicine, Durham, NC, USA.

Background: Genome-wide association studies (GWAS) of HIV-1 infected individuals have identified variants in the MHC and CCR5 regions as associated with viral load and disease progression, but none elsewhere. However, these studies are limited to analyzing common genetic variation represented on commercial arrays. Here we investigate the role of rare and functional variants in mediating HIV-1 control through exome sequencing.

Methods: We captured and sequenced all coding exons to high coverage (> 70x) in 131 HIV-1 infected individuals of recent Western European ancestry using the SureSelect Human All Exon 50Mb enrichment kit and the Illumina HiSeq2000. Paired-end, 120 bp reads were aligned using the Burrows-Wheeler Aligner (BWA) and quality control and variant calling were performed using a combination of Picard and Samtools. Variant functional annotation was performed using snpEff version 2.1. As a correlate of HIV-1 control, we used viral load at set point (spVL), calculated as the average of at least 3 longitudinal measurements obtained during the chronic phase of infection, in absence of antiretroviral therapy.

Results: An average of 10,109 non-synonymous variant alleles were observed per individual. No association was observed between mean virus load and genome-wide burden of non-synonymous variants after accounting for the top eigenstrat coordinates (from GWAS), age and sex. In total, 358 genes were found to carry homozygous stop gain mutations in at least 1 individual. Of these genes, 22 are reported as having a direct interaction with at least one viral protein in the HIV-1, Human Protein Interaction Database (http://www.ncbi.nlm.nih.gov/RefSeq/HIVInteractions). No association was observed between presence of a homozygous stop-gain mutation in a known HIV-1 interacting gene and spVL.

Conclusions: Preliminary analysis shows no association between spVL and genome-wide burden of missense variants or carriage of a homozygous stop mutation in a gene known to directly interact with HIV-1. Addition of further samples and other testing methods (e.g. individual variant testing, collapsing of variants by gene and pathways) will be required to fully address the role of rare, functional sequence variants in controlling HIV-1 infection.

2051F

The effect of genetic and interaction factors for risk prediction in type 2 diabetes. *D. Shigemizu[1], T. Abe[1], T. Morizono[1], TA. Johnson[1], KA. Boroevich[1], M. Kubo[1], Y. Nakamura[2], S. Maeda[1], T. Tsunoda[1].* 1) Center for Genomic Medicine, RIKEN, Yokohama, Kanagawa, Japan; 2) Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan.

Many genome-wide association studies (GWAS) have recently identified and validated novel single nucleotide polymorphisms (SNPs) associated with type 2 diabetes (T2D). Numerous models and scores using genetic risk factors have been developed for T2D risk prediction, such as a prognostic model, because multiple genetic and clinical risk factors are expected to contribute the pathogenesis of T2D. However, risk prediction models composed of genetic factors have shown only a low ability to predict the development of T2D compared to models composed clinical risk factors. To our knowledge, exhaustive analysis considering not only algorithms for risk factor detection but also regression methods for model construction and, in furthermore, interactions of the risk factors have not been investigated so far. Here, we exhaustively examined risk prediction models considering three algorithms for risk factor detection, three regression methods for the model construction and interactions of risk factors. The best risk prediction model was constructed using a Bayes factor approach, considering information of the square of the correlation coefficient between SNPs and the lasso method. This method achieved a highest area of a receiver operating characteristic curve (AUC), an improvement over analysis composed of only clinical risk factors. With the addition of two significant interaction factors, the model was further improved with an increase in AUC. Our results suggest that current risk prediction models can be improved by taking into consideration genetic risk factors and interaction factors.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2052W

Genome-Wide Association Study of Structural Foot Disorders (lesser-toe deformity, hallux valgus and plantar soft-tissue atrophy) in Older Caucasian Populations. *YH. Hsu[1, 2], CH. Cheng[1], Y. Liu[3], LA. Cupples[4], J. Jordan[3], M. Hannan[1].* 1) HSL Institute for Aging Res and Harvard Medical School, Boston, MA; 2) Molecular and Integrative Physiological Sciences, Harvard School of Public Health, Boston, MA; 3) University of North Carolina, Chapel Hill, NC; 4) Boston Univ. Sch Pub Hlth, Boston, MA.

Structural foot disorders, such as lesser-toe deformity, hallux valgus and plantar soft-tissue atrophy, commonly affect ~ 60% of older adults and are associated with chronic mobility limitations, disability and foot pain. Although, ageing is considered the most important cause of these foot disorders, the importance of genetics is commonly suspected in clinical observations of family aggregation. Previously, we reported strong heritability (h^2) for lesser-toe deformity (61% in men; 85% in women) and moderate h^2 for hallux valgus (~35%) and plantar soft-tissue atrophy (~20%) in older men and women, suggesting potential genetic predisposition to structural foot disorders. To identify their genetic determinants, we have undertaken a GWAS using 2.5M imputed SNPs (HapMapII CEU reference panel) to localize susceptible genes in the Framingham Study. Structural foot disorders were indicated as present or absent and were assessed based on an atlas of pictorial depictions. Plantar soft tissue atrophy was determined by standardized palpatory clinical examination. Among 2,446 Framingham participants (mean age 66 yrs; 57% women; Caucasian), we identified 753 (31%), 764 (31%) and 665 (27%) participants with lesser-toe deformity, hallux valgus and plantar soft-tissue atrophy, respectively. A mixed-effect regression model was performed and adjusted for age, sex, weight and principal components of ancestral genetic background. A kinship covariance matrix was used to take into account within family correlations among siblings. We filtered out SNPs with low imputation quality (O/E variance ratio of allele frequency < 0.3) and SNPs with MAF < 1%. In addition, p-values were also adjusted for λGC. We found several associations achieved genome-wide significance ($p < 5x10^{-8}$), i.e. SNPs on TBC1D22A and OR5D13 gene for lesser-toe deformity. For hallux valgus, the most significant SNP ($p= 4.9x10^{-7}$) is located in GATAD2B gene. For plantar soft-tissue atrophy, the most significant SNP ($p=4.76x10^{-7}$) is located near ADAMTS16 gene. Pathway and gene-set analyses for the genome-wide significant and suggestive genes suggested significant clustering of genes involved in connective tissue disorders (such as oligoarticular arthritis, osteoarthritis and osteosclerosis). Of note, few SNPs reported to associate with longevity. The replication in independent samples is undergoing. In conclusion, our results reveal novel genes to further elucidate the etiology of structural foot disorders.

2053T

Multi-stage genome-wide association meta-analyses identified two new loci for bone mineral density. *L. Zhang[1,2], J. Li[2], YF. Pei[1,2], Y. Lin[1], H. Shen[2], K. Estrada[3,4,5], F. Rivadeneira[3,4,5], AG. Uitterlinden[3,4,5], CS. Shin[6], HJ. Choi[6], EL. Duncan[7,8], PJ. Leo[7], MA. Brown[7], YZ. Liu[2], YJ. Liu[2], JG. Zhang[2], Q. Tian[2], YP. Wang[2], XZ. Zhu[1], SY. Wu[1], CJ. Papasian[9], HW. Deng[1,2].* 1) Center of System Biomedical Sciences, University of Shanghai for Science and Technology, Shanghai, China; 2) Center for Bioinformatics and Genomics, Department of Biostatistics and bioinformatics, Tulane university, New orleans, LA; 3) Department of Internal Medicine, Erasmus Medical Center, Rotterdam, The Netherlands; 4) Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands; 5) Netherlands Genomics Initiative (NGI)-sponsored Netherlands Consortium for Healthy Aging (NCHA), Leiden, The Netherlands; 6) Department of Internal Medicine, College of Medicine, Seoul National University, Seoul, Korea; 7) Human Genetics Group, University of Queensland Diamantina Institute, Brisbane, Queensland, Australia; 8) Department of Endocrinology, Royal Brisbane and Women's Hospital, Brisbane, Queensland, Australia; 9) Department of Basic Medical Science, University of Missouri-Kansas City, Kansas City, USA.

Aim: Bone mineral density (BMD) is a highly heritable trait and a major risk factor for osteoporosis. We aim to identify genetic variants associated with BMD. Methods: We conducted a multi-stage genome-wide association (GWA) meta-analysis in 26,988 individuals. In Stage 1, we meta-analyzed 7 GWA samples and 11,140 individuals for BMDs at the lumbar spine, hip, and femoral neck. In Stage 2 we analyzed in silico a selected set of single nucleotide polymorphisms (SNPs) from Stage 1 in 3 independent samples encompassing 9,185 individuals. Novel SNPs at or near the genome-wide significance (GWS, $5.0x10^{-8}$) level in joint analyses of Stages 1+2 were replicated de novo in 2 additional samples encompassing 6,663 individuals. Results: Combining evidence from all stages, we have identified 2 novel loci at the GWS level: 14q24.2 (rs227425, p-value $3.98x10^{-13}$, SMOC1) and 21q22.13 (rs170183, p-value 4.15x10−9, CLDN14) that were associated with BMD in the combined sample of males and females and in the female specific sample respectively. We have also replicated 12 previously reported loci at the GWS level: 1p36.12 (ZBTB40), 1p31.3 (GPR177), 4p16.3 (FGFRL1), 5q14.3 (MEF2C), 6q25.1 (C6orf97, ESR1), 7q21.3 (FLJ42280, SHFM1), 7q31.31 (FAM3C, WNT16), 8q24.12 (TNFRSF11B), 11p15.3 (SOX6), 11q13.4 (LRP5), 13q14.11 (AKAP11), and 16q24 (FOXL1). Gene expression analysis in osteogenic cells implied some potential functions of the newly identified genes in bone metabolism. Conclusion: In total, these findings contribute to uncover the biological pathways underlying bone metabolism, and thus provide valuable insights into the intervention and treatment of osteoporosis.

2054F

Genetics of left- and right-hemisphere subcortical structures in the human brain. *M.E. Renteria[1,2], A. Wallace[1], L. Strike[1], K. Johnson[1], D.P. Hibar[3], J.L. Stein[3], G. De Zubicaray[2], K.L. McMahon[4], G.W. Montgomery[1], P.M. Thompson[3], N.G. Martin[1], S.E. Medland[1], M.J. Wright[1].* 1) Queensland Inst Medical Research, Herston, Queensland, Australia; 2) Laboratory of Neuro Imaging, UCLA School of Medicine, Los Angeles, CA, USA; 3) School of Psychology, University of Queensland, Brisbane, QLD, Australia; 4) Center for Advanced Imaging, University of Queensland, Brisbane, QLD, Australia.

The human brain displays great individual variation in its morphology and function. Subcortical structures such as the hippocampus, caudate nucleus, amygdala, putamen or pallidum play important roles in both basic and higher-order cognitive functions and behaviour. Previous studies have reported that individual differences in subcortical volumes are highly heritable and that significant phenotypic correlation exists between them. In this study, we obtained whole-brain MRI scans and extracted in a sample of 1032 twins and their siblings (655 females and 377 males). MRI images were processed with the FSL/FIRST software package to calculate volumes of 7 subcortical structures. We performed univariate and multivariate twin modelling to study the genetic contributions to variation and covariation across structures. Furthermore, DNA was extracted from blood samples, followed by SNP genotyping with the Illumina 610K Quad Bead chip and imputation to the latest reference panel of the 1000 Genomes Project was carried out using Mach. Finally, we performed 14 genome-wide association studies in our sample (7 left-hemisphere structures and 7 right-hemisphere structures).

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2055W

Variant in the HLA-DPB1 region inversely associated with follicular lymphoma risk in meta-analysis of two genome-wide association studies. *L. Conde[1], K.E. Smedby[2], J.N. Foo[3], J. Riby[1], K. Humphreys[4], F.C.M. Sillé[1], H. Darabi[4], S. Sanchez[1], H. Hjalgrim[5], J. Liu[3], P.M. Bracci[6], C.F. Skibola[1].* 1) School of Public Health, Division of Environmental Health Sciences, University of California, Berkeley, CA 94720; 2) Unit of clinical epidemiology, Dept of Medicine Solna, Karolinska Institutet, Stockholm, Sweden; 3) Human Genetics, Genome Institute of Singapore, A*STAR, Singapore 138673; 4) Dept of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden; 5) Department of Epidemiology Research, Statens Serum Institut, Copenhagen, Denmark; 6) Department of Epidemiology and Biostatistics, University of California, San Francisco, CA 94143.

Introduction: Follicular lymphoma (FL) represents the second most common form of B-cell non-Hodgkin lymphoma (NHL). An important role for inherited genetic susceptibility in FL has been supported by recent genome-wide association studies (GWAS) where three independent susceptibility alleles have been identified on chromosome 6p21.33–32 in the HLA class I and class II regions. To identify new genetic variants, we conducted a meta-analysis of the top 1000 SNPs from existing GWAS data in 592 FL cases and 1541 controls from Denmark/Sweden (SCALE) and the San Francisco Bay Area (SF-NHL2). Validation genotyping of associated SNPs was conducted in 107 FL cases and 681 controls from an independent NHL case-control study population (SF-NHL1). The effect of validated SNP genotypes on gene expression levels was also investigated using publicly available microarray data.

Methods: We selected for meta-analysis the top 1000 SNPs most significantly associated with FL in the SF-NHL2 and SCALE GWAS. Selected SNPs in one dataset that were not genotyped or did not pass direct genotyping quality control in the other, were imputed using BEAGLE 3.0.3 and IMPUTEv1 with haplotype data from HapMap phase II-CEU or 1000 Genomes pilot1 CEU as reference. Odds ratios (ORs) for each individual study were estimated using unconditional logistic regression under an additive model in PLINK. OR estimates were combined in a meta-analysis under fixed- and random-effects inverse variance models using the metagen function from the meta package in R. Heterogeneity across studies also was tested with the Cochran's Q test and quantified with the I2 heterogeneity index.

Results: An additional SNP in the 6p21.32 HLA class II region was validated and strongly inversely associated with FL risk (OR=0.66, 95CI=0.57–0.77, p=1.02x10-7). The association was independent of previously established FL SNPs, but not of HLA-DPB1*0301 and also correlated significantly with HLA-DPB1 expression.

Conclusion: These results further support a major role for HLA genetic variation in FL pathogenesis.

2056T

Genome-wide linkage and association analyses in uterine leiomyomata reveal *FASN* as a risk gene. *S.L. Eggert[1], K.L. Huyck[4], P. Somasundaram[2], R. Kavalla[2], E.A. Stewart[5], A.T. Lu[6], J.N. Painter[7], G.W. Montgomery[7], S.E. Medland[7], S.A. Treloar[7,8], K.T. Zondervan[9], D.R. Nyholt[10], A.C. Heath[10], P.A.F. Madden[10], L. Rose[11], P.M. Ridker[11,12], D.I. Chasman[11,12], N.G. Martin[7], R.M. Cantor[6], C.C. Morton[2,3,12].* 1) Dept Genetics, Harvard Med Sch, Boston, MA; 2) Dept of Obstetrics, Gynecology and Reproductive Biology, Brigham and Women's Hospital, Boston, MA; 3) Dept of Pathology, Brigham and Women's Hospital, Boston, MA; 4) Dept of Medicine, Dartmouth-Hitchcock Medical Center, Lebanon, NH; 5) Dept of Obstetrics and Gynecology and Surgery, Mayo Clinic and Mayo Medical School, Rochester, MN; 6) Dept of Human Genetics, David Geffen School of Medicine at UCLA, Los Angeles, CA; 7) Queensland Institute of Medical Research, Brisbane, Australia; 8) The University of Queensland, Centre for Military and Veterans' Health, Herston, Australia; 9) Wellcome Trust Centre for Human Genetics and Nuffield Dept of Obstetrics and Gynaecology, University of Oxford, UK; 10) Washington University School of Medicine, St. Louis, MO; 11) Division of Preventive Medicine, Brigham and Women's Hospital, Boston, MA; 12) Harvard Medical School, Boston, MA.

Uterine leiomyomata (UL), the most prevalent pelvic tumors in women of reproductive age, pose a major public health problem given their high frequency, associated morbidities and most common indication for hysterectomies. A genetic component to UL predisposition is supported by analyses of ethnic predisposition, twin studies, and familial aggregation. A genome-wide SNP linkage panel was genotyped and analyzed in 261 white UL sister pair families from the Finding Genes for Fibroids study (FGFF). Two significant linkage regions were detected in 10p11 (LOD=4.15) and 3p21 (LOD=3.73) while five additional linkage regions were identified with LOD scores >2.00 in 2q37, 5p13, 11p15, 12q14 and 17q25. Genome-wide association studies were performed in two independent cohorts of white women and a meta-analysis was conducted. One SNP was identified with a p-value that reached genome-wide significance (rs4247357, P = 3.05E-08, odds ratio (OR) = 1.299). The candidate SNP is under a linkage peak and in a block of linkage disequilibrium in 17q25.3 which spans the genes fatty acid synthase (*FASN*), coiled-coil domain containing 57 (*CCDC57*) and solute carrier family 16, member 3 (*SLC16A3*). By tissue microarray immunohistochemistry, we found FAS protein expression elevated (3-fold) in UL when compared to matched myometrial tissue. *FAS* transcripts and/or protein levels are up-regulated in various neoplasms and implicated in tumor cell survival. *FASN* represents the first UL risk allele identified in white women by a genome-wide, unbiased approach and opens a path to management and potential therapeutic intervention.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2057F

X Chromosome markers associated with COPD: a meta-analysis in 3 cohorts. *M. Hardin[1], M. McDonald[1], M. Matthiessen[1], M. Cho[1], E. Wan[1], P. Castaldi[1], D. Lomas[2], P. Bakke[3], A. Gulsvik[3], J. Crapo[4], T. Beaty[5], C. Lange[1,6], E.K. SIlverman[1], D. DeMeo[1], The COPDGene and ECLIPSE investigators.* 1) Network Medicine, Brigham and Womens Hospital, Boston, MA; 2) Cambridge Institute for Medical Research, University of Cambridge, UK; 3) Institute of Medicine, University of Bergen, Norway; 4) Department of Medicine, University of Colorado, Denver CO; 5) Johns Hopkins Bloomberg School of Public Health, Baltimore, MD; 6) Harvard School of Public Health, Boston, MA.

Rationale: Chronic obstructive pulmonary disease (COPD) is a progressive disorder of the lungs. COPD susceptibility is influenced by genetic factors. Several GWAS have identified genes associated with COPD. However, these studies have generally excluded variants on the X Chromosome. Analysis of the X chromosome poses challenges compared to standard approaches to analyzing the autosomes. Recently, several investigators have developed methods to address these issues. We hypothesized that by utilizing methods specifically addressing X chromosome analysis, we would identify previously unrecognized associations between X chromosome SNPs and COPD. Methods: We analyzed the non-pseudo-autosomal region of the X chromosome in three Caucasian case/control cohorts of COPD. The three cohorts included ECLIPSE (1763 cases and 178 controls), GensKOLs (850 cases and 855 controls), and COPDGene (2774 cases and 2493 controls). All cases had COPD with GOLD stage II or higher. All controls were current or former smokers with normal lung function. Genotyping was performed using Illumina SNP arrays. Our primary analysis was via the method of Clayton et al as implemented in INTERSNP, using an allele-based test with 1 degree of freedom, counting hemizygous males as equivalent to homozygous females and adjusting for age, pack-years, and X-chromosome specific MDS components, followed by a meta-analysis of all three cohorts. A second analysis was performed using the logistic regression in PLINK with autosomally derived PCs. Results: 132,77 SNPs and 1941 subjects were analyzed from ECLIPSE, 12,014 SNPs and 1705 subjects were analyzed from GenKOLS, and 16966 SNPs and 5354 subjects were analyzed from COPDGene. The most significant associations included variants in PASD1 (ECLIPSE, OR 1.6, p = 9.96x10-5), MAGEB2 (COPDGene, OR 0.87, p=4.21x10-5), and DMD (GensKOLs, OR 0.81, p=4.9x10-4). The most significant variant in a meta-analysis of all cohorts was located in the PHEX gene (OR=1.15, P=4x10-4). In a meta-analysis of the results using logistic regression in PLINK, the most significant association was in the PASD1 gene (OR 0.8, p=1x10-4). Conclusions: Although no single variant achieved genome-wide significance, several associations were present among all three cohorts and using both analytical methods. Extending these analyses to a larger sample size and imputed data could improve statistical power to detect associations between markers on the X chromosome and COPD.

2058W

Periodontitis shares *IL2RA* as a genetic risk factor with rheumatoid arthritis, multiple sclerosis, type 1 diabetes, and Crohn's disease. *A.S. Schaefer[1], M. Nothnagel[2], C. Graetz[3], Y. Jockel[4], I. Harks[5], I. Staufenbiel[6], J. Eberhardt[6], E. Guzeldemir[7], N. Cine[8], E. Yilmaz[9], R. Nohutcu[9], B. Ehmke[5], P. Eickholz[10], M. Folwaczny[11], J. Meyle[12], U. Schlagenhauf[4], M. Laine[13], H. Dommisch[14], C. Bruckmann[15], B. Noack[16], B. Groessner-Schreiber[3], C. Doerfer[3], A. Franke[1], S. Jepsen[14], B.G. Loos[13], S. Schreiber[1].* 1) Institute of Clinical Molecular Biology, Christian-Albrechts-University, Kiel, Germany; 2) Institute of Medical Informatics and Statistics, Christian-Albrechts-University, Kiel, Germany; 3) Department of Operative Dentistry and Periodontology, University Medical Center Schleswig-Holstein Kiel, Germany; 4) Department of Periodontology, Clinic of Preventive Dentistry and Periodontology, University Medical Center of the Julius-Maximilians-University, Würzburg, Germany; 5) Center of Periodontology, Operative and Preventive Dentistry, Clinic of Preventive Dentistry, University Medical Center Münster, Germany; 6) Department of Conservative Dentistry, Periodontology and Preventive Dentistry, Hannover Medical School, Hannover, Germany; 7) Department of Periodontology, Faculty of Dentistry, Kocaeli University, Kocaeli, Turkey; 8) Department of Medical Genetics, Faculty of Medicine, Kocaeli University, Kocaeli, Turkey; 9) Department of Periodontology, Hacettepe University, Ankara, Turkey; 10) Center of Periodontology, Operative and Preventive Dentistry (Carolinum,) Clinic of Periodontology, Johann Wolfgang Goethe-University, Frankfurt am Main, Germany; 11) Department of Preventive Dentistry and Periodontology, University of Munich, Germany; 12) Department of Periodontology, University Medical Center Giessen and Marburg, Germany; 13) Department of Conservative and Preventive Dentistry, Section of Periodontology, Academic Center for Dentistry Amsterdam (ACTA), University of Amsterdam and VU University Amsterdam, The Netherlands; 14) Department of Periodontology, Operative and Preventive Dentistry, University of Bonn, Germany; 15) Department of Periodontology, Dental Clinic of the Bernhard-Gottlieb-University, Vienna, Austria; 16) University Medical Center Carl Gustav Carus der Technischen Universität Dresden, Center of Periodontology, Operative and Preventive Dentistry, Clinic of Preventive Dentistry, Dresden, Germany.

Periodontal diseases are highly prevalent and the major cause of tooth loss in adults above 40 years. The oral cavity is colonized by myriads of bacteria, which grow as complex colonies on tooth surfaces and on the oral mucosa. To some extent, it can be considered as a gateway for pathogens to other parts of the body, which may induce the development of systemic diseases. Periodontal diseases have been associated with a higher incidence of autoimmune conditions like rheumatoid arthritis (RA), systemic lupus erythematosus (SLE), and indirectly, of multiple sclerosis (MS). In the last five years, many disease-associated variants of genome-wide significance were identified, which contributed to the genetic basis of these autoimmune diseases, some of which showed a considerable overlap between different phenotypes. These factors often cluster within known regulatory pathways, and expose pathophysiologically relevant networks that are centered on specific transcriptional regulators. In an interdisciplinary medical research approach, we conducted the to date most comprehensive evaluation of the oral inflammatory disease periodontitis to identify shared genetic risk factors with the autoimmune diseases RA, SLE, and MS. We finemapped all major genetic risk loci (N=51) including their complete -cis regulatory regions by genotyping single nucleotide polymorphisms (SNPs) and insertions/deletions within these regions using the Illumina custom genotyping array immunochip. By a genetic association study of the 51 candidate regions in a total of 1,032 cases of the most severe and early-onset form aggressive periodontitis (AgP), further 1,351 cases of the more moderate but widespread form chronic periodontitis (CP), and 3,917 controls, we provide evidence for the intronic SNP rs11597237 of *IL2RA* to be associated with AgP (p=0.002, odds ratio=0.7 [95% confidence interval 0.6–0.9]). Our findings propose a genetic link between the oral cavity and the organism and further suggest the interrelation of inflammation and autoimmune disease.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2059T

Genetic factors underlying birth weight and cardiovascular disease: a study of 17 048 Finns. *K. Auro[1,2], E. Widen[3], J. Eriksson[1,4], A. Palotie[5], T. Lehtimäki[6], O. Raitakari[7], M. Perola[1,2].* 1) National Institute for Health and Welfare, Helsinki, Finland; 2) Helsinki University Central Hospital, Helsinki, Finland; 3) Institute for Molecular Medicine, Finland; 4) University of Helsinki, Finland; 5) Welcome Trust Sanger Institute, Cambridge, UK; 6) University of Tampere, Finland; 7) University of Turku, Finland.

Low birth weight has been solidly linked to cardiovascular diseases (CVD) at epidemiological level. Genetic factors linking birth weight and CVD are, however, poorly recognized. Our aim was to elucidate possible common genetic background underlying both fetal growth and CVD burden with combined genome and metabolome approach in large Finnish cohorts containing detailed phenotypic information. The study sample consists of three Finnish cohorts: The Young Finns Study, Helsinki Birth Cohort Study, and a subcohort of FINRISK 2002 sample. Together these cohorts contain total of 17 048 individuals having reliable birth measurements and data of CVD risk factors. Information of CVD events was gathered from hospital registers. Genome-wide genotypic (GWAS) data was available for 5 752 participants and metabolome data containing >100 serum metabolites for 8 288 participants. The study cohorts were pooled to gain statistical power. Gestation age, sex and maternal smoking were used as covariates. Only singleton pregnancies were included. The total amount of cardiovascular events in our study sample is 1800. Genotypic or metabolome data was available for 443 CVD cases. Birth weight associated with a novel gene locus (p=2.37x10-8) in CVD cases. In controls free of CVD, no association was seen. Birth weight also associated with several serum metabolites. With lipid particles, associations were seen with VLDL, LDL, and HDL levels measured in adulthood (P=0.0023–0.019). In addition, birth weight showed association with serum leptin (P=0.0013) and creatinine (p=0.0064) levels. Our results suggest a genetic link for fetal growth and CVD. Birth weight, when stratified for CVD status, associated with a novel gene locus with genome-wide statistical significance in a pooled study sample. Metabolome analysis also revealed association of birth weight with several serum metabolites, many involved in lipid metabolism. Since dyslipidemias significantly contribute to the CVD risk, these findings may indicate underlying, yet unknown genetic mechanisms common for both CVD and fetal growth. These novel findings elucidate the importance of fetal environment in future disease burden: factors affecting fetal growth may also guide the metabolic signals of the fetus into certain paths, contributing to the CVD risk in later life.

2060F

Genome-wide association study of orthostatic hypotension and supine-standing blood pressure changes in two Korean populations. *K. Hong, Y. Kim.* Center for Genome Science, Korea National Institute of Health, Cheongwon-gun, Chungbuk, South Korea.

Orthostatic hypotension is the failure of the cardiovascular reflexes to maintain blood pressure on standing from a supine position. The prevalence of orthostatic hypotension varies from 5% to 30% greatly depending on the population being studied. In the current study, we conducted genome-wide association studies for orthostatic hypotension case-control using the genome-wide SNP genotypes and also tested the supine-standing blood pressures as the quantitative traits. The study subjects were consisting of two population-based cohorts, Ansung and Ansan. The phenotypes and genotypes were obtained from Korean Genome and Epidemiology Study. The genome-wide association studies were conducted for each cohort, and the target phenotypes were the blood pressure differences in supine-standing systolic (dtSBP) and diastolic (dtDBP) and orthostatic hypotension. Among the target phenotypes, genome-wide association study for dtSBP detected a significant SNP, rs6736587 (meta-analysis p-value: 5.3 × 10-8) that located on chromosome 2p12. Interestingly, this region is known as the chromosome 2p12-11.2 deletion syndrome region and reported mental or psychomotor retardations. There is only one gene, CTNNA2 (alpha-N-catenin) around rs6736587±1Mbp. The CTNNA2 is a linker between cadherin adhesion receptors and the actin cytoskeleton, and is essential for stabilizing dendritic spines in rodent hippocampal neurons. Although there is no report about the function at blood pressure regulation, the hippocampal neurons interact primarily with the autonomic nervous system and might be related to the orthostatic hypotension.

2061W

Genetic association of erythrocyte and platelet phenotypes in Mexican Americans. *P.E. Melton[1], J.E. Curran[2], M. Carless[2], M.P. Johnson[2], T.D. Dyer[2], J.W. MacCluer[2], E.K. Moses[1], H.H.H. Goring[2], R. Duggirala[2], D. Glahn[3], J. Blangero[2], L. Almasy[2].* 1) Centre for Genetic Epidemiology and Biostatistics, University of Western Australia, Crawley AU; 2) Department of Genetics, Texas Biomedical Research Institute, San Antonio, TX; 3) Department of Psychiatry, Yale University School of Medicine, New Haven, CT.

Erythrocyte and platelet measurements are associated with risk of hematological disorders and early onset of cardiovascular disease. Most genetic studies on these phenotypes have been conducted in cohorts of European ancestry and little is known regarding genetic variants influencing these traits in other populations. This study reports family based genome-wide association study (GWAS) results for seven erythrocyte (hemoglobin, hematocrit, mean corpuscular volume, mean corpuscular hemoglobin, mean corpuscular hemoglobin concentration, and red blood cell count) and three platelet (platelet count, mean platelet volume, plateletcrit) phenotypes in 827 Mexican Americans from the Genetics of Brain Structure (GOBS) Study. Erythrocyte and platelet measures were averaged over three readings taken at time of blood draw. GWAS was conducted utilizing 931,219 SNPs from Illumina microarrays under an additive genetic model. Covariates included age, sex, their interactions as well as the first 10 principal components to account for population stratification. The strongest association identified was for plateletcrit on chromosome 17q24 with the intergenic SNP rs10853057 (p=2.45x10e-10) near GNA13. We found three additional significant intergenic signals for plateletcrit on chromosomes 1q43 (rs1361277, p= 2.0x10-8), 6p12 (rs7751328, p=2.0x10-8), and 8q24 (rs7006137, p=1.0x10-8). QQ plots were done for all phenotypes demonstrating significant SNPs and lambda values did not exceed 1.02. None of these observed SNPs for plateletcrit have been identified in other GWAS analyses. Additional associations were detected in intronic SNPs for mean platelet volume in STK38 on chromosome 6p21 (rs4236051, p =1.0x10-8) in a previously identified region for platelet function. Analysis of hematocrit (rs6807319, p= 5x10e-8) demonstrated a previously unidentified significant association in SLC9A9 on chromosome 3q23. This SNP also demonstrated suggestive association (p=2x10e-7) with hemoglobin. Further analysis of these results is being undertaken with whole-genome sequencing in a subset of GOBS participants. This study identified new genetic regions that may have important implications for better understanding of genetic variants influencing erythrocyte and platelet variation.

2062T

GWAS in an isolated Sardinian population contribute to elucidate the genetic control of serum Angiotensin Converting Enzyme (ACE) level. *I. Persico[1], M.P. Concas[1], G.B. Maestrale[1], L. Portas[1], F. Murgia[1], M. Cosso[2], D. Serra[2], M. Pirastu[1,2].* 1) Institute of Population Genetics, CNR, Sassari, Italy; 2) SharDNA Life Science, Pula, Italy.

INTRODUCTION: ACE, encoded by DCP1 gene, plays a key role in the Renin Angiotensin System which is involved in the control of fluid electrolyte balance and cardiovascular system. A QTL at DCP1 is the major determinant of serum ACE levels and more recently the AB0 locus was found associated with ACE levels by GWAS in hypertensive patients. However these two loci do not fully explain the variability of the trait suggesting a role of other genetic factors. PURPOSE: We aimed to detect the genetic components which determine the serum ACE levels in healthy individuals through GWAS of a genetically homogeneous population. METHOD. Talana is an isolated village characterized by centuries of isolation and high genetic homogeneity. We phenotyped 517 individuals (>18 yrs), genotyped with 2.5 million SNPs, without diseases and not taking medications that may interfere in ACE levels. The GWAS for discovery was carried out in two steps: 1) log10-transformed ACE levels adjusted for sex and principal components; 2) adding the DCP1 gene as covariate. The best hits were replicated in two villages of the same area (Baunei, N=848; Seulo, N=346), with different genetic structure and almost absent intermarriages with Talana, and in 1000 unrelated individuals from outbreed population, all selected with the same criteria. RESULTS: We confirmed the DCP1 locus on chr 17q23 (rs4311, p = 2.3E-18). We identified several loci with p values < 5.0E-8 spread in 10 Mb region (17q23.2–17q24.2) containing the DCP1. We did not observe linkage disequilibrium between all these loci. In addition we observed suggestive signals on chr 9q34.2, in 180 Kb region spanning AB0-SURF4-ADAMTS13 genes which reached a significant p value adding the DCP1 as covariate in GWAS second stage. We replicated the signals of the entire region 9q34.2 both in the two villages and in the outbreed population and we observed a strong effect of the A, B and O alleles on ACE levels. Finally we replicated some of the chr 17 hits (i.e. rs9895941, p=2.18E-9, MARCH10) in the two villages but not in the outbreed population. CONCLUSION: This study confirms the independent role played by DCP1 and AB0-Surf4-ADAMTS13 loci in the genetic control of the ACE level. It is interesting to note that both loci have been associated with cardiovascular disease. We also highlighted, thanks to the peculiar genetic structure of our isolated population, new genes/loci that should be confirmed with new genetic and functional studies.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2063F

Flipping sign test of GAWS summary statistics on multiple correlated traits. *Z. Zhang[1], N. Franceschini[2], T. Edwards[3], B. Keating[4], B. Tayo[5], E. Fox[6], A. Johnson[7], Y. Sun[8], Y. Sung[9], M. Nalls[10], J. Hunter[11], A. Dresbach[6], S. Musani[6], G. Papanicolaou[7], G. Lettre[12], A. Adebowale[13], R. Cooper[5], A. Reiner[14], D. Rao[9], D. Levy[7], X. Zhu[1].* 1) Department of Epidemiology and Biostatistics, School of Medicine, Case Western Reserve University, Cleveland, Ohio 44106, USA; 2) Department of Epidemiology, UNC Gillings School of Global Public Health, University of North Carolina at Chapel Hill, NC 27514, USA; 3) Center for Human Genetics Research, Vanderbilt Epidemiology Center, Department of Medicine, Vanderbilt University, Nashville, TN; 4) University of Pennsylvania School of Medicine, The Institute for Translational Medicine and Therapeutics, Philadelphia, PA, USA; 5) Department of Epidemiology and Preventive Medicine, Loyola University Stritch School of Medicine, Maywood, IL 60153, USA; 6) Department of Medicine, University of Mississippi Medical Center, Jackson, MS 39126, USA; 7) Center for Population Studies, National Heart, Lung, and Blood Institute, Framingham, MA 01702, USA; 8) Department of Epidemiology, Emory University, Atlanta, GA 30322; 9) Division of Biostatistics, Washington University in St. Louis, MO; 10) Molecular Genetics Section, Laboratory of Neurogenetics, NIA, NIH, Bethesda, MD 20892, USA; 11) Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA; 12) Montreal Heart Institute, Montréal, Canada; 13) Center for Research on Genomics and Global Health,National Human Genome Research Institute, Bethesda, MD 20892, USA; 14) Department of Epidemiology, University of Washington School of Public Health, Seattle, WA, 98195, USA.

GWAS is a routine approach to detect the genetic determinants of complex traits, but requires a large sample size in order to overcome multiplicity and subtle effect sizes. Successful GWAS are often conducted through meta-analysis by large consortia, which usually operate by sharing summary statistics from cohorts with the same phenotypes. When multiple correlated phenotypes, such as hypertension status, systolic and diastolic blood pressure (BP) are available, multivariate analysis is often more powerful than single trait analysis. However, it is difficult to perform multivariate analysis with summary statistics. We propose a novel permutation procedure, META (Multiple-outcome Empirical Trait Analysis), for multiple correlated traits using GWAS summary statistics. META preserves the linkage disequilibrium patterns among SNPs as well as the correlation structure among traits, and adjusts significance for both multiple testing and trait correlation. META accomplishes this by randomly flipping the sign of regression effect estimates from each cohort to obtain the distribution of meta-analysis test statistics for each SNP under the null hypothesis and empirically estimating significance. Our simulations show that the proposed permutation method can improve power up to 40%; compared to single trait analysis. We applied our method to the GWAS summary statistics from the Continental Origins and Genetic Epidemiology Network (COGENT) consortium study of BP traits, which includes 19 African ancestry cohorts. The proposed method identified two additional novel BP/hypertension loci that were not detected by the single trait analysis. Our results suggest that a multivariate approach should be used when multiple correlated traits are studied, and furthermore that this enhanced analysis can also be done using summary statistics.

2064W

Association study on the X chromosome identifies novel genes associated with systemic lupus erythematosus (SLE) in Asian populations. *W. Yang[1], Y. Zhang[1], N. Hirankarn[3], X. Zhang[2], Y. Lau[1].* 1) Paediatrics & Adolescent Med, Univ Hong Kong, Hong Kong, Hong Kong; 2) Anhui Medical University, Hefei, China; 3) Chulalongkorn University, Bangkok, Thailand.

SLE is an autoimmune disease with striking female predominance. The sex chromosomes may play an important role in the increased female prevalence of the disease. The prevalence rate of SLE increases by 14 folds in males with Klinefelter's syndrome (47, XXY), to a level similar to that of females. Several genes on the X chromosome were confirmed by candidate gene approach as associated with SLE. In this study, we performed a meta-analysis of X chromosome variants of the two Genome-wide association studies (GWAS) in Asian populations from Hong Kong and Anhui, China. Based on the GWAS data, together with replication from more samples from different research centers in Hong Kong, Anhui and Bangkok, Thailand, we have shown a much clearer picture of X-linked genes related to lupus. A novel locus, PRPS2 (P=3.9E-09, OR=0.84), and a risk haplotype spanned a region from ARHGAP4 to MECP2 (P=1.8E-05) were confirmed by the meta-analysis and replication in three cohorts. Clear difference between the susceptibility loci identified in Asians and those from Caucasians is also demonstrated. These results may help us understand the gender difference and the role of the X chromosome in this disease.

2065T

Genetic modifiers of the palatal phenotype in 22q11 microdeletion syndrome. Preliminary results of a genome-wide association study. *C. Vial[1], G.M. Repetto[1,2], M. Palomares[3,4], S. McGhee[5], N.K. Henderson-MacLennan[6], M.L. Guzman[1], K. Espinoza[1], G. Lay-Son[1,2].* 1) Center for Human Genetics, Facultad de Medicina, Universidad del Desarrollo-Clinica Alemana, Santiago, Chile; 2) Hospital Padre Hurtado, Santiago, Chile; 3) Fundación Gantz, Chile; 4) Hospital Calvo Mackenna, Santiago, Chile; 5) Stanford University, Stanford, CA;; 6) UCLA, Los Angeles, CA.

Chromosome 22q11 microdeletion syndrome (del22q11) is one of the most common pathogenic genomic rearrangements in humans, with an estimated frequency of 1/4000 live births. Most patients share a common 3Mb deletion, but the clinical phenotype shows marked variable expressivity. Approximately 70–80% of patients have palatal anomalies, ranging from overt cleft palate to occult submucous cleft leading to velopharyngeal insufficiency and hypernasal speech. The cause of the incomplete penetrance of the palatal anomalies is unknown. We performed a genome wide association study (GWAS) to search for genetic modifiers of the palatal phenotype in Chilean patients with del22q11. DNA samples from 91 Chilean patients with del22q11 demonstrated by FISH with TUPLE1 probe (Abbott) and known palate anatomy by physical examination and nasopharyngoscopy were analyzed with Affymetrix ® v. 6.0 SNP array. We compared cases with any palate anomaly (n=69) to controls with normal palates (n=22). After filtering, genotypes, over 728,000 SNPs were analyzed with PLINK v1.07, using Fisher exact test for association. A population stratification analysis was done using the principal component multi-dimensional scaling (MDS) test. This results were used to perform an association with MDS information using the Cochran-Mantel-Haenszel test. A Bonferroni-corrected p value of 1 x 10-7 was considered as evidence of association. Two regions were found to have an association with the palatal phenotype. The first one was between 7q11.23–7q21.11 and included 4 SNPs with a p value between 7 x 10-8 and 1.3 x 10-5. Another region of 120kb in 20q13.12 included 8 SNPs with p values between 1.3 x 10-6 and 5.6 x 10-5. Although a larger sample size is needed to reach statistical significance in the latter region, these results point out two potential regions of interest, both within gene rich regions which will have to be further studied. Funded by Fondecyt-Chile Grant #1100131.

2066F

Genome-wide copy number variation in anorectal malformations. *S.S. Cherny[1], E.H.M. Wong[1], L. Cui[2], C.L. Ng[2], C.S.M. Tang[1], M.T. So[2], B.H.K. Yip[1], G. Cheng[2], V.C.H. Liu[2], P.C. Sham[1], P.K.H. Tam[2], M.M. Garcia-Barcelo[2].* 1) Psychiatry, Univ Hong Kong, Pokfulam, Hong Kong; 2) Surgery, Univ Hong Kong, Pokfulam, Hong Kong.

Anorectal malformations (ARMs, congenital obstruction of the anal opening) are among the most common birth defects requiring surgical treatment (2–5/10,000 live-births) and carry significant chronic morbidity. ARMs present either isolated or as part of the phenotypic spectrum of some chromosomal abnormalities or monogenic syndromes. The etiology is unknown. To assess the genetic contribution to ARMs, we investigated single nucleotide polymorphisms (SNPs) and copy number variants (CNVs) at genome-wide scale. A total of 363 Han Chinese sporadic ARM patients and 4,006 Han Chinese controls were included. Overall, we detected a 1.3-fold significant excess of rare CNVs in patients. Stratification of patients by presence/absence of other congenital anomalies showed that while syndromic ARM patients carried significantly longer rare duplications than controls ($p < .05$), non-syndromic patients were enriched with both rare deletions and duplications when compared to controls ($p < .0004$). Twelve chromosomal aberrations and 114 rare CNVs were observed in patients but not in 868 controls or 11,943 healthy individuals from the Database of Genomic Variants (DGV). Importantly, the aberrations were observed in isolated-ARM patients. Gene-based analysis revealed 79 genes interfered by CNVs in patients only. In particular, we identified a *de novo DKK4* duplication and an *INTU* hemizygous deletion. These genes are members of pathways (*Shh* and *Wnt*, respectively) involved in the development of the anorectal region. In mice, *Shh* or *Wnt* disruption results in ARMs. Our data suggest a role for rare CNVs not only in syndromic but also in isolated ARM patients and provides a list of plausible candidate genes for the disorder.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2067W

A genome-wide meta-analysis of copy number variation identifies novel childhood obesity loci. *D. Hadley[1], J. Bradfield[1], H. Hakonarson[1,2], S. Grant[1] on behalf of the Early Growth Genetics Consortium.* 1) The Center for Applied Genomics, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA; 2) Division of Genetics and Department of Pediatrics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania, USA.

Recently, single nucleotide polymorphisms (SNPs) have been associated with the common form of early-onset obesity through a collaborative meta-analysis of North American, Australian and European pediatric populations (Bradfield et al, 2012). However, only limited progress has been made in understanding the influence of copy number variants (CNVs) in common childhood obesity. To that end, we developed a novel method to perform CNV meta-analyses and applied it to a large subset of the same three global populations studied by Bradfield et al to conduct the first meta-analysis of CNVs in common early-onset obesity. We compared over 2.2K cases to over 5.8K controls that passed quality control to discover genome-wide significant signals of association. We found common (frequency ≥ 5%) deletions at chromosomes 4q13.1 ($P \leq 2.32x10^{-6}$, OR = 1.31) and 6q14.1 ($P \leq 7.85x10^{-6}$, OR = 1.25), rare (frequency < 5%) deletions at chromosomes 6p22.3 ($P \leq 3.02x10^{-5}$, OR = 3.37), 14q11.2 ($P \leq 8.05x10^{-7}$, OR = 1.76), and 20q13.2 ($P \leq 8.69x10^{-5}$, OR = 2.71), and rare duplications at chromosome 1p32.2 ($P \leq 3.95x10^{-5}$, OR = 6.54). Most notably, the deletions at chromosome 6p22.3 were highly penetrant and span 25Kb of the intergenic region immediately downstream of *SOX4* which encodes a HMG box transcription factor best known to contribute to the development of the endocrine pancreas. Moreover, the deletions at *SOX4* were robustly identified from at least one site across all three continental populations analyzed. Our work is significant not only because it sheds light on the contributions of CNVs in the pathophysiology of common, early-onset childhood obesity, but also because, to the best of our knowledge, it is the first multi-site, genome-wide meta-analysis of CNVs performed for any trait.

2068T

Copy number variation analysis implicates locus 5q21.2 with raised intra-ocular pressure. *A. Nag[1], P.G. Hysi[1], C. Venturini[2], S. MacGregor[3], A.W. Hewitt[4], T.L. Young[5], P. Mitchell[6], A.C. Viswanathan[2], D.A. Mackey[7], C.J. Hammond[1], WTCCC2.* 1) King's College London, London, United Kingdom; 2) NIHR Biomedical Research Centre for Ophthalmology, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, UK; 3) Genetics and Population Health, Queensland Institute of Medical Research, Brisbane, Australia; 4) Centre for Eye Research Australia, University of Melbourne, Royal Victorian Eye and Ear Hospital, Melbourne, Australia; 5) Centre for Human Genetics, Duke Univ Med Ctr, Durham, NC; 6) Centre for Vision Research, Westmead Millennium Institute, University of Sydney, Australia; 7) Lions Eye Institute, University of Western Australia, Perth, Australia.

Glaucoma is a major cause of visual impairment and the second leading cause of blindness in the world. Rare sequence mutations in the genes myocilin (MYOC) and optineurin (OPTN) explain about 2–5 percent of primary open angle glaucoma cases. Lately, GWAS have identified common genetic variants for glaucoma, but still a significant heritability gap remains. We hypothesized that copy number variants (CNVs) might influence part of the susceptibility to glaucoma or its related quantitative phenotypes. The study examined the association between CNV and intra-ocular pressure (IOP), the major modifiable risk factor for glaucoma, in a sample of 781 unrelated individulas from the TwinsUK cohort. The samples were genotyped with the Illumina Human Hap610 Quad. The software package PennCNV was used to detect CNV regions for each sample. Appropriate quality control measures were implemented and CNV regions encompassing less than 10 SNPs were filtered out. Association testing between the CNVs detected and IOP was carried out using the software package GenABEL. Regression was performed to test for the association between a CNV and IOP. CNV regions which were less than 1% in frequency were excluded from further analysis. We found association with IOP for two CNV regions: 5q21.2 (p=0.013) and 22q11.2 (p=0.03). All the 5q21 CNVs overlapped the gene RAB9BP1 and all the 22q11 CNVs overlapped the gene CRYBB2P1. Replication studies were carried out in a sample of 467 Australian twins and in a sample 1620 individuals from the Blue Mountains Eye Study (BMES) cohort. Both studies successfully replicated the 5q21.2 CNV association and direction of effect (p=0.001 and p=0.003 respectively). A meta-analysis of this locus across the 3 cohorts showed that per allele copy number change increased the IOP by 1.62mmHg (p=8.67 x 10-7). The frequency of this CNV region in the TwinsUK cohort, the Australian twins cohort and the BMES cohort was 1.3%, 2.8% and 1% respectively. Interestingly, this IOP locus overlaps the previously widely replicated GLC1G linkage locus for glaucoma, for which subsequent studies have not reached consensus on the causal gene (including WDR36). Our results suggest that copy numbers changes of the RAB9BP1 gene may explain the linkage signal at this locus. .

2069F

Genome-wide Analysis of Copy Number Variants in Down Syndrome Associated Congenital Heart Defects. *D. Ramachandran[1], J. Mulle[2], AE. Locke[3], P. Bose[1], LJ. Bean[1], S. Le[1], T. Rosser[1], K. Dooley[4], DJ. Cutler[1], E. Feingold[5], SY. Cheong[6], CL. Cua[7], CL. Maslen[8], RH. Reeves[9], SL. Sherman[1], ME. Zwick[1].* 1) Human Genetics, Emory University, Atlanta, GA; 2) Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA; 3) Center for Statistical Genetics and Department of Biostatistics, University of Michigan School of Public Health, MI; 4) Sibley Heart Center Cardiology, Children's Hospital of Atlanta, GA; 5) Department of Human Genetics, University of Pittsburgh, Pittsburgh, PA; 6) Battelle Center Mathematical Medicine, Nationwide Children's Hospital, Columbus, OH; 7) Heart Center, Nationwide Children's Hospital, Columbus, OH; 8) Department of Molecular and Medical Genetics, Division of Cardiovascular Medicine, Oregon Health & Science University, Portland, Ore; 9) Department of Physiology, McKusick Nathans Institute for Genetic Medicine, School Of Medicine, John Hopkins University, MD.

Our project is focused on identifying genetic variants contributing to congenital heart defects (CHD), especially atrioventricular septal defect (AVSD) among individuals with Down Syndrome (DS). Non-syndromic AVSDs occur in approximately 1 in 10,000 births in the general population. However, among infants with DS nearly 20% have an AVSD, representing a nearly 2000-fold increased risk of AVSD compared to the general population. We hypothesize that in the presence of extra chromosome 21, otherwise benign copy number variants (CNVs) become susceptibility alleles and hence could explain the variable penetrance observed in DS-associated AVSD. We have used Affymetrix SNP 6.0 genotyping platform to comprehensively characterize the CNVs in a large DS population to test this hypothesis. The samples were collected through a multi-site recruitment effort. Cases were individuals with DS and a complete AVSD (DS+AVSD) and controls were those with DS and the documented absence of CHD (DS-CHD). We have developed an analysis pipeline for the quality control (QC) of samples and CNV calling aimed at minimizing false positive calls, including 3 algorithms to make the CNV calls, namely BEAST, GADA and GLAD. After QC filtering, we have a total of 831 samples consisting of 239 cases, 268 controls and 324 case parents for downstream analyses. Preliminary analyses identified a total of 434 CNVs. Among those, 139 CNVs were common to both cases and controls, 114 were present only in cases and 119 in controls. The size of the CNVs range from 964bp to 29.9Mb. In-depth analyses of these loci are presently underway for DS cases, controls and parents. Any interesting candidate loci identified will be validated independently. This study has the potential to shed light on as-yet unknown pathways that give rise to AVSD.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2070W

Rare genetic variation in relation to circulating adiponectin. *S. Gustafsson[1], L. Lind[2], A. Mahajan[3], S. Söderberg[4], A. Flyvbjerg[5], C.M. Lindgren[3], A.P. Morris[3], I. Ingelsson[1].* 1) Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden; 2) Department of Medical Sciences, Uppsala University, Uppsala, Sweden; 3) Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom; 4) Department of Public Health and Clinical Medicine, Cardiology and Heart Centre, Umeå University, Umeå, Sweden; 5) The Medical Research Laboratories, Department of Clinical Medicine, Faculty of Health, Aarhus University, Aarhus, Denmark.

Background: Previous studies have identified genetic variants associated with the phenotypic variance of adiponectin, mainly focusing on common genetic variation which to date explains a small fraction of the expected adiponectin variance. Our aim was to investigate the impact of rare genetic variants on circulating adiponectin. **Methods**: We investigated the relationship between circulating adiponectin levels and the proportion of minor alleles at rare genetic variants within a gene carried by an individual in two population-based cohorts, including 948 individuals from the Prospective Investigation of the Vasculature in Uppsala Seniors (PIVUS) and 1106 individuals from the Uppsala Longitudinal Study of Adult Men (ULSAM). Samples were genotyped on Illumina OmniExpress (PIVUS), Illumina Omni 2.5M (ULSAM) and Illumina Metabochip (both samples), and further imputed using haplotypes from the 1000 Genomes, June 2011 release. The software GRANVIL was used to perform the regression analysis and the results were combined across cohorts in a fixed-effect sample size-weighted Z-score meta-analysis. **Results**: In models adjusted for population stratification the proportion of minor alleles at rare variants within a transcript of *RPS27L* was significantly and positively associated with levels of adiponectin (P-value=$5.1\text{x}10^{-7}$), which remained significant after adjustment for BMI (P-value= $2.3\text{x}10^{-7}$), but more modestly associated in a subset of non-diabetic individuals. The transcripts of *FBXL14* and *WNT5B* were significantly associated with adiponectin in models adjusting for population stratification and BMI in a subset of non-diabetic individuals, when restricting the analysis to non-synonymous markers. With these settings only one non-synonymous marker at chr12:1702929 was included in the analysis of transcript regions of *FBXL14* and*WNT5B*. **Discussion**: Our results indicate that individuals who carry a larger proportion of rare alleles within a transcript of *RPS27L* have on average higher levels of adiponectin. A significant positive association was observed for each of the overlapping transcript regions of *FBXL14* and *WNT5B*, all adjacent to *ADIPOR2*, which included only one non-synonymous marker which is equivalent with a dominant effect single-marker association test. Prior studies have not been able to establish these regions to be associated with circulating adiponectin levels, so our results are of importance for understanding of adiponectin biology if replicated.

2071T

Comprehensive evaluation of type 2 diabetes susceptibility loci in East Asians by using 1000 Genomes Project data. *K. Hara[1], H. Fujita[1], T.A. Johnson[2], S. Maeda[2], T. Tsunoda[2], M. Kubo[2], T. Kadowaki[1].* 1) Department of Diabetes and Metabolic Diseases, Graduate School of Medicine, University of Tokyo, Tokyo 113-8655, Japan; 2) RIKEN Center for Genomic Medicine, Yokohama Kanagawa 230-0045.

Although over 50 type 2 diabetes (T2D) loci have been identified through genome-wide association studies (GWAS), the vast majority of the genetic predisposition to T2D still remains to be clarified. One explanation is that despite the use of high-throughput genotyping arrays, only a small proportion of genetic variants in the human genome are actually surveyed, especially in non-European populations. We explored the comprehensive catalog of genomic variations provided by the 1000 Genomes Project to identify variations conferring susceptibility to T2D in the Japanese population that were not detected in previous scans. We imputed 10,524,368 variants derived from 194 East Asian subjects (June 2011 release) into 5,976 cases and 20,829 controls genotyped by 610K single-nucleotide polymorphism (SNP) array. Overall concordance was good, although imputed SNPs with minor allele frequency (MAF) below 5 % apparently contained poorly imputed SNPs. We then tested associations for T2D before and after adjusting for age, sex, and body mass index. We confirmed that 8 previously reported loci were associated with T2D at the level of genome-wide significance ($P<5\times10^{-8}$). To briefly assess the similarities of genetic architecture between ethnic groups, we compared the effect sizes of previously reported and newly discovered variants between European and East Asian populations. There was significant concordance in the direction of effects among the two populations, indicating that the genetic factors contributing to T2D were substantially shared among these populations. In addition, we found 30 loci harboring multiple variants associated with T2D at $P<5\times10^{-5}$, seven of which had been reported previously. The remaining 23 previously unreported loci were subsequently selected for in silico (2,799 cases and 3,793 controls) and de novo genotyping (10,319 cases and 6,795 controls) follow-up studies focused on East Asian subjects. We also sought to define the most relevant SNPs for susceptibility to T2D in the previously identified genes. We did not find any stronger association with T2D than the originally reported SNPs in our population. Our study highlights the benefit of using data derived from next-generation sequencing of the human genome such as the 1000 Genomes Project to explore T2D loci more comprehensively.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2072F

Genome-wide association study of age-related macular degeneration identifies associated variants in the TNXB-FKBPL-NOTCH4 region of chromosome 6p21.3. *V. Cipriani[1,2], H.-T. Leung[3,4], V. Plagnol[5], C. Bunce[1,2,6], J.C. Khan[3,7], H. Shahid[3,8], A.T. Moore[1,2], S.P. Harding[9], P.N. Bishop[10,11], C. Hayward[12], S. Campbell[12], A.M. Armbrecht[13], B. Dhillon[13], I.J. Deary[14], H. Campbell[15,16], M. Dunlop[15], A.F. Dominiczak[17], S.S. Mann[1,2,18], S.A. Jenkins[1,2,19], A.R. Webster[1,2], A.C. Bird[1,2], M. Lathrop[20,21], D. Zelenika[20], A.J. Cree[22], J. Gibson[23], S. Ennis[23], A.J. Lotery[22,24], A.F. Wright[12], D.G. Clayton[3], J.R.W. Yates[1,2,3].* 1) Institute of Ophthalmology, University College London, London, EC1V 9EL, UK; 2) Moorfields Eye Hospital, London, EC1V 2PD, UK; 3) Department of Medical Genetics, Cambridge Institute for Medical Research, University of Cambridge, Cambridge, CB2 0XY, UK; 4) Department of Ophthalmology and Visual Sciences, The Chinese University of Hong Kong, Hong Kong, China; 5) UCL Genetics Institute, London, WC1E 6BT, UK; 6) Department of Infectious and Tropical Diseases, London School of Hygiene & Tropical Medicine, University of London, London, WC1E 7HT, UK; 7) Department of Ophthalmology, Royal Perth Hospital, Perth, Western Australia 6001, Australia; 8) Department of Ophthalmology, Addenbrooke's Hospital, Cambridge, CB2 0QQ, UK; 9) Department of Eye and Vision Sciences, Institute of Ageing and Chronic Disease, UCD Building, University of Liverpool, Liverpool, L69 3GA, UK; 10) School of Biomedicine, Faculty of Medical and Human Sciences, University of Manchester, Manchester, M13 9PT, UK; 11) Centre for Advanced Discovery and Experimental Therapeutics, Manchester Academic Health Science Centre, Central Manchester Foundation Trust, Manchester, M13 9PL, UK; 12) Medical Research Council Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, EH4 2XU, UK; 13) Department of Ophthalmology, University of Edinburgh and Princess Alexandra Eye Pavilion, Edinburgh, EH3 9HA, UK; 14) Centre for Cognitive Ageing and Cognitive Epidemiology, Department of Psychology, University of Edinburgh, Edinburgh, EH8 9JZ, UK; 15) Colon Cancer Genetics Group, Institute of Genetics and Molecular Medicine, University of Edinburgh and MRC Human Genetics Unit, Edinburgh, EH4 2XU, UK; 16) Centre for Population Health Sciences, University of Edinburgh, Edinburgh, EH8 9AG, UK; 17) Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow, G12 8QQ, UK; 18) St Thomas' Hospital, London, SE1 7EH, UK; 19) Department of Inherited Cardiovascular Disease, The Heart Hospital, London, W1G 8PH, UK; 20) Centre National de Génotypage, Institut Génomique, Commissariat à l'Énergie Atomique, Evry, France; 21) Fondation Jean Dausset-CEPH, Paris, France; 22) Clinical Neurosciences Research Grouping, Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton, SO16 6YD, UK; 23) Genetic Epidemiology and Genomic Informatics Group, Human Genetics, Faculty of Medicine, University of Southampton, Southampton, SO16 6YD, UK; 24) Southampton Eye Unit, Southampton General Hospital, Southampton, SO16 6YD, UK.

Age-related macular degeneration (AMD) is the commonest cause of blindness in Western populations. Susceptibility is influenced by age, environmental and genetic factors. Known genetic risk loci do not account for all the heritability. We therefore carried out a genome-wide association study of AMD in the UK population with 893 cases of advanced AMD and 2199 controls. This showed association with the well established AMD risk loci ARMS2-HTRA1 ($P = 2.7 \times 10^{-72}$), CFH ($P = 2.3 \times 10^{-47}$), C2-CFB ($P = 5.2 \times 10^{-9}$), C3 ($P = 2.2 \times 10^{-3}$) and CFI ($P = 3.6 \times 10^{-3}$) and with more recently reported risk loci at VEGFA ($P = 1.2 \times 10^{-3}$) and LIPC ($P = 0.04$). Using a replication sample of 1411 advanced AMD cases and 1431 examined controls we confirmed a novel association between AMD and single nucleotide polymorphisms on chromosome 6p21.3 at TNXB-FKBPL (rs12153855/rs9391734; discovery $P = 4.3 \times 10^{-7}$, replication $P = 3.0 \times 10^{-4}$, combined $P = 1.3 \times 10^{-9}$, OR = 1.4, 95% CI = 1.3–1.6) and the neighbouring gene NOTCH4 (rs2071277; discovery $P = 3.2 \times 10^{-8}$, replication $P = 3.8 \times 10^{-5}$, combined $P = 2.0 \times 10^{-11}$, OR = 1.3, 95% CI = 1.2–1.4). These associations remained significant in conditional analyses which included the adjacent C2-CFB locus. Further research will be needed to identify the causal variants and determine whether any of these genes are involved in the pathogenesis of AMD. TNXB, FKBPL and NOTCH4 are all plausible AMD susceptibility genes, and NOTCH4 in particular merits further investigation as the Notch signalling pathway has already been suggested as a potential therapeutic target for AMD.

2073W

Evidence for association and biologic support of a novel locus for serum albumin. *N. Franceschini[1], F.J.A. van Rooij[2], B.P. Prins[3], M.F. Feitosa[4], M. Karakas[5], Q. Wu[6], C.A. Winkler[7], J. Kopp[8], J.F. Felix[2], K.E. North[1], B.Z. Alizadeh[3], L.A. Cupples[9], J.R.B Perry[10], A.P. Morris[11] on behalf of the CHARGE Consortium Protein Working Group.* 1) Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, NC; 2) ErasmusAGE, Department of Epidemiology, Erasmus MC, Rotterdam, the Netherlands; Netherlands Consortium for Healthy Aging (NGI-NCHA), The Netherlands Genomics Initiative, Leiden, the Netherlands; 3) Department of Epidemiology, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands; 4) Division of Statistical Genomics, Department of Genetics, Washington University School of Medicine, St. Louis, MO, USA; 5) Department of Internal Medicine II-Cardiology, University of Ulm Medical Center, Ulm, Germany; 6) Molecular Cardiology/Nephrology & Hypertension Lerner Research Institute/NB20 Cleveland Clinic, Cleveland, OH, USA; 7) Molecular Genetics Epidemiology Section, National Institute of CancerFrederick National Laboratory for Cancer Research, National Institutes of Health, Frederick, MD, USA; 8) National Institute of Diabetes and Digestive and Kidney Diseases, National Institute of Health, Bethesda, MD, USA; 9) Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA and National Heart, Lung, and Blood Institute's (NHBLI's) Framingham Heart Study; 10) Genetics of Complex Traits, Peninsula Medical School, University of Exeter, UK; 11) Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK.

Serum protein concentrations are highly heritable and are altered in many disorders including kidney diseases, malnutrition, cancer, and inflammatory diseases. To identify loci associated with serum albumin concentration, the major plasma protein, we performed a genome-wide association study of up to 53,190 European ancestry individuals from 19 studies. Each study performed association analysis using linear regression or mixed models (for family data) while adjusting for age, sex and measures of population stratification. These results were then combined using fixed-effects inverse-variance weighted meta-analysis. We identified four novel loci for serum albumin at genome-wide significance ($p<5\times10^{-8}$), including *HPN-SCN1B*($p=3.3\times10^{-15}$) which showed evidence of two independent association signals using approximate conditional analyses (GCTA): rs4806073, $p=1.6\times10^{-12}$ and rs11671010, $p=1.5\times10^{-11}$; CEPH Europeans from Utah $r^2=0.02$, 4.3kb apart). The lead SNP at the *HPN-SCN1B* locus maps to an intron of *HPN*, a gene encoding hepsin, a membrane-bound serine protease with substrate specificity for basic amino acids, similar to those involved in proalbumin processing to albumin. In *hepsin* -/- mice, we observed reduced serum albumin concentration compared with wild-type litter-mates ($p=9.1\times10^{-12}$). We also confirmed the previously reported associations in East Asian populations at *GCKR-FNDC4* ($p=2.9\times10^{-14}$) and *RPS11-FCGRT* ($p=3.2\times10^{-8}$) for serum albumin. *FCGRT* encodes the heavy alpha chain of the FcRn, which prevents lysosomal degradation of albumin and immunoglobulins in lysosomes and thereby extends their serum half-life. KO *fcgrt* mice manifest low serum albumin. Our findings support a biologic role for the *HPN* locus on serum albumin homeostasis and confirm the *RPS11-FCGRT* locus identified in animal models. These loci may have potential roles in disease states.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2074T

Gene-based meta-analysis of genome-wide association studies implicates new susceptibility loci for obesity. *S. Hägg, Y. Pawitan, E. Ingelsson on behalf of the GIANT consortium.* Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden.

Genome-wide association studies (GWAS) have identified 36 loci associated with body mass index (BMI) to date. The majority of these have been identified through standard meta-analyses of single variants. This approach may miss loci with allelic heterogeneity. Therefore, the aim of the present study was to discover additional susceptibility loci for BMI using gene-based meta-analysis to identify regions with high allelic heterogeneity. We included data from ~124,000 adult individuals of European descent from 46 cohorts of the Genetic Investigation of Anthropometric Traits (GIANT) consortium in the discovery stage, and data from additionally 112,000 individuals from 43 cohorts in the replication stage. Each cohort was tested for association between ~2.4 million imputed or genotyped single nucleotide polymorphisms (SNPs) and BMI, and summary statistics were subsequently meta-analyzed in 17,941 autosomal genes based on UCSC Genome Browser hg18 assembly. We used a software package called `Versatile gene-based test for genome-wide association studies' (VEGAS), which assigns SNPs to genes by accounting for LD-block structure using simulations, to calculate gene-based p-values. The VEGAS method was applied to each cohort separately before a gene-based meta-analysis using the Fisher sum statistics with chi-squared distribution was performed. Of all genes tested, two loci which have already associated with BMI (*FTO* and *TMEM18*), as well as six novel loci (*TXNDC12, PEX2, FAM54A*, *SSFA2*, *IARS2* and *KIAA1731*) were associated with BMI at genome-wide significance (p-value < $2*10^{-6}$). Further, replication following the same procedure confirmed six of the eight loci to be genome-wide significant, and all eight loci well below a Bonferroni-corrected threshold of replication (alpha=0.00625 for eight tests). Using a new approach to do gene-based meta-analyses of large-scale GWAS data, we have reproduced results for two already identified BMI susceptibility loci, as well as contributed with six novel potential target genes associated with BMI. Our results indicate that gene-based meta-analysis of GWAS can provide a useful strategy to find additional susceptibility loci that are not identified in standard single-marker analyses due to high allelic heterogeneity.

2075F

Homozygosity mapping in a Sindh Pakistan family identifies new candidate regions for primary autosomal recessive microcephaly. *P. Lemay[1], M. Sindhi[2], N. Sehar[2], S. Kashif[3], Q. Brohi[4], J. Michaud[1], Z. Kibar[1].* 1) CHU Ste-Justine, Montreal, Quebec, Canada; 2) Department of Zoology, University of Sindh Jamshoro, Sindh, Pakistan; 3) Autism Institute Karachi Sindh, Pakistan; 4) Sir Cowasjjee Jehangir Institute of Psychiatry, Hyderabad, Sindh, Pakistan.

Microcephaly is defined as a condition in which the head circumference of an affected individual is >3 standard deviation below the population age and sex related mean. Autosomal recessive primary microcephaly (MCPH) is a neurodevelopmental disorder which is characterized by two principal features, microcephaly present at birth and non-progressive mental retardation. Primary microcephaly is further defined by the absence of other syndromic features or significant neurological deficits (other than mild-to-moderate mental retardation. MCPH has an estimated incidence of 1 in 1 000 000 in the Caucasian populations but is particularly frequent in northern Pakistan where it occurs in approximately 1 in 10 000 individuals. To date, 8 loci and 5 genes were associated with microcephaly in the Northern Pakistan region. We have recruited 10 extended families affected with MCPH from the Sindh Pakistan region. This southern region of Pakistan was never investigated for MCPH and could result in identification of new associated loci. A homozygosity mapping analysis was performed using the Illumina Human660w quad beadchip and the HomozygocityMapper program on 2 distantly related affected individuals from one large family of our Pakistani cohort. A total of 7 homozygous regions were identified: chr5: 27564–1936443, chr10: 53806602–58948207, chr10:113735304–116047921, chr11:7714879–115 31547, chr11:11592794–14944645, chr13: 101755185–104391808, chr19: 48557013–51595182. No known MCPH genes were found in these regions. Using a combination of homozygosity mapping and whole exome sequencing, we will reduce the number of identified potentially pathogenic coding variants facilitating the identification of the causative MCPH gene in this family.

2076W

Large-scale rare-variant analyses of eight quantitative traits reveal novel loci for triglycerides and fasting insulin in genome-wide studies imputed from 2188-haplotype 1000 Genomes reference panel. *R. Mägi[1], M. Horikoshi[2,3], I. Surakka[4,5], S. Wiltshire[2,3], A.-P. Sarin[4,5], T. Esko[1], A. Mahajan[2], T. Ferreira[2], M. Beekman[6,7], S. Gustafsson[8], S. Hägg[8], C. Ladenvall[9], L. Marullo[2,10], C.P. Nelson[11,12], J.S. Ried[13], G. Thorleifsson[14], N. Tsernikova[1], S.M. Willems[15], C. Willenborg[16], T. Winkler[17], C.M. Lindgren[2], M.I. McCarthy[2,3], S. Ripatti[4,5], I. Prokopenko[2,3], A.P. Morris[2] for ENGAGE Consortium.* 1) Estonian Genome Center, University of Tartu, Tartu, Estonia; 2) Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK; 3) Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford, UK; 4) Institute for Molecular Medicine Finland FIMM, University of Helsinki, Helsinki, Finland; 5) National Institute for Health and Welfare, Helsinki, Finland; 6) Department of Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands; 7) Netherlands Consortium for Healthy Ageing, Leiden, The Netherlands; 8) Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden; 9) Department of Clinical Sciences, Diabetes and Endocrinology, Lund University and Lund University Diabetes Centre, CRC at Skåne University Hospital, Malmö, Sweden; 10) Department of Evolutionary Biology, Genetic Section, University of Ferrara, Ferrara, Italy; 11) Department of Cardiovascular Sciences, University of Leicester, Leicester, UK; 12) National Institute for Health Research (NIHR) Leicester Cardiovascular Disease Biomedical Research Unit, Glenfield Hospital, Leicester, UK; 13) Institute of Genetic Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany; 14) deCODE Genetics, Reykjavik, Iceland; 15) Department of Genetic Epidemiology, Erasmus MC, Rotterdam, The Netherlands; 16) AG Kardiovaskuläre Genomik, Medizinische Klinik II, Universität zu Lübeck, Lübeck, Germany; 17) Public Health and Gender Studies, Institute of Epidemiology and Preventive Medicine, Regensburg University Medical Center, Regensburg, Germany.

Genome-wide association studies (GWAS) have been successful in identifying loci contributing to a range of complex human traits. However, there is still large proportion of unexplained trait heritability. Current large-scale association studies have mainly been targeting common variants due to the design of genome-wide genotyping arrays. Nevertheless, it is possible to gather information about rare variants by imputing markers up to high-density reference panels. We have tested for association between eight quantitative traits and accumulations of minor alleles at rare variants with minor allele frequency (MAF)<1% within gene regions (defined by the UCSC Human Genome database) using GRANVIL software. 14 cohorts with genome-wide association data contributing to the ENGAGE Consortium were imputed up to the cosmopolitan reference panel from the 1000 Genomes Project (June 2011 release). Altogether 17,098 men and 16,616 women were used in analyses for 8 traits: high density lipoprotein, low density lipoprotein, triglycerides, total cholesterol, body-mass index (BMI) and waist-to-hip ratio (WHR), fasting glucose and fasting insulin (FI) adjusted for BMI. Association summary statistics were combined across studies in a weighted Z-score based meta-analysis. We observed genome-wide significant evidence ($p<1.7x10^{-6}$, Bonferroni correction for 30,000 genes) for association with two traits. For triglycerides, associations with *ZNF259* ($p=1.4x10^{-11}$) and *APOA5* ($p=4.7x10^{-8}$) genes were detected. The excess of rare variants in *APOA5* has already been described for hypertriglyceridemia. Both, *ZNF259* and *APOA5*, are also associated with metabolic syndrome. For fasting insulin, a male-specific association was observed with *TAS2R60* ($p_{men}= 6.2x10^{-10}$; $p_{women}=0.22$, $p_{men+women}=0.00084$). This gene belongs to a family of taste receptors that may play a role in the perception of bitterness, and is gustducin-linked. Our results highlight the potential of imputation into exiting GWAS data up to high-density reference panels as a tool for discovering rare-variant associations with complex traits and an alternative to costly re-sequencing experiments.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2077T

Common genetic variants underlie endometriosis and fat distribution but not overall adiposity. *N. Rahmioglu[1], S. Macgregor[2], H.R. Harris[3], A. Morris[1], S.A. Missmer[3], G.W. Montgomery[2], C.M. Lindgren[1], K.T. Zondervan[1], The International EndoGene Consortium (IEC), The GIANT Consortium.* 1) Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom; 2) Queensland Institute for Medical Research, Herston, Queensland, Australia; 3) Brigham and Women's Hospital and Harvard Schools of Medicine and Public Health, Boston, MA, USA.

Lower body mass index (BMI) and waist-hip ratio (WHR) have been associated with increased risk of endometriosis in epidemiological studies. Both endometriosis and these obesity-related traits are heritable. The top endometriosis associated risk increasing SNP on chromosome 7, had a p-value of 4.4x10-5 in the GIANT consortium WHRadjBMI and marked one of 14 loci associated with lower WHRadjBMI. Therefore, we aimed to investigate the potential commonality in genetic variants underlying their aetiology through, (1) genetic enrichment analysis comparing three large-scale GWAS of endometriosis (3194 cases; 7060 controls; 12.5m SNPs imputed using 1000G), WHRadjBMI (n=77,167; 2.85m SNPs imputed using HapMap2) and BMI (n=123,865; 2.85m SNPs imputed using HapMap2), to identify potential pleiotropic regions, (2) polygenic prediction analysis, to evaluate the aggregate effect of WHR alleles on endometriosis risk, and (3) pathway analysis using the PANTHER database, to investigate potential common etiological pathways between enriched loci underlying endometriosis and WHR. The genetic enrichment analyses revealed significant enrichment of genetic variants between endometriosis and fat distribution (p-values for binomial tests 0.02–0.004) but not with overall adiposity (p-values 0.12–0.98). However, the directions of effect of the enriched variants associated with endometriosis and WHRadjBMI was not consistently in either the same or opposite direction (p=0.31–0.60). Some of the enriched variants include: (1) the endometriosis risk increasing allele at PPARG on chromosome 3, showing a strong association with higher WHRadjBMI, (2) the endometriosis risk increasing allele upstream of WNT4, showing an association with lower WHRadjBMI. Polygenic prediction revealed no en mass shared common genetic component between endometriosis and fat distribution. Pathway analysis generated some evidence for over-representation of developmental processes (p=1x10-4), WNT signaling (p=1.7x10-3) and cadherin signaling (p=1.2x10-3) pathways. We have identified support for potential pleiotropy underlying endometriosis and fat distribution, and identified several genetic loci with plausible biological candidacy. Even though the overall polygenic component of fat distribution does not contribute to the risk of endometriosis, there is evidence for the presence of some key pleiotropic loci contributing to the endometriosis risk and regulation of fat distribution, which require replication.

2078F

Genome-wide association of single nucleotide polymorphisms with weight loss outcomes following Roux-en-Y Gastric Bypass surgery. *E. Rinella[1], C. Still[2], Y. Shao[3], G.C. Wood[2], X. Chu[2], B. Salerno[2], G.S. Gerhard[2], H. Ostrer[4].* 1) New York University Langone Medical Center, New York, NY; 2) Geisinger Obesity Research Institute, Geisinger Clinic, Danville PA; 3) Division of Biostatistics, New York University School of Medicine, New York, NY; 4) Department of Pathology, Albert Einstein College of Medicine, Bronx, NY.

Objective: Bariatric surgery offers an effective treatment for obesity and obesity-related complications when other interventions fail. However, despite its success, many patients do not maintain sufficient weight loss following bariatric surgery. In order to better understand the long-term response to Roux-en-Y Gastric Bypass (RYGB) we designed a genome-wide association study (GWAS) to examine the genetic component of variable weight loss after RYGB. **Subjects:** 175 obese patients (BMI>35) receiving RYGB were selected for a Discovery cohort: 86 “Regainers” who lost the least excess body weight (EBW) and 89 “Retainers” who lost the most EBW at 2 years post-surgery and genotyped using Affymetrix 6.0 SNP arrays. A second cohort of 164 Regainers and 169 Retainers was selected to validate the Discovery cohort results using custom SNP arrays. **Results:** We identified 111 SNPs in the Discovery cohort whose frequencies were significantly different between Regainers and Retainers (allelic χ^2-test p<0.0001). Linear regression of %EBW remaining at 2 years post-surgery revealed 17 SNPs that were validated in the Replication cohort. These SNPs cluster in or near several genes with potential biological relevance including HTR1A, IGF1R, NMBR andPKHD1. **Conclusion:** We conclude that there are genetic components in addition to known clinical factors with utility in predicting weight loss outcome following RYGB.

2079W

Identification of common and specific genetic determinants to skin prick test reactivity using genome-wide association study and gene-based test approaches. *P-E. Sugier[1,2,3], A. Vaysse[1,2,3], C. Loucoubar[1,2,3], C. Sarnowski[1,2,3], P. Margaritte-Jeannin[1,2,3], M-H. Dizier[1,2,3], M. Lathrop[3], F. Demenais[1,2,3], E. Bouzigon[1,2,3], the EGEA collaborative group.* 1) U946, INSERM, Paris, France; 2) Université Paris Diderot, Sorbonne Paris Cité, Institut Universitaire d'Hématologie, France; 3) Fondation Jean Dausset-Centre d'Etude du Polymorphisme Humain (CEPH), France.

The prevalence of allergic diseases such as asthma has reached pandemic proportions in industrialized countries. The genetic component of allergy is substantial but it has been rarely investigated at the genome-wide level. Two genome-wide association studies (GWAS) of atopy defined by increased specific IgE levels or skin prick test (SPT) reactivity to allergens have been conducted in population-based cohorts with inconsistent results. We aimed to identify common and specific genetic determinants of SPT response to a panel of aeroallergens using a GWAS approach combining a single SNP analysis and gene-based association tests. We conducted GWAS of atopy phenotypes in 1,660 subjects (925 atopics and 735 non-atopics) from the French Epidemiological study on the Genetics and Environment of Asthma (EGEA) which includes families ascertained through asthmatics. These subjects were genotyped by Illumina 610K Array. Four atopy traits were investigated: 1) atopy as a whole and defined by a positive SPT response to at least one of 11 aeroallergens, and then three groups of aeroallergens: 2) indoors (Dermatophagoides pteronyssinus, cat, Blattella germanica), 3) outdoors (timothy grass, olive, birch, Parietaria judaica, ragweed), and 4) molds (Cladosporium herbarum, Aspergillus, Alternaria tenuis). For each phenotype, we used logistic regression for the single SNP analysis followed by gene-based association tests using VEGAS (Lui et al Am J Hum Genet 2010). Six atopy loci were detected at P-value$\leq 5x10^{-7}$ by single SNP analyses. One gene in 3q24 region was strongly associated to atopy as a whole ($P=3x10^{-7}$) and in a lesser extend to each SPT group (P ranging from $5x10^{-6}$ to $3x10^{-4}$), suggesting this locus harbors common genetic determinants shared across SPT responses to allergens. Five loci were mainly detected in sub-groups of allergens: 5q13 with indoors, 12q24 with outdoors and 4q24, 15q13 and 19q13 with molds. Gene-based tests both strengthened the evidence for association of 19q13 locus with SPT to molds and increased evidence for two other loci on chromosome 19 with this phenotype. Further investigation will be needed to confirm our findings in order to reliably identify key genetic factors underlying atopy predisposition. This study highlights that combining at the genome-wide level single SNP analyses and gene-based tests may facilitate the identification of new susceptibility genes. Funding: ANR-11-BSV1-027-01 grant.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2080T

A genome-wide meta-analysis of circulating levels of IGF-I and IGFBP-3. *A. Teumer[1], R.S. Vasan[2], M. Bidlingmaier[3], T. Tanaka[4], N.L. Glazer[2], M. Nethander[5], T. Harris[4], M. Beekman[6], R. Westendorp[6], E. Slagboom[6], J.I. Rotter[7], P.P. Pramstaller[8], Q. Sun[9], C. Zhang[10], C.S.P. Lam[2], M.O. Goodarzi[7], K. Rice[11], B.M. Psaty[12], G. Li[12], J.B. Meigs[13], L. Kuller[14], M.N. Pollak[15], C. van Duijn[16], A. Tönjes[17], S. Berndt[18], E. Ziv[19], X. Xue[20], C. Ohlsson[21], H. Wallaschofski[22], R.C. Kaplan[20] on behalf of the CHARGE Insulin-like growth Factor Working Group.* 1) Interfaculty Institute for Genetics and Functional, University of Greifswald, Greifswald, Germany; 2) Boston University School of Medicine, Boston MA USA; 3) Medizinische Klinik - Innenstadt, Ludwig-Maximilians University, Munich, Germany; 4) National Institute on Aging, Baltimore, MD USA; 5) Genomics Core Facility, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden; 6) Leiden University Medical Center, 2300 RC Leiden, The Netherlands; 7) Cedars-Sinai Medical Center, Los Angeles CA USA; 8) Center for Biomedicine, European Academy Bozen/Bolzano (EURAC), Bolzano, Italy, Department of Neurology, General Central Hospital, Bolzano, Italy, Department of Neurology, University of Lübeck, Lübeck, Germany; 9) Harvard School of Public Health; 10) the Eunice Kennedy Shriver National Institute for Child Health and Development, Bethesda MD USA; 11) Department of Biostatistics, University of Washington, Seattle WA USA; 12) Cardiovascular Health Research Unit, University of Washington, Seattle, WA USA; 13) Massachusetts General Hospital and Harvard Medical School, Boston MA USA; 14) University of Pittsburgh Graduate School of Public Health, Pittsburgh, PA USA; 15) Lady Davis Research Institute of Jewish General Hospital and McGill University, Montreal, Quebec, Canada; 16) Erasmus University Medical Center, the Netherlands; 17) University of Leipzig, Department of Medicine, Division of Endocrinology and Nephrology Liebigstraβe 20, Leipzig, Germany; 18) National Cancer Insitute, Bethesda MD USA; 19) Department of Medicine, University of California-San Francisco, USA; 20) Albert Einstein College of Medicine, Bronx NY USA; 21) Department of Internal Medicine, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden; 22) Institute of Clinical Chemistry and Laboratory Medicine, University Medicine Greifswald, Germany.

Insulin-like growth factor-I (IGF-I) and its predominant carrier IGF binding protein 3 (IGFBP-3) are implicated in a complex biological system important in diverse biological processes including cell replication, proliferation or differentiation, protein synthesis and carbohydrate and bone metabolism. IGF-I and IGFBP-3 serum concentrations are associated with cardiovascular disease, cancer and all-cause mortality. IGF-I and IGFBP-3 levels are known to have a high degree of heritability.

We conducted a genome-wide association study (GWAS) of circulating concentrations of IGF-I and IGFBP-3, in a meta-analysis of up to 22,000 middle-aged adults of European ancestry. Analyses were adjusted for age and sex, and sex-specific associations were also investigated. Because most circulating IGF-I is bound to IGFBP-3, the concentration of free circulating IGF-I may be estimated by adjusting blood IGF-I concentrations for the concentrations of circulating IGFBP-3. Therefore, we also performed GWAS of IGF-I concentrations after adjustment for IGFBP-3 concentrations, and vice versa.

In analysis of men and women together, we identified two new loci *GCKR* (rs780093) and *NUBP2* (rs1065656) where variants reached genome-wide significance ($p<5x10^{-8}$) for association with IGF-I, and one new locus *CELSR2* (rs12740374) where variants were associated with IGFBP-3 concentrations. In analysis of women only, variants at one additional locus *VPS13C*(rs12593844) were significantly associated with IGF-I adjusted for IGFBP-3. SNPs near *IGFBP3* showed effects on IGFBP-3 concentrations that differ by sex.

The findings revealed potential new loci which gene products are related to glucokinase and insulin-like growth factor binding. Furthermore, the associations of known variants with both IGF-I and IGFBP-3 concentrations near *IGFBP3* (rs11977526) and *TNS3* (rs700752) as well as the associations with IGFBP-3 concentrations in *SORCS2* (rs4234798) and in close vicinity of*NUBP2/IGFALS* (rs1065656) were confirmed.

2081F

Phenotypic refinement and genome-wide association analysis identifies a functional variant in ADCY3 associated with BMI. *N.J. Timpson[1], R. Gaillard[2], E. Stergiakouli[1], HR. Taal[2], DM. Evans[1], F. Rivadeneira[3], B. St Pourcain[5], AG. Uitterlinden[3], J. Kemp[1], A. Hofman[4], S. Ring[5], VWV. Jaddoe[2], G. Davey Smith[1].* 1) MRC CAiTE Centre, University of Bristol, Bristol, Bristol, United Kingdom; 2) Erasmus Medical Centre, Department of Epidemiology, Rotterdam, Netherlands; 3) Erasmus Medical Centre, Department of Internal Medicine and Epidemiology, Rotterdam, Netherlands; 4) Erasmus Medical Centre, Department Epidemiology & Biostatistics, Rotterdam, Netherlands; 5) School of Social and Community Medicine, University of Bristol, Bristol, Bristol, United Kingdom.

Genomewide association studies tend to be undertaken for recognised phenotypes based on either convention or available instrumentation. Body mass index (BMI; weight(kg)/height(m)2) is a routine measure of weight given height despite the known failure of the square of height to precisely account for this covariable. Work here aimed to re-assess the contribution of common genetic variation to BMI taking into account height. Genomewide association analysis (GWA) for BMI was undertaken in 6076 non-related participants from the Avon Longitudinal Study of Parents and Children (mean age 7.5 years) adjusting for height, age and sex. rs1172294 in the 3' UTR of RBJ was the strongest of a clear association signal on chr2, showing a 0.204kg/m2(0.133,0.274)p=1.1e-08 change per common allele. This signal contained rs11676272 (0.201kg/m2(0.132,0.270)p=1.4e-08), a missense SNP coding for a ser/pro change in exon1 of ADCY3. This association is larger then the established association at FTOrs9939609 for BMI at this age (0.128kg/m2(0.054,0.202)p=0.0007) and equivalent to that seen at later ages where the full effect of rs9939609 is realised. Results could be re-created when using a revised version of BMI based on the equation weight(kg)/height(m)3.05 which removed residual correlation with height in this population. In 2089 non-related participants from an independent study (Generation-R, average age 6.1 years) there was consistent evidence for association between rs11676272 and BMI adjusted for height (0.14kg/m2(0.056,0.224)p=0.0001; meta-analysis 0.177(0.123,0.230)p=7.6e–11, I217.4). Transformed lymphocyte derived analysis of the ADCY3 transcript showed evidence of association between the coding variant rs11676272 and levels of ADCY3 expression (r2=0.21, p<1x10-30). Furthermore, analyses conditional on rs11676272 obliterated both this signal and phenotypic association across this region of chr2. The ser/pro change coded by rs11676272 lies within the second transmembrane spanning alpha-helix of this membrane associated adenylate cyclase. It incorporates a helix-breaking residue into an important regulatory region of the enzyme and is likely to disrupt cAMP formation and triglyceride metabolism in adipose tissue. Together this provides evidence of a direct link between GWA results and a functional genetic change in ADCY3. Analyses serve as an important reminder as to the value of precise measurement in undertaking genetic association studies.

2082W

A genome-wide association study of kidney transplant survival: Donors recipients and interactions. *C. Franklin[1], M. Hernandez Fuentes[2], J. Mollon[2,3], I. Rebollo Mesa[2], E. Perucha[2], P. Conlon[4], N. Anyanwu[2], S.H. Sacks[2], M.E. Weale[2,3], N. Soranzo[1], G. Lord[2], United Kingdom and Ireland Renal Transplant Consortium, the Wellcome Trust Case-Control Consortium 3.* 1) Wellcome Trust Sanger Institute, Cambridge, Cambridgeshire, United Kingdom; 2) MRC Centre for Transplantation, King's College London, London, UK; 3) Department of Medical and Molecular Genetics, King's College London, London, UK; 4) Royal College of Surgeons in Ireland, Dublin, Ireland.

Approximately 1,600 kidney transplants from cadaveric donors are performed in the UK each year. There is a large variance in the survival times of these transplants with some grafts failing in the first few months, whereas around 70% survive 10 years or longer. Matching of donor and recipient for blood group and HLA types is used routinely to reduce the risk of rejection. However, the 100% acceptance rate for transplants between monozygotic twins indicates that other genetic influences on graft survival must exist.

We have genotyped samples from the UK and Ireland Renal Transplant Consortium collected from 24 transplant centres between 1981 and 2007. Samples were genotyped on Illumina 660 or 610 chips and after QC 512,535 SNPs were retained for analysis. Extensive post transplant phenotype information was collected on these samples through the NHS blood and transplant service and individual transplant centres. Genotype and phenotype data for a total of 2,678 transplant recipients, 2,200 donors and 2,081 complete transplant pairs was available.

We have conducted survival analysis of the time to graft failure using cox proportional hazards models to estimate the effect of each SNP. Effects were estimated for three models: recipient genotype, donor genotype and the interaction between donor and recipient genotype. Likelihood ratio tests were then used to estimate the significance of each effect. Several strong signals were found. The strongest was a recipient-only effect, but several weaker donor-only and interaction effects were also identified. None of the signals overlap with the HLA or ABO loci. This study represents the discovery phase of the first genome-wide association analysis of genetic effects on kidney transplant survival. Replication is on-going.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2083T

Admixture mapping of Vitamin D in African Americans. *I. Halder[1], L. Pearson[2], .S. Mulukutla[1], M. Shriver[2], V. Causer[1], G. Huggins[3], S. Reis[1].* 1) Cardiovascular Institute, Univ Pittsburgh, Pittsburgh, PA; 2) Dept. of Anthropology, Pennsylvania State University, State College, PA; 3) Tufts University Medical Center, Boston, MA.

Background: Vitamin D is crucial for a multitude of Cardiovascular and musculoskeletal functions. African Americans (AAs) generally have lower circulating levels of 25 hydroxyvitamin D [25(OH)D] than Whites, the cause for which is unknown. While darker skin pigmentation and dietary factors have been attributed to this difference, little is known about genetic determinants of 25(OH)D in AAs. Neither GWAS nor admixture mapping studies of 25(OH)D have been reported in AAs. We examined whether genetic ancestry was associated with 25(OH)D in AAs and performed genome wide admixture mapping. Methods: Circulating 25(OH)D was assessed by radioimmunoassay in 220 unrelated AAs from the HeartSCORE study. All individuals were genotyped for 1695 ancestry informative markers as part of a novel CVD gene chip. Individual admixture was ascertained by maximum likelihood method and admixture-phenotype associations were performed. The ADMIXMAP software was used for admixture mapping with age, gender, BMI, vitamin D supplementation status and season of blood draw as covariates. Results: Mean West African individual admixture in the sample was 69% (SD 16%). Covariate adjusted 25(OH)D levels showed a significant negative association with increasing individual African admixture (B = -12.5, SE = 3.7, t = -3.4, P = 0.001). Measured genetic ancestry explained ~ 5% of variation in vitamin D levels. In admixture mapping studies, significant excess European ancestry was observed in one region on chromosome 4 (4q), with largest excess seen at locus rs2545258 (Z score: -2.79, P = 9.47E-07); and one on chromosome 11 with largest excess seen at locus rs12289502 (Z score: -2.51, P = 6.61E-05). Excess West African ancestry was observed in one region on chromosome 22 with largest excess seen at locus rs1009787 (Z score: 2.3, P = 0.002). Conclusions: Genetic ancestry explains part of the variability in circulating 25(OH)D levels in AAs. Two loci associated with 25(OH)D in this study are near candidate genes identified in populations of European ancestry: *GC* (4q) and *CYP2R1* (11p). The SNP on chromosome 22 is in the *HIRA* gene which has been associated with the histone chaperone complex. *HIRA* in association with other cellular regulators acts as a transcriptional regulator of vitamin D receptors. In conclusion, we have identified three regions which appear to be associated with circulating 25(OH)D levels in AAs. We are in the process of confirming these results in other independent cohorts.

2084F

A genome-wide association study establishes multiple susceptibility loci for Sjögren's syndrome. *C.J. Lessard[1], H. Li[1], I. Adrianto[1], J.A. Ice[1], R. Jonsson[2], G.G. Illei[3], M. Rischmueller[4], G. Nordmark[5], X. Mariette[6], C. Miceli-Richard[6], M. Wahren-Herlenius[7], T. Witte[8], M. Brennan[9], R. Omdal[2], P.M. Gaffney[1], J.A. Lessard[10], W.-F. Ng[11], N. Rhodus[12], B. Segal[12], R.H. Scofield[1], J.A. James[1], J.-M. Anaya[13], J.B. Harley[14], C.G. Montgomery[1], K. Moser Sivils[1].* 1) Arthritis & Clinical Immunology, OMRF, Oklahoma City, OK; 2) University of Bergen, Norway; 3) National Institute of Dental and Craniofacial Research, NIH, Bethesda, MD; 4) The Queen Elizabeth Hospital and Health Service, Australia; 5) Uppsala University Hospital, Sweden; 6) Université Paris-Sud, France; 7) Karolinska Institute, Sweden; 8) Hannover Medical School, Germany; 9) Carolinas Medical Center, Charlotte, NC; 10) Valley Bone and Joint Clinic, Grand Forks, ND; 11) Newcastle University, England; 12) University of Minnesota, Minneapolis, MN; 13) Center for Autoimmune Diseases Research (CREA), Universidad del Rosario, Colombia; 14) Cincinnati Children's Hospital Medical Center, Cincinnati, OH.

Sjögren's syndrome (SS) is a common (0.7–1.0% of the European population), clinically heterogeneous autoimmune disease characterized by exocrine gland dysfunction that involves both innate and adaptive immune responses. A complex genetic architecture has been hypothesized; however, genetic studies to date have been limited to candidate gene approaches. We used high-density Illumina OMNI1-Quad genotyping arrays in a discovery cohort of 424 European-derived SS cases and 2120 healthy controls to perform the first genome-wide association scan (GWAS) in an unbiased manner to identify SS susceptibility loci. Stringent quality control (QC) criteria, adjustments for population stratification, and standard GWA statistical methodologies were used to compare allele frequencies between cases and controls. A total of ~650,000 single nucleotide polymorphisms (SNPs) were tested for association to SS (P_{omni}). For replication, an independent cohort of 1194 SS cases and 2930 healthy controls were genotyped using the ImmunoChip (IC; P_{IC}) with ~26,000 overlapping SNPs with the GWAS after QC. Meta-analysis between the GWAS and IC was done using METAL (P_{meta}). The most significantly associated region with risk of SS was the major histocompatibility complex (MHC), with 1071 overlapping SNPs exceeding a genome-wide significance (GWS) threshold of 5x10E-8, and the peak association was observed in MSH5 (rs3117574 P_{meta}=5.33x10E-79). Additional results across the extended MHC support association with multiple loci throughout this region. In the GWAS, two SNPs, rs485497 and rs4680536, were identified near IL12A with P_{omni}<6x10E-4. Both rs485497 (P_{IC}=1.88x10E-7) and rs4680536 (P_{IC}=2.06x10E-5) replicated yielding P_{meta}=4.81x10E-10 and P_{meta}=1.69x10E-8, respectively. In addition, we observed associations surpassing GWS for the first time with loci previously implicated in SS, including IRF5 (rs10488631 P_{meta}=5.25x10E-13), BLK (rs922483 P_{meta}=1.50x10E-8), and STAT4 (rs10168266 P_{meta}=3.59x10E-8). IL12A, IRF5 and STAT4 are involved in type 1 interferon responses. IL12A encodes the p35 subunit of IL12 and is secreted by monocytes and dendritic cells ultimately stimulating the production of IFN-γ. Interestingly, responses to IL12A are mediated through STAT4. We present the first GWAS of SS identifying and confirming IL12A as a novel susceptibility locus. We also observed IRF5, BLK and STAT4 for the first time at GWS confirming them as risk loci for SS.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2085W

Pooling-based genome-wide association study for intracranial aneurysms in the Portuguese population. *P.C. S. Abrantes[1,2,3], I. Sousa[1,2], M. M. Santos[4], V. Francisco[1,2], T. Krug[1,2], J. M. Xavier[1,2], A. Jacinto[3], D. Coiteiro[4], S. A. Oliveira[1,2].* 1) Instituto Medicina Molecular, Lisbon, Portugal; 2) Instituto Gulbenkian de Ciência, Oeiras, Portugal; 3) Centro de estudos de Doenças Crónicas, Lisbon, Portugal; 4) Hospital de Santa Maria, Lisbon, Portugal.

Subarachnoid hemorrhage is a life-threatening event that most frequently leads to severe disability and death. Its most frequent cause is the rupture of an intracranial aneurysm (IA), which is a dilation of a blood vessel caused by disease or weakening of the vessel wall. Although the genetic contribution to IA is well established, to date no single gene has been unequivocally identified as responsible for IA formation or rupture. The main goal of this work was to identify for the first time new IA susceptibility genes in the Portuguese population. To this end, we conducted a pooling-based GWAS in 100 IA cases and 92 controls matched for age and gender. Pool-based GWAS is an efficient strategy to reduce the high costs associated with individual-based GWAS but is also effective at identifying major genetic contributions to disease that could be lost in individual underpowered studies. In the present study, 2 pools of cases and 2 pools of controls were genotyped in triplicate using the Affymetrix Human SNP Array 6.0. Probe intensities were analyzed using the SNPMaP package and the R freeware. Two methods were used to select variants that increase IA risk: a) a Welch t-test between cases and controls and b) the absolute differences of averaged normalized RAS (Relative Allele Score) between cases and controls. Top 50 hits from the Welch t-test and RAS difference approach, as well as the convergent SNPs from both methods, were analyzed. We replicated the 9p21.3 region found in previous IA GWAS, and identified 29 novel loci associated with IA (P-value $<5x10^{-8}$ or RAS difference $\geq$ 15%). Moreover, we found 6 putatively loci that were convergent in both analyses and due to their biological function could have a role in IA etiology. With a better understanding of the contribution of causative variants in the IA Portuguese population, we expect to contribute to unravel the genetic etiology of IA.

2086T

Genetic variants at the IGHC locus are associated with IgG levels in multiple sclerosis patients. *E. Albrecht[1], D. Buck[2], M. Aslam[2], A. Goris[3], N. Hauenstein[2], A. Jochim[2], S. Cepok[2], V. Grummel[2], B. Dubois[3], A. Berthele[2], P. Lichtner[4,5], C. Gieger[1], J. Winkelmann[2,4,5], B. Hemmer[2].* 1) Institute of Genetic Epidemiology, Helmholtz Center Munich - German Research Center for Environmental Health, Neuherberg, Germany; 2) Department of Neurology, Technische Universität München, Munich, Germany; 3) Laboratory for Neuroimmunology, Section of Experimental Neurology, KU Leuven, Leuven, Belgium; 4) Institute of Human Genetics, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany; 5) Institute of Human Genetics, Technische Universität München, Munich, Germany.

Multiple Sclerosis (MS) is considered a chronic inflammatory disease of the central nervous system of autoimmune origin involving T and B cells. Intrathecal IgG synthesis is observed in the majority of patients with MS. Whereas the amount of intrathecal IgG synthesis varies largely between patients, intrathecal IgG remains rather constant in the individual patient. Based on this observation it seems reasonable to assume that genetic factors may impact on the extent of intrathecal IgG synthesis. To investigate the genetic determinants of intrathecal IgG synthesis in MS patients, we performed a genome-wide association study (GWAS) based on 526,014 SNPs of the Human660-Quad chip in 229 MS patients. For replication, we genotyped 18 SNPs, showing an association with intrathecal IgG synthesis (p<1E-5), in an independent validation sample of 256 MS patients using Sequenom. Five of the 18 SNPs, which could be replicated in the first validation sample, were additionally analyzed in a second validation sample containing 152 MS patients genotyped on the Illumina Human660-Quad chip. Patients of all three samples are of European descent. We identified five SNPs showing a significant association with intrathecal IgG synthesis in the discovery and replication samples. These SNPs are in high LD and are located on the IGHC locus on chromosome 14. The strongest effect is seen for rs10136766 with a combined p-value of 7.5E-16 and 8.9% explained variance. The results of this study suggest that intrathecal IgG synthesis in MS patients is genetically influenced by variants in the IGHC gene cluster, with a remarkably high proportion of explained variance.

2087F

A web-based initiative to accelerate research on genetics and disease in African Americans. *K.E. Barnholt[1], A.K. Kiefer[1], H.L. Gates, Jr.[2], M. Nelson[1], M. Mullins[1], E. Baker[3], J. Frank[1], C.D. Bustamante[4], T.W. Love[5], R.A. Kittles[6], N. Eriksson[1], J.L. Mountain[1].* 1) 23andMe, Inc., Mountain View, CA; 2) W.E.B. Du Bois Institute for African and African American Studies, Harvard University, Cambridge, MA; 3) 23andYou.com; 4) Department of Genetics, Stanford University School of Medicine, Stanford, CA; 5) Onyx Pharmaceuticals, Inc., South San Francisco, CA; 6) College of Medicine, University of Illinois at Chicago, Chicago, IL.

Little is known about the connections between DNA and disease in African Americans, in part because most genetics research has involved only those of European ancestry. Greater understanding of such connections could improve diagnoses and lead to opportunities for more personalized health care. In 2011 23andMe, Inc., a personal genomics and research company, launched the Roots into the Future initiative, which aims to enroll 10,000 African Americans in an innovative research project. The study seeks to determine whether genetic associations previously identified in Europeans are relevant to African Americans and to discover other genetic markers linked to conditions of particular relevance to the African American community. Currently the 23andMe cohort includes nearly 10,000 African Americans, over 5700 of whom were recruited through the Roots into the Future initiative. Each of these individuals (58% female, 42% male; mean age: 44) has submitted a saliva sample for genotyping via 23andMe's custom genotyping array, which includes approximately 1 million single nucleotide polymorphisms. Participants are currently contributing information about their health and traits through online surveys. To date over 6200 participants have completed an average of 10.6 surveys. Using the genetic data we estimated the percent African and European ancestry of each participant. Median estimates were 73% and 23% respectively (with 4% uncertain). As expected, the higher a person's proportion of European ancestry, the greater the chance that person carries variants that are more common among Europeans than among Africans, such as those linked to HIV-resistance and alpha-1 antitrypsin deficiency. Furthermore, the higher a person's proportion of African ancestry, the more likely that person reported having curly hair, high blood pressure and type 2 diabetes, and the less likely that person reported having facial wrinkles, rosacea and Parkinson's Disease. Based on data for over 8700 individuals likely to self-identify as African American, we replicated over 25 genetic associations reported previously for African Americans, including those for body-mass index, type 2 diabetes, lupus, height, and osteoporosis. For conditions for which we have already accrued at least 500 cases among this cohort, such as asthma, migraines, and uterine fibroids, we anticipate having power either to replicate associations identified through previous studies of Europeans or to find new associations.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2088W

First systematic association study of achalasia points to a strong involvement of the *HLA* region in the disease process. *J. Becker[1,2], M. Knapp[3], M.M. Wouters[4], G. Trynka[5], V. Kumar[5], L. Franke[5], H.-J. Westra[5], C. Wijmenga[5], G.E. Boeckxstaens[4], M.M. Nöthen[1,2], I. Gockel[6], J. Schumacher[1].* 1) Institute of Human Genetics, University of Bonn, Bonn, Germany; 2) Department of Genomics, Life & Brain Center, University of Bonn, Bonn, Germany; 3) Institute for Medical Biometry, Informatics, and Epidemiology, University of Bonn, Bonn, Germany; 4) Translational Research Center for Gastrointestinal Disorders, Catholic University of Leuven, Leuven, Belgium; 5) Genetics Department, University Medical Center and University of Groningen, Groningen, The Netherlands; 6) Department of General and Abdominal Surgery, University of Mainz, Mainz, Germany.

Achalasia is a severe disorder of the lower esophageal sphincter (LES) characterized by degeneration of neurons in the myenteric plexus. This leads to the development of a megaesophagus with irreversible loss of LES function. Furthermore, the retention of food can result in aspiration pneumonia and esophagitis, which is a precancerous condition leading to esophageal carcinoma. Although the cause of achalasia is mainly unknown, autoimmune processes appear to be involved in individuals with a genetic susceptibility. The purpose of the present study was to identify genes contributing to achalasia. We therefore performed the first systematic association study using Illumina's Immunochips, which capture genetic variants at immune-relevant loci. To the best of our knowledge, we analyzed the largest achalasia sample that has been studied so far. Our sample consisted of 633 cases and 2,653 controls from Central Europe. After quality control 126,899 markers remained for association testing. The analysis yielded 24 markers reaching genome-wide significance (P<5x10^{-08}) which are all located within the *HLA* region on chromosome 6p. The most significant marker reached a P-value of 7.53x10^{-15} and a relative risk (RR) of 2.84. Of all markers, a total of four association signals are contributing independently from each other to achalasia according to a stepwise logistic regression analysis. The markers are located within or near the genes *HLA-DPA1*, *HLA-DQA2*, *HLA-DQB2*, and *MUC21*. In order to investigate whether the associated variants have an effect on gene expression, we analyzed one of the largest available expression quantitative trait loci (eQTL) data sets (Fehrmann *et al*., 2011 PLoS Genetics). This study indicates that achalasia risk variants represent significant eQTLs for transcripts that are functionally involved in immunological processes. Finally, we performed an interaction analysis using all 126,899 markers and observed one marker-pair that showed epistatic association reaching the threshold of Bonferroni-corrected significance (P=4.67x10^{-12}, RR=5.47). Both markers are located upstream of the genes *HLA-DQA1* and *HLA-DQB1*, respectively. Interestingly, the corresponding proteins are also interacting on the cellular level by forming the HLA-II molecule on antigen-presenting cells. In summary, we have identified the first risk genes for achalasia and our data emphasize the hypothesis that autoimmune-relevant processes are involved in the disease process.

2089T

Genetic variants associated with breast size also influence breast cancer risk. *G. Benton, N. Eriksson, C.B. Do, A.K. Kiefer, J.L. Mountain, D.A. Hinds, U. Francke, J.Y. Tung.*23andMe, Mountain View, CA.

While some factors of breast morphology, such as density, are directly implicated in breast cancer, the relationship between breast size and cancer is less clear. Breast size is moderately heritable, yet the genetic variants leading to differences in breast size have not been identified. To investigate the genetic factors underlying breast size, we conducted a genome-wide association study (GWAS) of self-reported bra cup size, controlling for age, genetic ancestry, breast surgeries, pregnancy history and bra band size, in a cohort of 16,175 women of European ancestry.

We identified seven single-nucleotide polymorphisms (SNPs) significantly associated with breast size ($p < 510^{-8}$): rs7816345 near ZNF703, rs4849887 and (independently) rs17625845 flanking INHBB, rs12173570 near ESR1, rs7089814 in ZNF365, rs12371778 near PTHLH, and rs62314947 near AREG. Two of these seven SNPs are in linkage disequilibrium (LD) with SNPs associated with breast cancer (those near ESR1 and PTHLH), and a third (ZNF365) is near, but not in LD with, a breast cancer SNP. The other three loci (ZNF703, INHBB, and AREG) have strong links to breast cancer, estrogen regulation, and breast development.

These results provide insight into the genetic factors underlying normal breast development and show that some of these factors are shared with breast cancer. While these results do not directly support any possible epidemiological relationships between breast size and cancer, this study may contribute to a better understanding of the subtle interactions between breast morphology and breast cancer risk.

2090F

Meta-analysis of genetic associations in up to 339,224 individuals identify 67 new loci for BMI, confirming a neuronal contribution to body weight regulation and implicating several novel pathways. *S.I. Berndt[1], S. Vedantam[2,3,4], F. Day[5], S. Gustafsson[6], A.E. Locke[7], C. Powell[8], B. Kahali[8], D.C. Croteau-Chonka[9], T.W. Winkler[10], A. Scherag[11], I. Barroso[12,13], J.S. Beckmann[14,15], A. Justice[16], C.M. Lindgren[17], T. Pers[4], P. Visscher[18], J. Yang[18], M. Boehnke[7], G. Abecasis[7], C.J. Willer[7], K.I. Mohlke[9], K.E. North[16], J.N. Hirschhorn[2,3,4], E. Ingelsson[6], R.J.F. Loos[5,19], E.K. Speliotes[8,20] for the GIANT Consortium.* 1) Division of Cancer Epidemiology & Genetics, National Cancer Institute, Bethesda, Maryland 20852, USA; 2) Divisions of Genetics and Endocrinology and Program in Genomics, Children's Hospital, Boston, Massachusetts 02115, USA; 3) Metabolism Initiative and Program in Medical and Population Genetics, Broad Institute, Cambridge, Massachusetts 02142, USA; 4) Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA; 5) MRC Epidemiology Unit, Institute of Metabolic Science, Cambridge, UK; 6) Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, 171 77 Stockholm, Sweden; 7) Department of Biostatistics, Center for Statistical Genetics, University of Michigan, Ann Arbor, Michigan 48109, USA; 8) Center for Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, Michigan, USA; 9) Department of Genetics, University of North Carolina, Chapel Hill, North Carolina 27599, USA; 10) Public Health and Gender Studies, Institute of Epidemiology and Preventive Medicine, Regensburg University Medical Center, Regensburg, Germany; 11) Institute for Medical Informatics, Biometry and Epidemiology, University of Duisburg-Essen, Germany; 12) Wellcome Trust Sanger Institute, Hinxton, Cambridge, CB10 1SA, UK; 13) University of Cambridge Metabolic Research Labs, Institute of Metabolic Science Addenbrooke's Hospital, CB2 OQQ, Cambridge, UK; 14) Service of Medical Genetics, Centre Hospitalier Universitaire Vaudois (CHUV) University Hospital, 1011 Lausanne, Switzerland; 15) Department of Medical Genetics, University of Lausanne, 1005 Lausanne, Switzerland; 16) Department of Epidemiology, School of Public Health, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27514, USA; 17) Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, OX3 7BN, UK; 18) University of Queensland Diamantina Institute, University of Queensland, Princess Alexandra Hospital, Brisbane, Queensland, Australia; 19) Mount Sinai School of Medicine, New York, NY 10029, USA; 20) Department of Internal Medicine, Division of Gastroenterology, University of Michigan, Ann Arbor, Michigan, USA.

Large-scale genome-wide association studies (GWAS) have identified > 50 loci unequivocally associated with obesity-susceptibility traits, including body mass index (BMI), waist-hip ratio, percent body fat, and extreme obesity. Of these, 32 loci were reported to be associated with BMI in our most recent meta-analysis, including 123,865 individuals at the discovery stage and 125,931 individuals at the follow up stage. To identify additional genetic loci for overall adiposity, we expanded the GIANT Consortium to include a total of 236,231 individuals from 82 GWAS as well as 103,046 individuals from 43 studies that were genotyped with the Metabochip, a custom-designed array comprised of SNPs with prior evidence of suggestive associations with metabolic traits. Summary statistics of the study-specific BMI-SNP association analyses assuming an additive genetic model were combined using a fixed-effects inverse variance meta-analysis approach. We confirmed 31 established BMI loci and identified 67 new independent loci (separated by at least 500 kb) associated with BMI ($P < 5x10^{-8}$). Joint conditional analyses from summary statistics identified additional independently associated common variants ($P < 5x10^{-8}$ for the secondary variant) at five loci (*MC4R*, *BDNF*, *GPRC5B*, *FANCL*, and *ADCY9*); analysis of rare variants at these loci may reveal additional independent genetic variants associated with BMI and help elucidate the nature of important functional units at each locus. Furthermore, we observed that among 1,929 independent variants with $P < 0.0001$ in our original GWAS and now included on the Metabochip, 1,608 demonstrated directionally consistent association signals, suggesting there are likely ~1,246 additional loci with BMI-associated common variants. Consistent with previous GWAS, many of the genome-wide significant BMI loci contain genes that have a potential neuronal role (e.g., *GRID1*, *NAV1*). However, this extended analysis also reveals loci near genes in pathways that were previously less apparent, such as those related to glucose and insulin homeostasis (e.g., *TCF7L2*, *GIPR*, *ADPGK*), mitochondrial processes (e.g., *MTCH2*, *MTIF3*, *PDK4*), lipid metabolism (e.g., *APO*-cluster, *NPC1*), and hormone regulation (*HSD17B12*). This large-scale meta-analysis has more than doubled the number of identified obesity-susceptibility loci and implicated additional pathways that further elucidate the complex biology of adiposity regulation.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2091W

Selective Immunoglobulin A Deficiency is Associated with IFIH1, TNFAIP3, PVT1, FAS, CDH23 and TM7SF3 in a Genomewide Association Study and Meta-Analysis. P.*G. Bronson[1], A.P. Manoharan[1], T.R. Bhangale[1], W. Ortmann[1], R.C. Ferreira[2], Q. Pan-Hammarström[3], L. Hammarström[3], R.R. Graham[1], T.W. Behrens[1].* 1) Genentech, Inc, South San Francisco, CA; 2) Cambridge Institute for Medical Research, Cambridge, UK; 3) Karolinska Institutet, Stockholm, Sweden.

The etiology of selective immunoglobulin A deficiency (IgAD) is unknown but a strong genetic component has been established. The most consistent associations are with HLA genes. We previously conducted a GWAS that tested 286k genotyped and 1.8M imputed variants in 430 cases and 1,090 controls (N=1,520) and a meta-analysis of three samples with 772 cases and 1,976 controls (N=2,748). IFIH1 rs1990760*T, a missense mutation associated with type 1 diabetes (T1D) and systemic lupus erythematosus (SLE), was associated with IgAD. The current GWAS tested 505k genotyped and 6.3M imputed variants (529 cases; 6,964 controls; N=7,493) and conducted a meta-analysis of four independent samples (1,289 cases; 8,932 controls; N=10,181). IFIH1 rs1990760*T was over-represented in cases (0.70) vs. controls (0.62) (OR=1.4, P=$3x10^{-14}$). In addition, 5 novel non-HLA loci were significant: TNFAIP3, CDH23, TM7SF3, FAS and PVT1. TNFAIP3 rs2230926*G, a missense mutation associated with SLE and rheumatoid arthritis (RA) in the opposite direction, was under-represented in cases (0.013) vs. controls (0.035) (OR=0.5, P=$4x10^{-9}$). CDH23 rs4747179*T was under-represented in cases (0.17) vs. controls (0.22) (OR=0.7, P=$1x10^{-8}$). Rs77540050*T (near TM7SF3) was under-represented in cases (0.003) vs. controls (0.013) (OR=0.3, P=$1x10^{-8}$). FAS rs2234978*T was over-represented in cases (0.37) vs. controls (0.32) (OR=1.3, P=$2x10^{-8}$). Rs72722767*A (near PVT1) was under-represented in cases (0.20) vs. controls (0.25) (OR=0.8, P=$5x10^{-8}$). Of 48 loci with suggestive results in the previous GWAS (P<$1x10^{-4}$), 6 loci were still suggestive in this GWAS (AHI1, CLEC16A, ADAM29, CD86, BCL6 and ODZ4). Over half of these loci are associated with autoimmunity. FAS is associated with a rare autoimmune lymphoproliferative syndrome. PVT1 and AHI1 are associated with multiple sclerosis (MS). CLEC16A is associated with MS and T1D. There is suggestive evidence that SLE is associated with CD86. The observed vs. expected numbers of low P-values in 278 independent autoimmune variants were examined. The expected number of P<$1x10^{-5}$ was 0.00278 and 18 were observed; the expected number of P<0.01 was 2.78 and 47 were observed (P<$1x10^{-5}$). In conclusion, results provide strong evidence for association with IFIH1, TNFAIP3, FAS, CDH23, PVT1 and TM7SF3, suggest that the AHI1, CLEC16A, ADAM29, CD86, BCL6 and ODZ4 may be involved, and provide additional support for the enrichment of association between IgAD and autoimmune disease loci.

2092T

Three new susceptibility loci for hyperuricemia identified through a genome-wide association analysis in Han Chinese. *C.-H. Chen, Y.-T. Chen, Y.-T. Chen, J.-Y. Wu.* Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan.

Background: Previous studies showed genetic variants contributed to hyperuricemia in various populations.

Objective: To investigate the genetic components of hyperuricemia, this study aimed to identify genes that might affect the level of uric acid in the blood in the Han-Chinese population.

Methods: DAN samples of a total of 1,947 unrelated subjects were extracted from a population-based study in the Taiwan. All subjects were of Han-Chinese ancestry and genotyped using the Axiom Genome-wide CHB Array. After excluding subjects with cryptic relatedness and population outliers, 471 subjects were classified into the hyperuricemia group for their levels of uric acid were equal to or greater than 7 mg/dL and 1,410 for the control group. Genome-wide association analysis was carried out using logistic regression models, adjust for age, sex and BMI.

Results: Several SNPs were found to be associated with hyperuricemia (P < 10^{-6}) and located in or near three gene loci, which were not previously reported to be susceptibility to hyperuricemia. In addition, several SNPs clustering in ABCG2, a previously reported major regulator of hyperuricaemia, showed nominal signals (P < 10^{-5}). SNPs in SLC2A9, another reported major regulator, however, did not show any significance in this study (P > 10^{-4}).

Summary: Our results identified 3 novel loci and 1 reported locus susceptible to hyperuricemia. Our findings might lead to a better understanding of differences in the molecular pathogenesis of hyperuricemia among various populations.

2093F

Candidate genetic loci for telomere length: A family based association study of Long Life Family Study (LLFS). *R. Cheng[1,2], J.H. Lee[1,2,3], M.S. Kang[1], M. Province[5], R. Mayeux[1,2,3,4], L.S. Honig[1,2,4].* 1) Taub Institute for Research on Alzheimer's Disease and the Aging Brain, Columbia University, New York, NY, USA; 2) Sergievsky Center, Columbia Univ, New York, NY, USA; 3) Department of Epidemiology, School of Public Health, Columbia Univ, New York, NY, USA; 4) Departments of Neurology and Psychiatry, P&S, Columbia Univ, New York, NY, USA; 5) Div. of Statistical Genomics, Dep. of Gen., Washington U. School of Med., St. Louis, MO, USA.

Shortening of chromosomal telomeres is a consequence of cell division, and is a biological factor that relates to cellular aging. Since shorter telomeres may relate to more rapid biological aging, we hypothesize that genetic and environmental factors contribute to shorter telomere length. To identify genetic factors that contribute to telomere length, we examined families participating in the Long Life Family Study (LLFS), which includes primarily white families from four different sites, including New York, Boston, Pittsburgh, and Denmark. To identify genetic factors that contribute to telomere length, we performed a family based genome wide association. For this purpose, we examined 4163 individuals from 586 families that are part of the Long Life Family Study (LLFS). Using 2.5M SNPs from the Illumina HumanOmni2.5, we applied an additive genetic model with linear mixed effects while taking into account kinship. The model adjusted for potential confounders, including age, sex, site, PCAs, a history of heart disease, smoking and drinking, and marriage. Heritability of telomere length was estimated to be 54.1% (SE=3.5%). One SNP (rs7680468) on chromosome 4 reached a genome-wide significant level (p=4.7x10E-8). Multiple SNPs had p-values less than 1x10E-5. Among them, two SNPs on chromosome 14 were located within the SYT16 gene. Other SNPs were located in TRDMT1, TMPRSS7, TRDMT1, ASCC2, and TSHZ2. In the present family based genome wide association study, we identified several candidate loci that are associated with telomere length. These loci are currently being confirmed in two independent datasets.

2094W

Common variations at chromosome 21q22 influences the risk of age-related nuclear cataract: The Singapore Malay Eye Study. *C.Y. Cheng[1,2,3], J. Liao[1], X. Li[3], C.C. Khor[2,4], W.T. Tay[3], A.G. Tan[5], J.J. Wang[5,6], P. Mitchell[5], Y.Y. Teo[2,7,8,9], E.S. Tai[2,10], S.M. Saw[2,3], T. Aung[1,3,11], T.Y. Wong[1,3,6,11].* 1) Dept of Ophthalmology, National University of Singapore, Singapore, Singapore; 2) Saw Swee Hock School of Public Health, National University of Singapore; 3) Singapore Eye Research Institute, Singapore; 4) Human Genetics, Genome Institute of Singapore; 5) Centre for Vision Research, Department of Ophthalmology, University of Sydney, Australia; 6) Centre for Eye Research Australia, University of Melbourne, Australia; 7) Centre for Molecular Epidemiology, National University of Singapore, Singapore; 8) Dept of Statistics and Applied Probability, National University of Singapore, Singapore; 9) Graduate School for Integrative Science and Engineering, National University of Singapore, Singapore; 10) Dept of Medicine, National University of Singapore; 11) Singapore National Eye Centre, Singapore.

Age-related cataract is the most common cause of blindness in the world. Although several mutations have been identified as causative for congenital and juvenile cataracts, little is known about the genes that influence the susceptibility of age-related cataract. We conducted a genome-wide association study on 2,259 individuals aged 40 years and older enrolled in the Singaporean Malay Eye Study (SiMES). Participants underwent a comprehensive eye examination including slit-lamp lens photography and provided a blood sample for genotyping. Lens photographs were graded using the Wisconsin Cataract Grading System based on 1 to 5 scale. Genotyping was performed with Illumina's 610K Quad Bead Chips. After standard quality control filtering, the association of nuclear cataract and 55,7824 genotyped SNPs was tested using ordered logistic regression analysis with adjustment for age, sex, and population substructure. The strongest evidence was observed at chromosome 21q22 (OR=1.48, *P*=$1.81x10^{-8}$), reaching genome-wide significance. The results were similar after further adjustment for smoking status. An additional SNP at 10q26 was suggestively associated with age-related nuclear cataract (*P*<$5x10^{-7}$). Further replication tests in other cohorts are needed to confirm these associations.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2095T

A genome-wide association study on progressive and bout-onset multiple sclerosis patients. *F. Clarelli[1], P. Brambilla[1], F. Esposito[1], G. Giacalone[1], M. Rodegher[1], M. Sorosina[1], C. Guaschino[1], N. Barizzone[5], P. Cavalla[3], E. Patti[2], D. Galimberti[4], E. Scarpini[4], S. Lupoli[8], R. Capra[6], G. Tedeschi[11], G. Mancardi[9], G. Coniglio[7], L. Grimaldi[10], A. Ghezzi[12], D. Cusi[8], V. Martinelli[1], M. Leone[5], S. D'Alfonso[5], G.C. Comi[1], F. Martinelli Boneschi[1].* 1) Department of Neurology, Scientific Institute and University Hospital San Raffaele, Milan, Italy; 2) Department of Neurological Sciences, University of Catania, Italy; 3) Department. of Neurology, University of Turin, Italy; 4) IRCCS Ospedale Maggiore Policlinico, Milan, Italy; 5) Department of Neurology, Ospedale Maggiore della Carità, Novara, Italy; 6) Department of Neurology, Spedali Civili, Brescia, Italy; 7) Department of Neurology, Madonna delle Grazie Hospital, Matera, Italy; 8) Department of Health Sciences, University of Milan, Italy; 9) Department of Neurological Sciences and Neurorehabilitation - University of Genova, Italy; 10) San Raffaele Hospital Giglio, Cefalù, Italy; 11) Department of Neurological Sciences, Second University of Naples, Italy; 12) Department of Neurology, S. Antonio Abate Hospital, Gallarate, Italy.

Background The role played by genetic factors in distinguishing between progressive multiple sclerosis (PrMS) and bout-onset multiple sclerosis (BOMS) is not yet well estabilished.**Objective** To identify genetic risk modifiers loci which could drive the risk of developing a progressive-onset versus a bout-onset course of multiple sclerosis. **Methods** We conducted a genome-wide association study (GWAS) in 530 patients affected by primary progressive (PPMS) and progressive-relapsing (PRMS) patients recruited in the context of a multicentric project conducted in continental Italy and compared the genetic profiling of these patients with 540 patients affected by a relapsing-remitting (RRMS) and secondary progressive (SPMS) multiple sclerosis of continental italian origin. Samples were genotyped on two different Illumina platforms (OmniExpress 700k, OmniQuad 610k), so only common markers across chips have been taken into account. **Results**. Results highlight the existence of a signal of weak association on chromosome 9p22.2 ($p=3.09 * 10^{-6}$),which needs to be validated and confirmed. While there are no evidences of any signal from the HLA region, supporting the knowledge that the contribution exerted by this region is independent of disease course. Further statistical analyses are ongoing to stratify results across platform type and other covariates. Other association studies are going to be conducted: genetic profiles of PrMS patients and Italian healthy controls are going to be compared, and the relation between genetic loci and PrMS progression index is going to be investigated. **Conclusions** This study represents one of the largest cohort study performed on PPMS. Results still need to be finalized to confirm or not the conclusion that the difference between disease courses can't be explained by genetic factors.

2096F

Novel loci for caffeinated coffee consumption revealed by a genome-wide meta-analysis of 91,000 individuals. *M. Cornelis on behalf of the Coffee & Caffeine Genetics Consortium.* Nutrition, Harvard School of Public Health, Boston, MA.

Background: Coffee, a major dietary source of caffeine, is one of the most widely consumed beverages in the world and has been implicated in various health conditions. We previously identified variants near *AHR* and *CYP1A2* associated with habitual caffeine and coffee intake. To identify additional loci, we now focus specifically on coffee consumption in an expanded sample. Methods: We included 28 population-based studies of European ancestry with self-reported intake of predominantly caffeinated coffee and genotype data imputed to ~2.5 million SNPs. A genome-wide (GW) analysis was performed for association of SNP allele dose with coffee consumption (cups/day) among coffee drinkers adjusting for age, sex, and smoking within each study. A secondary GW analysis was performed comparing high coffee consumers to none/low consumers. For both phenotypes, genetic effect estimates across studies were combined using an inverse-variance fixed-effects approach after applying quality control filters to the GW results. Results: Among ~91,000 coffee drinkers, coffee intake was associated at GW-significance (p<5E-8) with SNPs near *AHR* (7p21), *CYP1A2* (15q24), *POR* (7q11.2) and *MLXIPL* (7q11.23). SNPs near *BDNF* (11p13) reached GW-significance in the secondary analysis. These 5 loci have effect sizes of 0.05 to 0.15 cups/day per allele and together explain ~1.1% of the phenotypic variance of coffee intake. POR encodes an electron transfer protein that catalyzes the activity of CYP450 enzymes and is plausibly linked to caffeine metabolism. Brain-derived neurotrophic factor (BDNF) regulates synaptic transmission and plasticity in the central nervous system and thus may facilitate the psychostimulant effects of caffeine. The role of *MLXIPL* in coffee drinking behavior is unknown. SNPs associated with increased coffee consumption have previously been associated with smoking initiation (*BDNF*), higher adiposity (*BDNF*) but lower blood pressure (*CYP1A2*), and favorable lipid, inflammatory and liver enzyme profiles (*MLXIPL*)(genome.gov/gwastudies). Replication of these 5 loci and ~40 more with sub-GW significance are currently underway in populations of European, African American and South Asian descent as are functional studies of all genes spanning these regions. Conclusions: SNP associations for habitual coffee intake reinforce the role of the caffeine metabolic pathway and suggest new pathways related to neuronal function.

2097W

Genome-wide association study of Tanner puberty staging in males and females. *D. Cousminer[1], N. Timpson[2], D. Berry[3], W. Ang[4], I. Ntalla[5], M. Groen-Blokhuis[6], M. Guxens[7], M. Kähönen[8], J. Viikari[9], T. Lehtimäki[10], K. Panoutsopoulou[11], D. Boomsma[6], E. Zeggini[11], G. Dedoussis[5], C. Pennell[4], O. Raitakari[12], E. Hyppönen[3], G. Davey Smith[2], M. McCarthy[13], E. Widén[1], The Early Growth Genetics (EGG) Consortium.* 1) Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Finland; 2) The Medical Research Council (MRC) Centre for Causal Analyses in Translational Epidemiology, School of Social and Community Medicine, University of Bristol, Bristol, UK; 3) Centre for Paediatric Epidemiology and Biostatistics, MRC Centre for Epidemiology of Child Health, UCL Institute of Child Health, London, UK; 4) University of Western Australia, Perth, Western Australia, Australia; 5) Harokopio University of Athens, Department of Dietetics and Nutrition, Athens; 6) Netherlands Twin Register, Department of Biological Psychology, VU University, Amsterdam, The Netherlands; 7) Center for Research in Environmental Epidemiology (CREAL), Barcelona, Catalonia, Spain; 8) Department of Clinical Physiology, University of Tampere and Tampere University Hospital, Finland; 9) Department of Medicine, University of Turku, Finland; 10) Department of Clinical Chemistry, Fimlab Laboratories, University Hospital and University of Tampere, Finland; 11) Wellcome Trust Sanger Institute, Hinxton, UK; 12) Department of Clinical Physiology and Nuclear Medicine, University of Turku, Finland; 13) Wellcome Trust Centre for Human Genetics, Roosevelt Drive, University of Oxford, Oxford, UK.

Puberty is a complex trait with large variation in timing and tempo in the population, and extremes in pubertal timing are a common cause for referral to pediatric specialists. Recently, large genome-wide association studies (GWAS) have revealed 42 common variant loci associated with age at menarche (AAM), and some implicated genes are known from severe single-gene disorders. However, little remains known of the genetic architecture underlying normal variation in the onset of puberty, especially in males.

Tanner staging, a 5-stage scale assessing female breast and male genital development, is a commonly used measure of pubertal development. While AAM is a late event during puberty, Tanner staging during mid-puberty may correlate more closely with the central activation of puberty. With Tanner scale data at the comparable ages of 11–12 yrs in girls and 13–14 yrs in boys, we performed GWAS meta-analyses across 10 cohorts with up to 9,900 samples. The combined male and female analysis showed evidence for association near *LIN28B* ($P=1.95 \times 10^{-8}$), previously implicated in AAM and height growth in both sexes. Our data confirms that this locus is also important for male pubertal development and may be part of the pubertal initiation program upstream of sex-specific mechanisms. A novel signal ($P= 4.95 \times 10^{-8}$) with a consistent direction of effect across contributing datasets locates on chromosome 1 at an intronic transcription factor binding-site cluster within the gene *CAMTA1*. Furthermore, the primary analyses revealed suggestive evidence for male-specific loci, e.g. nearby *MKL2* ($P= 4.68 \times 10^{-7}$), which may be confirmed by follow-up genotyping. MAGENTA gene-set enrichment analysis of the combined-gender GWAS results showed enrichment of genes involved in expected pathways given the known biology underlying activation of puberty via the HPG axis. Novel genes near suggestively associated loci may also pinpoint novel regulatory mechanisms; *CAMTA1* is a calmodulin-binding transcriptional activator, while *MKL2* is also a transcriptional activator involved in cell differentiation and development. These results suggest the presence of multiple real signals beneath the genome-wide significant threshold, and further exploration of enriched pathways may reveal new insights into the biology of pubertal development.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2098T

Genetic variation associated with circulating monocyte count in the eMERGE Network. *D. Crosslin[1,19], A. McDavid[2], N. Weston[3], X. Zheng[4], E. Hart[3], M. de Andrade[5], I. Kullo[6], C. McCarty[7,8], K. Doheny[9], E. Pugh[9], A. Kho[10], M. Hayes[11], M. Ritchie[12], A. Saip[13], D. Crawford[14,15], P. Crane[16], K. Newton[3], R. Li[17], D. Mirel[18], A. Crenshaw[18], E. Larson[3], C. Carlson[2], G. Jarvik[1,19].* 1) Department of Medicine, Division of Medical Genetics, University of Washington, Seattle, WA; 2) Fred Hutchinson Cancer Research Center, Public Health Sciences Division, Seattle WA; 3) Group Health Research Institute, Center for Health Studies, Seattle, WA; 4) Department of Biostatistics, University of Washington, Seattle, WA; 5) Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, MN; 6) Division of Cardiovascular Diseases, Mayo Clinic, Rochester, MN; 7) Essentia Institute of Rural Health, Duluth, MN; 8) Center for Human Genetics, Marshfield Clinic Research Foundation, Marshfield, WI; 9) Center for Inherited Disease Research, Johns Hopkins University, Baltimore, MD; 10) Divisions of General Internal Medicine and Preventive Medicine, Northwestern University, Chicago, IL; 11) Division of Endocrinology, Metabolism, and Molecular Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL; 12) Center for Systems Genomics, Department of Biochemistry and Molecular Biology, Pennsylvania State University, University Park, PA; 13) Department of Biomedical Informatics, Vanderbilt University, Nashville, TN; 14) Center for Human Genetics Research, Vanderbilt University, Nashville, TN; 15) Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, TN; 16) Division of General Internal Medicine, University of Washington, Seattle, WA; 17) Office of Population Genomics, National Human Genome Research Institute, Bethesda, MD; 18) Program in Medical & Population Genetics, Broad Institute of Harvard & MIT; Cambridge, MA; 19) Department of Genome Sciences, University of Washington, Seattle, WA.

With white blood cell count (WBC) emerging as an important risk factor for chronic inflammatory diseases, genetic associations of differential leukocyte types, specifically monocyte count, are providing novel candidate genes and pathways to further investigate. Circulating monocytes play a critical role in vascular diseases such as in the formation of atherosclerotic plaque. We performed a joint and ancestry-stratified genome-wide association analyses to identify variants specifically associated with monocyte count in 13,923 subjects in the electronic Medical Records and Genomics (eMERGE) Network. In the joint and European ancestry samples, we identified novel significantly-associated variants on chromosome 16 from the interferon regulatory factor 8 (*IRF8*) gene (p-value = 2.78e-16, β = -0.22.) Other findings include novel missense variants in the chemokine binding protein 2 (*CCBP2*) gene (p-value = 1.88e-7, β = 0.30) and a region replication found in ribophorin I (*RPN1*) (p-value = 2.63e-16, β = –0.23) on chromosome 3. The *CCBP2* and *RPN1* region is located near GATA2 gene, which has been previously shown to be associated with coronary heart disease. On chromosome 9 we found a novel association in the prostaglandin reductase 1 (*PTGR1*) gene (p-value = 2.29e-7, β = 0.16,) which is downstream from lysophosphatidic acid receptor 1 (*LPAR1*.) This region has previously been shown to be associated with monocyte count. We also replicated monocyte associations of genome-wide significance (p-value = 5.68e-17, β = -0.23) at the *ITGA4* gene on chromosome 2. The novel *IRF8* results and further replications provide excellent candidate genes and pathways to further investigate.

2099F

A Genome-wide Association Study of Systemic Lupus Erythematosus in North Americans of European Ancestry. *F.Y. Demirci[1], X. Wang[1], A.H. Kao[2], A. Clarke[3], R. Ramsey-Goldman[4], S. Manzi[2], M.M. Barmada[1], M.I. Kamboh[1].* 1) Human Genetics, GSPH, Univ Pittsburgh, Pittsburgh, PA, USA; 2) Medicine, West Penn Allegheny Health System, Pittsburgh, PA, USA; 3) Medicine, McGill Univ, Montreal, Quebec, Canada; 4) Medicine, Northwestern Univ, Chicago, IL, USA.

Systemic lupus erythematosus (SLE) is a complex autoimmune disease that is influenced by the interplay of heritable, hormonal, and environmental factors. Genome-wide association studies (GWAS) in individuals of European ancestry have identified a number of SLE susceptibility loci using moderate-size samples and earlier versions of Illumina or Affymetrix platforms. Follow-up studies on GWAS regions showing suggestive associations using larger samples and more markers have implicated additional loci. In this study, we report the results of a genome-wide association analysis of SLE in a North American case-control sample of European ancestry (n= 1,166) genotyped by using the Affymetrix Genome-Wide Human SNP Array 6.0 consisting of 906,600 SNPs. As expected, the most significant associations were observed at the major histocompatibility complex (MHC) locus on chromosome 6p21, from which the SNPs easily surpassed the stringent genome-wide significance threshold of $5x10^{-8}$. Among several other previously implicated SLE loci also supported in our study, the strongest associations were observed in the *STAT4* (2q32) and *BLK* (8p23) regions (p= $4.9x10^{-7}$ and $5.0x10^{-6}$, respectively). We also observed several novel loci with p<$5x10^{-5}$ and the most compelling ones were at 2p22 (p=$2.1x10^{-6}$) and 12p13 (p=$6.7x10^{-6}$). These new suggestive loci harbor candidate genes potentially relevant to SLE and related subphenotypes based on their functions. The newly implicated SLE loci in our GWAS warrant further investigation and confirmation in other SLE case-control samples.

2100W

Genome Wide Association Study of Sexual Orientation in a Large, Web-based Cohort. *E.M Drabant, A.K. Kiefer, N. Eriksson, J.L. Mountain, U. Francke, J.Y. Tung, D.A. Hinds, C.B. Do.* 23andMe, Mountain View, CA.

There is considerable variation in human sexual orientation. Heritability studies have differed on the exact scope of genetic contributions for sexual orientation, but it appears that both genetics and environment play a role. Though a few linkage studies have pointed at a possible role for certain genes on the X chromosome, the strength of that evidence is limited due to the conflicting nature of the reports and small sample sizes. We sought to clarify some of the questions surrounding the possible genetic underpinnings of sexual orientation by deploying a web-based survey to the large 23andMe database and conducting the first ever genome-wide association study (GWAS) on sexual orientation.

We adapted the Klein Sexual Orientation Grid to examine seven elements of sexual orientation. All items were rated on a seven point scale by participants. Initial analyses focused on the "self identification" item as a continuous variable in response to the question "How do you label, identify or think of yourself?" In a sample of 7,887 men and 5,570 women, 77.2% of men 74.6% of women identified as heterosexual only, 7.3% of men and 15.3% of women as heterosexual mostly, 1.1% of men and 2.7% of women as heterosexual somewhat more, 1.3% of men and 3.5% of women as bisexual, 0.7% of men and 0.5% of women as homosexual somewhat more, 2.9% of men and 1.6% of women as homosexual mostly, and 9.5% of men and 1.8% of women as homosexual only. In both men and women, sexual identity was most significantly correlated with sexual attraction (men r=0.97, women r=0.90), sexual behavior (men r=0.95, women r=0.83), sexual fantasies (men r=.96, women r=.75), and emotional attraction (men r=0.79, women r=0.45), and the least strongly correlated with heterosexual/homosexual lifestyle (men r=.54, women r=.37), and social preference (men r=.15, women r=.08).

We carried out GWAS stratified by sex in a cohort of 7887 unrelated men and 5570 unrelated women of European ancestry collected in the two months since the initial survey release. No clear genome-wide significant associations have been found thus far, and the current data do not show any direct association for markers within chromosome band Xq28. However, data collection is still ongoing, and increased sample size may help to clarify the roles for currently suggestive associations.

2101T

Identification of multiple novel susceptibility regions for primary sclerosing cholangitis through dense genotyping of autoimmunity risk loci. *E. Ellinghaus[1,4], J.Z. Liu[2,4], T. Folseraas[3,4], A. Franke[1,5], C.A. Anderson[2,5], T.H. Karlsen[3,5] on behalf of The International Primary Sclerosing Cholangitis Study Group &The Immunochip Consortium.* 1) Institute of Clinical Molecular Biology, Christian-Albrechts-University Kiel, Kiel, Germany; 2) Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, UK; 3) Medical Department, Rikshospitalet University Hospital and University of Oslo, Oslo, Norway; 4) These authors contributed equally; 5) Shared senior authorship.

Primary sclerosing cholangitis (PSC) is a rare bile duct and liver disease of unknown etiology. In lack of medical therapy, PSC has become a major indication for liver transplantation in Northern Europe and the USA. An important clinical feature of PSC is the frequent occurrence of inflammatory bowel disease (IBD), which is reported in 60% to 80% of PSC patients of Northern European descent. So far, two genome-wide association studies (with up to 715 patients included at the discovery stage of analysis) have identified four genome-wide significant (P<5×10^{-8}) associations for PSC (HLA, *MST1*, *IL2RA* and *MMEL1*).

To identify additional PSC risk loci and systematically interrogate the genetic relationship between PSC and other immune-mediated diseases, we performed genotyping at 186 distinct risk loci previously implicated in 12 autoimmune diseases using a custom high-throughput SNP genotyping array, the Immunochip. This array was specifically designed to perform fine-mapping and deep replication of previously established risk loci by investigating both common and rare SNPs.

After genotyping of 196,524 SNPs in 4,228 PSC cases and 27,077 controls from throughout Europe, Canada and the USA, stringent quality control left 3,788 PSC cases and 25,079 controls for association analysis by logistic regression with the first 10 principal components included as covariates. Our analysis yielded seven novel PSC risk loci at genome-wide significance (P<5×10^{-8}) within or near the following genes: *CTLA4*, *PSMG1*, *BACH2*, *PRKD2*, *IL2*/*IL21*, *SH2B3* and *CLEC16A*. Six of the seven newly identified PSC risk loci overlap with previously established susceptibility loci for type 1 diabetes, with a less pronounced overlap with celiac disease, rheumatoid arthritis and IBD. Further analysis of the obtained data is ongoing and will be presented. The outcome is likely to provide unprecedented insight into the biological pathways underlying PSC as well as the relationship between PSC and several more prevalent autoimmune diseases, both of which may lead to the development of novel therapeutic strategies for a rare, but debilitating, condition.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2102F

Screening of inflammation-related SNPs identifies a functional IL6 receptor variant as a risk factor for Atopic Dermatitis. *J. Esparza Gordillo[1,2], H. Schaarschmidt[3], L. Liang[4], W.O.C.M. Cookson[5], M.L. Lee-Kirsch[6], J. Henderson[7], L. Paternoster[7], J.I. Harper[8], E. Mangold[9], M.M. Nothen[9,10], F. Rüschendorf[2], T. Kerscher[1,2], I. Marenholz[1,2], A. Matanovic[1,2], S. Lau[11], T. Keil[12], C.P. Bauer[13], M. Kurek[14], A. Ciechanowicz[15], M. Macek[16], A. Franke[3], M. Kabesch[17], N. Hubner[2], G. Abecasis[18], S. Weidinger[19], M. Moffat[5], Y.A. Lee[1,2].* 1) Experimental and Clinical Research Center, Charite Univ Med, Berlin, Germany; Lindenberger Weg 80; 13125-Berlin. Germany; 2) Max-Delbrück-Centrum (MDC) for Molecular Medicine, Robert-Rössle-Strasse 10; 13125-Berlin; 3) Institute for Clinical Molecular Biology, Christian-Albrechts-University, Kiel, Germany Institut für Klinische Molekularbiologie Christian-Albrechts-Universität zu Kiel Universitätsklinikum Schleswig-Holstein Schittenhelmstr. 12 24105 Kiel; 4) Department of Epidemiology and Department of Biostatistics, Harvard School of Public Health 677 Huntington Avenue Boston, MA 02115-6018, USA; 5) National Heart and Lung Institute, Imperial College, London, UK Imperial College London, South Kensington Campus, London SW7 2AZ; 6) Klinik fur Kinder- und Jugendmedizin, Technical University Dresden, Dresden, Germany 01304 Dresden; 7) Department of Community-Based Medicine, University of Bristol, Bristol BS8 2BN, UK; 8) Department of Paediatric Dermatology, Great Ormond Street Hospital for Children, London, UK London WC1N 3JH; 9) Institute of Human Genetics, University of Bonn, Bonn, Germany 53105 Bonn; 10) Department of Genomics, Life & Brain Center, University of Bonn, Bonn, Germany; 11) Pediatric Pneumology and Immunology, Charité, Berlin, Germany D13353 Berlin; 12) Institute for Social Medicine, Epidemiology and Health Economics, Charité, Berlin, Germany; 13) Department of Pediatrics, Technical University of Munich, Munich 80333 Munich; 14) 14 Department of Clinical Allergology, Pomeranian Medical University, Szczecin, Poland 70-204 Szczecin Poland; 15) Department of Laboratory Diagnostics and Molecular Medicine, Pomeranian Medical University, Szczecin, Poland; 16) Department of Biology and Medical Genetics, Charles University Prague, 2nd Medical School and Faculty, Hospital Motol, Prague, Czech Republic; 17) Department of Pediatric Pneumology, Hannover Medical School, Allergy and Neonatology, Hannover, Germany; 18) Department of Biostatistics, Center for Statistical Genetics, University of Michigan School of Public Health, Ann Arbor, Michigan, USA; 19) Department of Dermatology, Venereology and Allergy, University Hospital Schleswig-Holstein, Kiel.

Genome wide association studies (GWAS) have provided great insight into the genetic architecture of complex diseases. Unfortunately, huge samples sizes will be needed to detect all the multiple low-risk variants likely to play a role in disease at the significance threshold required in GWAS. Therefore, a major challenge is the identification of new risk factors among the thousands of markers that show evidence for association but do not reach genome-wide significance (the so called “GWAS grey zone”). The parallel increase in prevalence observed in common inflammatory disorders together with the identification of SNPs simultaneously influencing the risk of multiple diseases points to overlapping etiological factors among these disorders. We aimed to use this overlap in order to identify new risk factors for the chronic inflammatory skin disease Atopic Dermatitis (AD). We inspected the GWAS repository (http://www.genome.gov/gwastudies/) and selected all 318 SNPs associated to any immune-related trait at genome-wide significance. These markers were then screened for association with AD in a GWAS meta-analysis including 2890 cases and 2030 controls. Interestingly a strong enrichment on AD-associated markers was observed among these 318 candidate SNPs (p<2x10-7), confirming our hypothesis of an overlap among the genetic determinants of AD and inflammatory traits. Eight markers moderately associated to AD (GWAS_p<0.01) were selected for follow-up in 3 cohorts including 4235 cases and 6805 controls. A missense functional variant in the IL6 receptor (rs2228145) affecting the shedding of the soluble IL6R (sIL6R) was significantly associated with AD (OR=1.15; P=5x10-9). rs2228145 correlated with serum sIL6R levels (P=4x10-14) and sIL6R was increased in AD patients (P=0.001). This finding suggests that targeted blocking of IL6R signaling with Tocilizumab may be a novel target for clinical intervention in AD, especially for individuals carrying this variant. Additionally variants in RAD50, RUNX3, and ERBB3 were also associated with AD. Strikingly, all 4 risk SNPs reported here showed only modest association with AD in the GWAS and would have been missed if selecting markers based on significance level alone. This study demonstrates that inclusion of relevant biological information is a valuable tool for GWAS prioritization and may help to identify novel susceptibility genes in already-existing GWAS data sets.

2103W

A novel sarcoidosis risk locus for Europeans on chromosome 11q13.1. *A. Fischer[1], B. Schmid[1], D. Ellinghaus[1], M. Nothnagel[2], K.I. Gaede[3], M. Schürmann[4], S. Lipinski[1], P. Rosenstiel[1], G. Zissel[5], K. Höhne[5], M. Petrek[6], V. Kolek[7], S. Pabst[8], C. Grohé[9], J. Grunewald[10], M. Ronninger[11], A. Ecklund[10], L. Padyukov[11], C. Gieger[12], H.-E. Wichmann[13,14,15], A. Nebel[1], A. Franke[1], J. Müller-Quernheim[5], S. Hofmann[1], S. Schreiber[1,16,17].* 1) Institute of Clinical Molecular Biology, Christian-Albrechts University, Kiel, Germany; 2) Institute of Medical Informatics and Statistics; Christian-Albrechts University; Kiel, Germany; 3) Department of Pneumology, Research Center Borstel, Borstel, Germany; 4) Institute of Human Genetics, University of Lübeck, Lübeck, Germany; 5) Department of Pneumology, University of Freiburg, Freiburg, Germany; 6) Laboratory of Immunogenomics and Immunoproteomics, Faculty of Medicine and Dentistry, Palacky University Olomouc, Czech Republic; 7) Dept. of Respiratory Medicine, Faculty of Medicine and Dentistry, Palacky University Olomouc, Czech Republic; 8) Medical Clinic II, Department of Pneumology, University of Bonn, Germany; 9) Department of Respiratory Medicine, Evangelische Lungenklinik Berlin-Buch, Germany; 10) Respiratory Medicine Unit, Department of Medicine, Karolinska Institutet, Stockholm, Sweden; 11) Rheumatology Unit, Department of Medicine, Karolinska Institutet, Stockholm, Sweden; 12) Institute of Genetic Epidemiology, Helmholtz Centre Munich, German Research Center for Environmental Health, Neuherberg, Germany; 13) Institute of Epidemiology I, Helmholtz Centre Munich, German Research Center for Environmental Health, Neuherberg, Germany; 14) Institute of Medical Informatics, Biometry and Epidemiology, Ludwig-Maximilians-University, Munich, Germany; 15) Klinikum Grosshadern, Munich, Germany; 16) Department of General Internal Medicine, University Hospital Schleswig-Holstein, Kiel, Germany; 17) Popgen Biobank, University Hospital Schleswig-Holstein; Kiel, Germany.

Sarcoidosis is a complex inflammatory disease with a heterogeneous clinical picture. Amongst others, an acute and chronic clinical course can be distinguished, for which specific genetic risk factors are known. In order to identify additional risk loci for sarcoidosis and its acute and chronic subforms, we analyzed imputed data from a genome-wide association scan (GWAS) for these phenotypes.

After quality control, the GWAS comprised nearly 1.3 million imputed single nucleotide polymorphisms based on an Affymetrix 6.0 Gene Chip dataset of 564 German sarcoidosis cases, including 176 acute and 354 chronic cases and 1575 controls. We identified chromosome 11q13.1 as a novel locus influencing susceptibility to sarcoidosis with genome-wide significance. The lead marker was significantly associated in three distinct German case-control populations and in an additional German family sample with ORs ranging from 0.67 to 0.77. This finding was further replicated in two independent European case-control populations from the Czech Republic (OR = 0.75) and from Sweden (OR = 0.79). In a meta-analysis of the included European case-control samples the marker yielded a p value of 2.68×10^{-18}. The locus was previously reported to be associated with Crohn's disease and alopecia areata. For sarcoidosis, fine-mapping and expression analysis suggest several transcripts as candidates for the underlying risk gene in the associated region.

In summary, this study provides striking evidence for association of chromosome 11q13.1 with sarcoidosis in Europeans, and thus identified a further genetic risk locus shared by sarcoidosis and Crohn's disease. Full results, including marker and gene names, will be shown in the oral presentation or on the poster, respectively.

2104T

A GWAS on Age-Related Hearing Impairment in a European population. *E. Fransen[1], S. Bonneux[1], I. Schrauwen[1], F. Di Berardini[2], I. Dhooge[3], J.J. Corneveaux[4], J.D. Ohmen[5], P.H. Van de Heyning[6], R. Friedman[5], M.J. Huentelman[4], G. Van Camp[1].* 1) Center Medical Genetics, University of Antwerp, Antwerp, Belgium; 2) Dipartimento di Scienze Chirurgiche Specialistiche, Università degli Studi di Milano, Milan, Italy; 3) ENT department, University Hospital of Ghent, Ghent, Belgium; 4) Neurogenomics Division, The Translational Genomics Research Institute, Phoenix, AZ, USA; 5) House Ear Institute, Gonda Research Center for Cell and Molecular Biology, Los Angeles, CA, USA; 6) ENT department, University Hospital of Antwerp, Antwerp, Belgium.

Age-related hearing impairment (ARHI) is characterized by a gradual, symmetric and sensineural loss of hearing acuity beyond age 50. Hearing loss is sensorineural, symmetric and primarily affecting the high frequencies. Heritability studies have indicated that ARHI is a complex trait, with approximately half of the variance attributable to genetic factors, and the remainder to environmental risk factors. There where several environmental risk factors have been identified including noise exposure and smoking, the genetic basis of ARHI remains to be elucidated. We and others have already performed several large-scale association studies into ARHI, pointing at several interesting candidate genes, but none of these candidate genes has ever shown a consistent replication across multiple studies. Here we present the results of a genome-wide association study on a population of 2161 samples collected in and around Antwerp, Belgium. We report the most promising candidate genes as scored by classic association test, show the results of pathway analysis and compare our results to the results of previous association studies into ARHI.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2105F

An extensive literature-based multilocus genetic score is associated with all-cause-mortality and major diseases: The Rotterdam and TwinGene studies. *A. Ganna[1], A. C. Janssens[2], A. Hofman[2], F. Rivadeneira[2], A. G. Uitterlinden[2], P. K. E. Magnusson[1], N. L. Pedersen[1], E. Ingelsson[1], H. Tiemeier[2].* 1) Department of medical epidemiology and biostatistics, Karolinska institutet, Sweden; 2) Department of epidemiology, Erasmus Medical Center, The Netherlands.

Background: Twin studies have shown that approximately 20–30 % of the overall variation in lifespan is caused by genetic differences, and that the genetic influence on lifespan increase with age. Moreover, centenarians' offspring have lower risk of all-cause mortality and cause-specific mortality from cancer and coronary heart disease than controls. Genome-wide association (GWA) studies have revealed a large number of single nucleotide polymorphisms (SNPs) associated with diseases and their risk factors, but have failed to robustly identify any longevity loci besides the well-established APOE locus. Objectives: To create a multilocus genetic score including SNPs that have been found to be associated with mortality-related traits in GWA studies, and to test for association of this score with all-cause-mortality and incidence of major diseases. Methods: Using the NHGRI-GWAS catalog, six medical doctors with expertise in epidemiology independently selected 125 traits moderately or highly associated with overall mortality. After extensive quality control, we included 709 SNPs in a unweigthed multilocus genetic score by adding the number of risk alleles. We prospectively studied the association of this multilocus genetic score with: (1) time-to-death; and (2) incidence of at least one of nine major diseases (coronary heart disease, stroke, heart failure, diabetes, dementia, lung, breast, colon and prostate cancers) in two population-based cohort studies of Dutch and Swedish individuals (Rotterdam study and TWINGENE; total N=15,039). We further investigated whether the strength of associations varied with age. Results: During a median follow-up of 6.0 years (max, 22.2 years), we observed 4,267 deaths and 2,132 incident disease events. The multilocus genetic score was significantly associated with time-to-death, but the association was very modest (hazard ratio [HR] per added risk allele, 1.002; P-value=0.01). The association remained significant after removing APOE rs2075650 from the multilocus genetic score, which was our top association in the single-marker analysis. The association between the multilocus genetic score and incidence of major diseases was stronger (HR per added risk allele, 1.005; P-value=0.002). Associations were more pronounced for individuals died at average life expectancy. Conclusions: Genetic determinants of longevity are partially overlapping with those of commons diseases.

2106W

Large-scale meta-analysis in up to 40,429 Europeans identifies genetic loci associated with non-fasting plasma glucose and reveals links to glucose uptake in muscle. *KS. Gutierrez[1], V. Lagou[2,3], A. Isaacs[1,4], JB. Meigs[6,7], YS. Aulchenko[1,5], I. Prokopenko[2,3] for MAGIC (the Meta-Analyses of Glucose and Insulin-related traits Consortium) investigators.* 1) Genetic Epidemiology Unit, Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands; 2) Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom; 3) Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford, United Kingdom; 4) Centre for Medical Systems Biology, Leiden, The Netherlands; 5) Laboratory of Recombination and Segregation Analysis, Institute of Cytology and Genetics SD RAS, Novosibirsk, Russia; 6) General Medicine Division, Massachusetts General Hospital, Boston, Massachusetts, USA; 7) Department of Medicine, Harvard Medical School, Boston, Massachusetts, USA.

Levels of circulating plasma glucose are strictly regulated. An unusually high non-fasting (random) plasma glucose (RG) level (measurement of plasma glucose at any time point during the day) might indicate abnormal glucose homeostasis and can be used to suggest a diagnosis of diabetes. We sought to identify genetic variants contributing to the variability of RG and to elucidate the genetic relationship between RG, other glycemic traits and type 2 diabetes (T2D). We conducted an inverse variance fixed-effects meta-analysis of genome-wide association studies (GWAS) for RG, adjusted for age and sex, at ~2.5M European HapMap II reference panel-imputed SNPs and SNPs genotyped on the Metabochip platform. Up to 40,429 non-diabetic European individuals from 14 studies were available for analysis. We found three variants associated with RG at a genome-wide significance level: at *G6PC2* (rs573225,P=1.27x10^{-20}), *GCK* (rs2908282, P=1.18x10^{-17}) and the muscle-specific tyrosine-protein kinase receptor (*MuSK*) (rs4500149, P=1.57x10^{-8}). The first two were also associated with fasting glucose (FG) and HbA1c in previous studies. Association at the *MuSK* gene locus is novel and a promising biological candidate, due to its important role in muscle contraction through increased glucose uptake leading to improved cellular insulin efficiency. We observed suggestive associations (P<10^{-5}) within *ADCY5* (rs7613951, P=2.43x10^{-7}), *SLC2A2* (rs16847990,P=1.21x10^{-6}), both previously associated with FG, and at *ANKRD32* (rs7718956, P=1.09x10^{-6}) and *TMEM132D* (rs10847533, P=1.09x10^{-6}). A model with an additional adjustment for the time since last meal produced similar results, but was available only in a subset of studies. We also evaluated consistency between the effects of known T2D and glycemic variants on T2D, FG and RG. To that end, using Kolmogorov-Smirnov test we revealed that T2D, fasting insulin, FG and 2 hours post load glucose (2hGlu) variants are enriched for the RG association signals (P=5.77x10^{-15}, N=158). Consistently, T2D (P=1.53x10^{-6}, N=57), FG (P=2.81x10^{-10}, N=39) and 2hGlu (P=0.001, N=7) loci also overlapped with RG associations when analyzed separately. In conclusion, RG shares genetic background with T2D and other glycemic traits. Association at the *MuSK* gene locus needs to be confirmed and suggests that analysis of RG will provide additional insights into regulation of glucose homeostasis and muscle glucose uptake.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2107T

Genome-wide association study detects multiple novel loci associated with Nuclear Magnetic Resonance spectroscopy detected metabolites in human serum. *P. Henneman[1], A. Verhoeven[2], H. Dharuri[1], J. van Klinken[1], A. Meissner[2], S. Göraler[2], A. Deelder[2], R. Frants[1], L. Karssen[3], B. Oostra[3], K. Willems van Dijk[1,4], C. van Duijn[3].* 1) Human Genetics, Leiden University Medical Center, Leiden, The Netherlands; 2) Dept. Parasitology, Leiden University Medical Center, Leiden, The Netherlands; 3) Dept. Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands; 4) Dept. Internal Medicine, Leiden University Medical Center, Leiden, The Netherlands.

Introduction: Recently reported GWA studies on classical Metabolic Syndrome (MetS) traits showed that only moderate proportions of the heritability of these traits could be explained. Recent reports hypothesized and showed that one possible approach to tackle this problem of "missing heritability" is to make use of intermediate or endo-phenotypes. Here we aim to find novel loci involved in MetS-related traits using intermediate metabolomic phenotypes detected by state of the art NMR technology. Methods: The present study applies to a Dutch genetically isolated population: The Erasmus Rucphen Family (ERF) study. Two-dimensional J-resolved NMR spectroscopy experiments were performed on human sera using a 14.09T 600 MHz ultra shield Plus from Bruker. Forty-three metabolites for 2482 ERF participants were successfully quantified and QC-ed. Heritability estimates were performed using SOLAR software. Genome-wide association was based on 300K genotypes imputed at 2.5 million SNPs. Association analysis was performed using the GenABEL "R" packages that enables adjustment for kinship. 265 ratios were derived from concentrations using novel bioinfomatic techniques, which selected ratios of metabolites only in the vicinity of a metabolite. Gene-based association was performed using VEGAS. Results: Heritability estimates for metabolites or the derived measures ranged between 10% and 70%. Genome-wide association for 43 metabolites and 265 derived measures were performed. Thirty-four SNPs in 23 metabolite traits showed P-value < 1e -8 of which 15 SNPs were located at genes that were reported earlier (e.g. *LPL*, *GCKR*, *ABO*). Six SNPs showed P-value < 1e-10. Moreover, a gene-based approach using the genome-wide results, detected 7 additional loci in 6 of the 43 metabolites. For the derived traits we found 268 SNPs in 30 traits associated with a P-value < 1e0–10. Novel bioinformatic techniques indentified promising MetS-related biochemical reactions and metabolic pathways. Replication (Finish cohorts of Kettunen et al. 2012, N~8000) and meta-analysis analysis are ongoing. Conclusion: Here we show successful detection of several novel and confirmation of earlier associated loci with NMR detected metabolites, which may provide substantial contribution to basal physiological knowledge of metabolic syndrome and cardiovascular disease.

2108F

Ontology, visual, and informatics enhancements to the NHGRI Genome-wide Association Study (GWAS) catalog. *L.A. Hindorff[1], J.A.L. MacArthur[2], D. Welter[2], T. Burdett[2], P. Hall[1], H.A. Junkins[1], H. Parkinson[2].* 1) Population Genomics, NHGRI, NIH, Bethesda, MD; 2) EMBL-EBI, Wellcome Trust Genome Campus, Hinxton, UK.

The National Human Genome Research Institute's (NHGRI) GWAS Catalog (http://www.genome.gov/gwastudies) contains over 1,300 publications and 6,400 disease- or trait-associated genetic variants, which have been manually curated from published genome-wide association studies (GWAS). The availability of this curated, regularly updated, and downloadable resource contributes to the knowledge base of genetic variants associated with important diseases and conditions and enables cross-linking with other commonly used resources such as Ensembl, Database of Genotypes and Phenotypes (dbGaP), the Phenotype-Genotype Integrator (PheGenI) and UCSC Genome Browser. We have implemented several improvements to increase the usability of catalog data and ease of integration with other resources. First, catalog traits have been integrated into the Experimental Factor Ontology (Malone et al., 2010), an application ontology based on a paradigm of reusing elements from reference ontologies including PATO and ChEBI. This standardization of trait organization will facilitate increased integration of NHGRI GWAS catalog data with other sources and improved querying of the catalog. Building upon this ontology, generation of the iconic GWAS catalog diagram has been automated (http://wwwdev.ebi.ac.uk/fgpt/gwas) such that visual representations of GWAS associations across the human genome are now available filtered by different criteria, including disease category and p-value. These customized diagram images can be downloaded for use in publications and presentations. SNP-trait associations in the online version of the diagram are also interactive, with dynamic links to the GWAS catalog entry and to external resources such as Ensembl and dbSNP. Informatics improvements have also increased the efficiency and quality of the curation process. For example, the previous limit of 50 SNP-disease associations per paper is no longer in effect, with all qualifying associations with a p-value < $1x10^{-5}$ now included in the catalog. Ultimately, the benefits of standardized disease annotation, customized views of the GWAS catalog diagram, and efficient curation will yield an enhanced user experience and facilitate further integration of GWAS catalog data with other resources.

2109W

Genome-wide association study of time-to diabetic retinopathy. *S.M. Hosseini[1,2], A.P. Boright[3], L. Sun[4], K. Howard[5], A.J. Canty[6], S.B. Bull[4,7], B.E. Klein[5], R. Klein[5], A.D. Paterson[1,2,4], the DCCT/EDIC research group.* 1) Gen & Genomic Biol, Rm 15–701, Hosp Sick Children, Toronto, ON, Canada; 2) Inst of Medical Science, Univ Toronto, Toronto, Canada; 3) Dept of Medicine, Univ Toronto and LMC Diabetes & Endocrinology, Toronto, Canada; 4) Dalla Lana School of Public Health, Univ Toronto, Toronto, Canada; 5) Dept of Ophthalmology & Visual Sciences, Univ Wisconsin, Madison, WI, USA; 6) Dept of Mathematics & Statistics, McMaster Univ, Hamilton, Canada; 7) Samuel Lunenfeld Research Institute, Toronto, Canada.

Diabetic retinopathy (DR) is a major health problem affecting up to 82% of patients with type 1 diabetes (T1D). Duration of diabetes and poor glycemic control are the major known risk factors for DR. Consistent with a genetic contribution, the severity of DR clusters in families, yet few genetic variants have shown convincing and consistent associations with DR.

White T1D participants in the Diabetes Complications and Control Trial (DCCT, n=1304) and Wisconsin Epidemiologic Study of Diabetic Retinopathy (WESDR, n=627) were genotyped by Illumina Human1M and HumanOmni1-Quad assays, respectively. Following quality control procedures, genotypes of ~1.5M autosomal SNPs from HapMap3 were imputed in both studies. Time-to severe non-proliferative DR [level 53/<53 or worse on ETDRS scale or panretinal laser photocoagulation treatment] was defined by examination of repeated retinal photographs taken over more than 2 decades (average 4.5 photos in WESDR and 16.8 in DCCT/EDIC). To address both left and interval censoring, survival analysis was performed using a complementary log-log model. Multivariate models were built to account for glycemia (measured by A1C), diabetes duration and other relevant covariates in the association analysis of time-to-DR with genotype. Analyses were performed separately in WESDR and the 3 subgroups of DCCT (primary cohort, secondary cohort by treatment group - conventional or intensive) before combining the results in a meta-analysis.

Two SNPs on 8q23.3, upstream of *TRPS1*, showed significant association with risk of DR (p<5E-8). Suggestive associations (p<1E-6) were also observed for SNPs on: 2q14.1 in *DPP10*, a member of voltage gated potassium channels, associated with asthma; 5q31.3 in protocadherin beta gene cluster (*PCDHB1–16*); 11p14.3 near *NELL1*; and 6p22.3 (intergenic). Attempts to replicate the results in independent cohorts are in progress.

This first GWAS of time-to diabetic retinopathy illustrates the value of survival analysis with explanatory covariates in genetic studies of diabetic complications.

2110T

Genome-wide screen for telomere length loci identifies a cancer risk associated gene in Finnish cohorts. *I. Hovatta[1,2,3], I. Surakka[2,3], I. Sirén[1], J. Kettunen[2,3], J.G. Eriksson[2,4,5,6,7], P. Knekt[2], V. Salomaa[2], J. Kaprio[2,3,8], S. Ripatti[2,3,9].* 1) Molec Neurology Program, University of Helsinki, Helsinki, Finland; 2) National Institute for Health and Welfare, Helsinki, Finland; 3) Institute for Molecular Medicine Finland, FIMM, University of Helsinki, Helsinki, Finland; 4) Dept of General Practice and Primary Health Care, University of Helsinki, Helsinki, Finland; 5) Helsinki University Central Hospital, Helsinki, Finland; 6) Folkhälsan Research Centre, Helsinki, Finland; 7) Vasa Central Hospital, Vasa, Finland; 8) Dept of Public Health, University of Helsinki, Helsinki, Finland; 9) Wellcome Trust Sanger Institute, Hinxton, Cambridge, United Kingdom.

We carried out a genome-wide association study to identify loci influencing leukocyte telomere length (LTL) in four Finnish sample sets with 5133 individuals. The relative LTL was measured using a qPCR based methods. We tested genome-wide associations between 38 million markers and LTL. In a meta-analysis of the results three SNPs in the intron of the telomerase transcriptase (TERT) gene associated (p-value <5×10–8) with LTL. Additionally, we found 10 suggestive loci from which the best SNP per area, together with the most significant TERT SNP, rs7726159, were taken forward to the replication genotyping of 3452 individuals from Health 2000 and 4031 individuals from FINRISK92-07 cohorts. In the replication analyses the association of TERT locus was replicated separately in both replication datasets leading to a screening and replication stage combined P-value of 6.0×10–10. Variants in TERT have previously been associated with multiple types of cancer. We therefore investigated the TERT SNP, LTL and risk of cancer in the Health 2000 data (447 incident cancer cases and 4939 controls) in which we observed a suggestive association between LTL and cancer risk (P-value = 0.03) with a simple logistic regression with age and sex as covariates. We also investigated if individuals with shorter LTL have higher mortality. We saw a significantly higher risk of mortality for individuals with shorter telomeres (P-value = 0.003) when combining data from Health 2000 and one of the screening datasets with 914 incident death cases altogether. Longitudinal studies, such as ongoing in the Health 2000, will be essential to understand the causal relationship between LTL, cancer risk, mortality, and genetic variation in the TERT locus.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2111F

GWAS Identifies 13 Polymorphisms Associated with Motion Sickness. *B.S. Hromatka, E.R. Chang, J.Y. Tung, J.L. Mountain, U. Francke, C.B. Do, N. Eriksson.* 23andMe, Inc., Mountain View, CA.

Motion sickness and its underlying causes are not well understood and yet this condition affects many; between 7% and 28% of people experience acute motion sickness while traveling. This condition is also medically relevant, as individuals who experience motion sickness while traveling are more likely to experience nausea and vomiting after surgery. A better understanding of this common condition could perhaps lead to improved travel and surgery experiences. Twin studies on motion sickness suggest high heritability (57–70%), but to date there are no known genetic factors associated with motion sickness. To discover which genes may be involved, we conducted a genome-wide association study on over 33,400 individuals who were genotyped on the 23andMe platform and surveyed on motion sickness. We found thirteen SNPs associated with motion sickness at a genome-wide-significant level. Of these 13 SNPs, several are found near genes associated with eye, ear, and neurological development. Of note is rs6926953 (p = 1.95 x 10-9), which is 300 kb upstream from the gene MUTED; mutations in the mouse homolog of MUTED are implicated in balance problems. In addition, rs4891066 (p = 3.24 x 10-9) is 99 kb downstream from TSHZ1, a gene involved in inner ear development. Furthermore, rs17720662 (p = 2.69 x 10-27), which was the most significant association found, is 1.1 Mb upstream from the gene PVRL3, which plays a role in eye development. One prevailing theory of motion sickness is that nausea results from contradictory information the brain receives while traveling. The inner ear, which can perceive acceleration and influences balance, signals "moving" to the brain, while the eye signals "stationary" because the car or boat is stationary relative to the viewer. The proximity of several of these motion sickness-associated SNPs to genes associated with eye, ear, and neurological development suggests that these genes may be involved in motion sickness. These findings are an important first step towards understanding the biological basis of not only motion sickness, but of conditions such as post-operative nausea and vomiting as well.

2112W

Identification of Leprosy Host Genetic Susceptibility Loci. *A. Irwanto[1,2], L. Hong[3], F. Zhang[3], J. Liu[1].* 1) Human Genetics, Genome Institute of Singapore, Singapore, Singapore; 2) Saw Swee Hock School of Public Health, National University of Singapore, Singapore; 3) Shandong Provincial Institute of Dermatology and Venereology, Provincial Academy of Medical Science, Jinan, China.

Interaction among the pathogen, environmental factors and host genetic makeup determine the success of an infection and its disease outcome. Genome-wide association studies (GWAS) have been successful in identifying host susceptibility genes that show significant association with the development of leprosy, including *CCDC122*, *LACC1* (*C13orf31*), *NOD2*,*TNFSF15*, *HLA-DR-DQ*, *LRRK2*, *RIPK2*, *IL23R*, *RAB32* in the Chinese population and *TLR1*, *HLA-DRB1*, *HLA-DQA1* in the Indian population, but many more remain to be discovered. Further GWAS analysis with better coverage of genetic variation through imputation and better statistical power through enlarged samples size can help to identify additional novel susceptibility loci for leprosy. Here we performed a large GWAS analysis of leprosy by analyzing over 2 million SNPs in 1623 cases and 6580 controls of Northern and Southern Chinese as well as Chinese minority population through a whole genome imputation using 1000 Genomes phased-haplotype information (from 1092 individuals of multiple ethnicities) and association analysis adjusting for population stratification via logistic regression. Our GWAS analysis replicated most of the previously identified loci and revealed additional novel suggestive associations at 48 novel loci (p-value < 10^{-4}; I^2 < 30). The validation study of the novel association signals is carried out using Sequenom MassArray platform in independent datasets and will likely discover additional new susceptibility genes to provide further biological insight into the pathogenic mechanism of leprosy.

2113T

Genetic risk score in Multiple Sclerosis: Back to the individual in the post-GWAS era. *N. Isobe[1], P. KhanKhanian[1], B. Vinod[1], S. Cailler[1], H. Harbo[2], A. Santaniello[1], S. Hauser[1], J. Oksenberg[1], P. Gourraud[1].* 1) UCSF, San Francisco, CA; 2) Department of Neurology University Hospital, Oslo Norway.

By the end of 2011, after 9 GWAS in Multiple Sclerosis (MS), there were more than 50 genomic regions associated with MS. In the well-known genomic region of HLA-DRB1, the description of the associations of multiple MS genetic variants with susceptibility, disease severity and response to treatment has become extremely elaborated. We are thus facing a new situation: the discovery of additional MS genes increase the critical need for summary metrics of the MS-specific genes configuration of patients. As of 2012, we developed a log-additive weighted statistical model, termed Multiple Sclerosis Genetic Burden, to take into account as many as 64 single nucleotide polymorphisms (SNPs). In addition, if real HLA typing is available, the contribution of the MHC region can be flexibly taken into account to provide a more accurate risk model. When computed in over 3,000 samples of the USCF MS repository, the MSGB analysis demonstrates a higher aggregation of susceptibility variants in multi-case compared to sporadic MS families (p=4.72.10-4). A greater MSGB score in siblings of MS patients was associated with an increased risk of MS (odds ratio, 2.1; p = 0.001), but MSGB differences between probands and their sibs show that case-control status cannot be predicted with the currently available genetic data. Using this score we applied another original statistical method, the equivalence test (Gourraud et al 2011 Gen. Epi.) to prove that the primary progressive and relapsing remitting forms of MS share common genetic architecture. Finally, we used MSGB score as a simplified way to evaluate the genetic contribution to various aspects of MS such as age of onset (p=2.4e-2) and presence of oligo-clonal band (p= 4.21e-6). The 2012 version of the MSGB analyses underline the high variability of individual and family load with known common variants. The primary interest of aggregated genetic risk score, such as MSGB resides in its capacity to integrate genetic contributions to MS risk in a summarized yet genome-wide way.

2114F

A genome-wide association study of HIV-1 infection in African American and European American injection drug users. *E. Johnson[1], D. Hancock[1], J. Levy[2], G. Page[3], S. Novak[1], C. Glasheen[1], N. Gaddis[2], N. Saccone[4], J. Rice[5], Q. Wang[6], M. Moreau[6], K. Doheny[7], J. Romm[7], A. Brook[6], L. Bierut[5], A. Kral[8].* 1) Behavioral Heallth Epidemiology, RTI International, Research Triangle Park, NC; 2) Research Computing Division, RTI International, Research Triangle Park, NC;; 3) Genomics, Statistical Genetics, and Environmental Research Program, RTI International, Atlanta, GA; 4) Department of Genetics, Washington University in St. Louis, MO; 5) Department of Psychiatry, Washington University in St. Louis, MO; 6) Rutgers University Cell and DNA Repository (RUCDR), Riscataway, NJ; 7) Center for Inherited Disease Research (CIDR), Johns Hopkins University, Baltimore, MD; 8) Urban Health Program, RTI International, San Francisco, CA.

Nearly 40 million people worldwide are infected with HIV. Host genetic factors likely influence susceptibility to acquiring HIV upon exposure. However, the only validated genetic association with HIV-1 infection is the 32 base-pair deletion in the CCR5 gene, which is rare in the population. To identify and characterize novel genetic risk factors, we conducted GWA studies of HIV-1 infection in African American and European American injection drug users. Unlike the prior genetic studies, we focused on injection drug users whose transmission may be primarily parenteral rather the than mucosal or mother-to-infant, and used HIV-1 negative controls with individually assessed high probability of exposure. HIV-1 positive cases and negative controls were frequency-matched on behavioral and demographic risks of exposure. DNA was extracted from stored serum at RUCDR and treated with Illumina FFPE restoration prior to genotyping at CIDR on the Illumina Omni1-Quad array. Following quality control procedures for genotyped SNPs and participants, the analysis data set included ~700,000 autosomal SNPs in 2,010 African Americans (630 cases, 1,380 controls) and 1,149 European Americans (335 cases, 814 controls). In the separate ethnic groups, additive SNP genotypes were tested for association with HIV-1 infection using logistic regression models adjusted for age, gender, behavioral risk class, recruitment pre- or post- availability of antiretroviral therapy, and 10 principal component eigenvalues for population stratification. The ethnic-specific GWA results were then combined in a fixed-effects meta-analysis. The lowest P-value (meta-analysis P=1.88x10-6) was observed for a regulatory SNP on chromosome 9p13, located in a transcription factor binding site upstream of a gene essential to B-cell differentiation. B cells, a primary component of the adaptive immune system, may play a role in cell-to-cell transmission of the HIV virus and are known to be affected during disease progression (e.g., expansion of immature transitional B-cells and B-cell exhaustion). The SNP is associated with HIV-1 infection in both ethnic groups, but occurs more frequently in African Americans (MAF=21.9%) than in European Americans (MAF=1.3%). Analyses using 1000 Genomes imputed SNPs and genotyping of Hispanic American samples are underway to further characterize this biologically plausible gene region and others that may help determine the complex etiology of host response to HIV-1.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2115W

Genetics of Myopia in a Participant Driven, Web-Based Cohort. *A.K. Kiefer, J.Y. Tung, C. Do, D.A. Hinds, J.L. Mountain, U. Francke, N. Eriksson.* 23andMe, Inc., Mountain View, CA.

Myopia, or nearsightedness, is the most common eye disorder worldwide. Myopia results from the elongation of the eye axis, which causes images to be focused in the vitreous part of the eye instead of on the retina. In the United States, roughly a third of the adult population has clinically relevant myopia (more severe than -1 diopter), and its prevalence has increased in recent decades. The etiology of myopia is largely unknown, but is believed to have a substantial genetic component, with prior heritability estimates ranging from 0.60–0.90. Education, reading, and lack of time spent outdoors are known environmental risk factors for myopia, and may contribute to its rapidly increasing prevalence. We collected self-reported information on myopia from the 23andMe participant cohort. We describe the results of a genome-wide association analysis on this convenience sample consisting of approximately 23,000 cases, defined as anyone who reported a diagnosis of nearsightedness, and 16,000 controls, who reported that they had not been diagnosed with nearsightedness, all of European ancestry. We replicate the only two previously identified associations with myopia in Europeans, in GOLGA8B, which belongs to a family of Golgi auto-antigens, and in RASGRFI, a transcription initiation site that is highly expressed in neurons and the retina. We also find novel genome-wide-significant associations ($P < 5E-08$) in 13 regions. Notable among the novel findings are variants in LAMA2 (rs12193446,) a gene coding for laminin, an extracellular protein that is a major component of the basement membrane and believed to be involved in the development and maintenance of different eye structures; in RDH5 (rs3138142), which encodes a retinol dehydrogenase that catalyzes the final step in the biosynthesis of 11-cis retinaldehyde; and in KCNQ5 (rs7744813), a member of the KCNQ potassium channel gene family that produces an M-type K(+) current channel and is expressed in monkey retinal pigmentation epithelial cells. This largest ever GWAS on myopia in Europeans—by an order of magnitude—identifies many new potential biological pathways involved in the development of myopia.

2116T

A meta-analysis of 52,439 individuals identifies four loci associated with leptin levels independent of adiposity. *T.O. Kilpeläinen[1, 2], Q. Sun[3, 4], Z. Kutalik[5], K. Kristiansson[6], M. Mangino[7], M. Su[8], P.P. Framstaller[9], D. Pasko[10], T. Tanaka[11, 12], Y.J. Sung[13], P. Henneman[14, 15], A. Mahajan[16], M. Beekman[17, 18], M.E. Kleber[19], T.I.A. Sørensen[1], L.J. Rasmussen-Torvik[20], O.T. Raitakari[21], L. Perusse[22], C. Ohlsson[23], M. Walker[24], S.M. Willems[14], R.H. Myers[25], J. Eriksson[26], V. Salomaa[27], H. Grallert[28], T.B. Harris[29], O. Pedersen[1], R.J. Loos[2, 30] for the LEPGen Consortium.* 1) Novo Nordisk Foundation Center for Basic Metabolic Research, University of Copenhagen, Copenhagen, Denmark; 2) MRC Epidemiology Unit, Institute of Metabolic Science, Cambridge, United Kingdom; 3) Department of Nutrition, Harvard School of Public Health, Boston, MA; 4) Channing Laboratory, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA; 5) University of Lausanne, Lausanne, Switzerland; 6) Public Health Genomics Unit, Department of Chronic Disease Prevention, National Institute for Health and Welfare, Helsinki, Finland; 7) Department of Twin Research & Genetic Epidemiology, King's College London, London, United Kingdom; 8) Department of Biostatistics, University of Washington, Seattle, WA; 9) Center for Biomedicine, European Academy Bozen/Bolzano (EURAC), Bolzano, Italy; 10) Peninsula College of Medicine and Dentistry, University of Exeter, Exeter, United Kingdom; 11) Medstar Research Institute, Baltimore, MD; 12) Longitudinal Study Section, National Institute on Aging, Baltimore, MD; 13) Division of Biostatistics, Washington University School of Medicine, St. Louis, MO; 14) Department of Epidemiology, Erasmus MC, Rotterdam, The Netherlands; 15) Department of Human Genetics, Leiden University Medical Center, The Netherlands; 16) Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom; 17) Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands; 18) Netherlands Consortium for Healthy Ageing, Leiden, The Netherlands; 19) Mannheim Institute of Public Health, Social and Preventive Medicine, Medical Faculty of Mannheim, University of Heidelberg, Mannheim, Germany; 20) Department or Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL; 21) Research Center of Applied and Preventive Cardiovascular Medicine, University of Turku and Department of Clinical Physiology, Turku University Hospital, Turku, Finland; 22) Division of Kinesiology, Department of Social and Preventive Medicine, Laval University, Ste-Foy, Quebec, Canada; 23) Center for Bone and Arthritis Research, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden; 24) Institute of Cellular Medicine, Newcastle University, Newcastle upon Tyne, United Kingdom; 25) Genome Science Institute, Boston University School of Medicine, Boston, MA; 26) Diabetes Prevention Unit, Department of Chronic Disease Prevention, National Institute for Health and Welfare, Helsinki, Finland; 27) Chronic Disease Epidemiology and Prevention Unit, Department of Chronic Disease Prevention, National Institute for Health and Welfare, Helsinki, Finland; 28) Institute of Epidemiology, Helmholtz-Zentrum München, Neuherberg, Germany; 29) Laboratory of Epidemiology, Demography, and Biometry, National Institute on Aging, National Institutes of Health, Bethesda, MD; 30) Mount Sinai School of Medicine, New York, NY.

Background: Leptin is an adipocyte-derived hormone for which the circulating levels correlate closely with body fatness. Leptin acts as a chronic signal to inform the brain about the stored body fat and as such it is involved in the regulation of long-term energy homeostasis. Complete leptin deficiency due to mutations in the leptin (*LEP*) gene causes severe obesity. We aimed to identify genetic variants associated with leptin levels in the general population. **Methods:** We performed a meta-analysis of associations between ~2.5 million SNPs and circulating leptin levels in 32,061 individuals from 22 studies and followed up SNPs in the six most significant loci ($P<10^{-6}$) in 20,278 additional individuals from 12 studies. To account for the effect of body fatness in the results, we performed the meta-analyses with and without adjustment for BMI. **Results:** We identified five loci reaching genome-wide significance ($P<5x10^{-8}$) for their association with leptin in the combined discovery and follow-up analyses. Variants near the *LEP*, *CCNL1*, and *SCL32A1* genes and in the *GCKR* gene showed a similar association with leptin between the BMI-unadjusted ($P=3x10^{-6}$, $P=1x10^{-9}$, $P=1x10^{-6}$, and $P=7x10^{-11}$, respectively) and BMI-adjusted ($P=1x10^{-13}$, $P=6x10^{-8}$, $P=2x10^{-8}$, $P=1x10^{-10}$, respectively) models. In contrast, variants in the *FTO* gene showed association with leptin in the BMI-unadjusted analysis only ($P=2x10^{-9}$); no association was observed after adjusting for BMI ($P=0.63$). **Conclusions:** Variation near the *LEP*, *CCNL1*, and *SLC32A1* genes and in the *GCKR* gene is associated with circulating leptin levels irrespective of the overall body fatness, whereas the association of the *FTO* locus with leptin is attributed to BMI. Interestingly, the near-*CCNL1* and *GCKR* loci have previously been reported to robustly associate with other metabolic or anthropometric phenotypes (near-*CCNL1* with birth weight; *GCKR* with e.g. lipids, glucose- and insulin levels) in previous meta-analyses of genome-wide association studies, whereas the near-*LEP* and near-*SLC32A1*loci have not been reported for association with such phenotypes. Further follow-up of the confirmed loci is ongoing. The identification of loci regulating leptin levels may provide new insights into the genetic regulation of obesity and metabolism.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2117F

Pathway-based analysis of genome-wide association studies for the personality traits. *H. Kim[1], S. Roh[1], C. Hong[4], E. Lee[4], Y. Sung[2], H. Lee[2], H. Chung[3], J. Lee[4], H. Kim[1].* 1) Dept Biochem, Sch Med, Ewha Womans Univ, Seoul, South Korea; 2) Dept Internal Medicine, Sch Med, Ewha Womans Univ, Seoul, South Korea; 3) Dept Obstetrics and Gynecology, Sch Med, Ewha Womans Univ, Seoul, South Korea; 4) Center for Genome Science, Korea National Institute of Health, Korea Centers for Disease Control and Prevention, Cheongwon.

Genome-wide association studies (GWAS) are now used as a major tool to identify SNPs associated with complex traits. In the initial GWAS, only a few SNPs within or near the *PTPRD*showed significant association and other loci showed the moderate association with personality. Single-marker-based analysis will miss many SNPs with moderate genetic effects. We performed pathway-based analysis on GWAS data sets of the Five Factor Model (FFM) personality trait to overcome these limitations and to find missing heritability of personality. We found pathways and gene sets with statistical significance of FDR< 0.05. More than 10 pathways associated with behavior and neurotransmitter pathway showed highly significance for Neuroticism, Extraversion, Openness, and Agreeableness, respectively. They include the association of Neuroticism with axon guidance (FDR< 1×10^{-5}), Extraversion with toll like receptor signaling pathway, nicotinate and nicotinamide metabolism and exon guidance (FDRs< 1×10^{-5}), Openness with adrenergic and monoamine G protein-coupled receptors (FDRs< 1×10^{-5}), and Agreeableness with adrenergic and monoamine G protein-coupled receptors (FDRs< 1×10^{-5}). Conscientiousness wasn't shown significant association. Personality trait is polygenic and at least 10,000 genes expressed in brain, so it is necessary to consider how many genes might conceivably influence personality traits. Interestingly, we also found that some significant pathways were overlapped in four personality traits regardless factor. Therefore we suggest that there may be common pathways associated with general personality which is not limited factors of personality. This research was supported by NRF (2010-0026606) and KNIH (4845-301).

2118W

A genome-wide association study identifies a novel susceptible locus for refractive error at 19q13.12 in Korean. *W.R. Kim[1], T.Y. Chung[2], Y.M. Song[3], K. Lee[4], J. Sung[1].* 1) Department of Epidemiology, School of Public Health and Institute of Health and Environment, Seoul National University, Seoul, Korea; 2) Department of Ophthalmology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea; 3) Department of Family Medicine, Samsung Medical Center and Center for Clinical Research, Samsung Biomedical Research Institute, Sungkyunkwan University School of Medicine, Seoul, Korea; 4) Department of Family Medicine, Busan Paik Hospital, Inje University College of Medicine, Seoul, Korea.

Introduction: Refractive error is the most common ocular disorder worldwide and the prevalences of myopia in Asians are much higher than those in Caucasians. In Korea, the prevalence of myopia is 53.7±0.6%. Although environmental factors such as sustained near visual work, higher education level and lack of outdoor activities have been reported to increase the risk of myopia, myopia is a highly heritable trait (h2=0.60~0.90) but the mechanism of myopia is not fully understood. Previous genome-wide association studies in Asian countries have identified risk variants for refractive error; in a Japanese population, 11q24.1 were linked with high myopia, which was not replicated in Chinese; in a Chinese population, 13q12.12 and 4q25 were associated with high myopia; in a Singapore Chinese population, variants at 5p15 were suggested. In Europeans whose myopia prevalence is generally lower than that of Asians, 15q14 and 15q25 were identified to be responsible for refractive error. To add to and compare with the existing body of evidence, we performed a GWA study involving the general families in Korea.**Method**: Among 3461 adult individuals (1493 men) who have participated in the Healthy Twin study Korea, full eye examination was undertaken on 1960 individuals. Refractive error was measured by Auto-Refractor (Canon Full Auto Ref-keratometer RK-F1, Japan). The spherical equivalent was calculated from the standard formula: spherical equivalent = sphere + (cylinder/2). We investigated the spherical equivalent as a quantitative trait. Individuals with history of corneal surgery such as cataract, LASIK and LASEK were excluded. The Affymetrix GeneChip Array version 6 (1 million SNPs) platform was used for generating genetic marker information. A GWA study was conducted by combining p-values of within-family information and population mean effects of between-families. Complete data on refractive error were available on 1022 persons with 537158 SNPs. **Result**: We identified a significant association between refractive error and a variant at 19q13.12 (P=7.95E-8). On the 8q21.11, loci also showed a suggestive association (P=4.18E-6). In a European population, 15q14 was reported to be a susceptibility locus for refractive error measured as a quantitative trait. In our study, rs688220 on 15q14 was not significantly associated with refractive error. (Psmallest=0.20) **Conclusion**: Our data suggests that a susceptible locus at 19q13.12 would be associated with refractive error.

2119T

A Genome-Wide Association Study on Hallux Valgus Angle (HVA) identifies candidate loci: The Healthy Twin Study Korea. *SJ. Lee[1], D. Lee[1], JH. Hwang[2], YM. Song[3], K. Lee[4], J. Sung[1].* 1) Department of Epidemiology, Graduate School of Public Health and Institute of Environment and Health, Seoul National University, Seoul, South Korea; 2) Department of Physical & Rehabilitation Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, South Korea; 3) Department of Family Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, South Korea; 4) Department of Family Medicine, Busan Paik Hospital, Inje University College of Medicine, Busan, Korea.

Introduction: Hallux valgus (HV) is one of the most common chronic structural foot deformities known to be associated with genetic predisposition, age, sex and constricting footwear. Genetic influences on foot structure, particularly HV, play an important role, with the heritability of 0.55. However, specific genetic variants underlying HV have not yet been identified. The objective of this study is to investigate common genetic variants that confer susceptibility to the degree of HVA a genome-wide association (GWA) study. **Methods**: Among the 3461 participants of The Healthy Twin Study Korea, 837 individuals (312 men, 525 women, aged 44.8 ± 13.7 years), who have taken the weight-bearing anteroposterior (AP) and lateral foot X-ray examinations with genetic information (Affymetrix GeneChip v6) were included. Hallux valgus angle (HVA) was measured from each foot X-ray film, and HVA exceeding 20 degree was defined as having HV. Information about narrow-toed high-heeled shoe use was collected through a self-reported questionnaire. A meta-analysis of GWA results was executed by combining the p-values of population mean effects using founders between families (PLINK) and those of generalized transmission disequilibrium test based on familial relations (FBAT). **Results**: We identified potential candidate loci at 5q34 (p=2.86 x 10-7), 7q11.21 (p=3.77 x 10-7), and 4p16.1 (p=2.53 x 10-6). These regions comprise a gene related to limb and bone formation in Drosophila, which may be a potential target not only for HV, but general foot development and disorders. **Conclusion**: Our study could successfully identified potential genetic loci influencing HV, of which functions are of potential interests in foot development. Our findings are waiting for further validation through replications studies, however, it is challenging because of the lack of compatible measurements in other studies.

2120F

Sciatica in young adult population: A genome-wide association study in the Young Finns. *S. Lemmelä[1], S. Solovieva[1], R. Shiri[1], T. Lehtimäki[2,3], I. Seppälä[2,3], M. Kähönen[2,3], M. Juonala[4,5], J. Viikari[4,5], O. Raitakari[4,5], E. Viikari-Juntura[1], K. Husgafvel-Pursiainen[1].* 1) Finnish Institute of Occupational Health, Helsinki, Finland; 2) Tampere University Hospital, Tampere, Finland; 3) University of Tampere, Tampere, Finland; 4) University of Turku, Turku, Finland; 5) Turku University Hospital, Turku, Finland.

Sciatic pain is one of the leading causes of disability in the working age population. Epidemiological data suggest a strong genetic component and complex genetic etiology. This study is part of the research consortium focused on musculoskeletal disorders (MSD) and risk factors MSDs may share with cardiovascular diseases. The current study is based on the follow-up of the Young Finns Study carried out in 2007 (subjects 30–45 years), where the question on sciatica diagnosed by physician was included. A Genome wide association study (GWAS) was conducted in 2442 Finnish individuals (46% men, 54% women) using custom-built Illumina 670K beadcip at Wellcome Trust Sanger Institute (UK). Study sample comprised 180 Finnish sciatica cases and 1840 unaffected controls. A total 546 674 SNP's passed quality control and were analysed for association using PLINK. Further imputation to ~23 million SNPs (1000 Genomes imputation reference) was performed using IMPUTE. After excluding SNPs with minor allele frequency <0.01 and/or with imputation quality <0.7, ~8 million SNPs were available for the analysis. The GWA analysis was performed using SNPTEST with additive genetic model and adjustment for age and parental socioeconomic status based on occupation. The locus 1q32.2 was the most significantly associated with sciatica ($p=8.7 \times 10^{-7}$) in the genotyped SNP set. After imputation, in total 21 additional SNPs in three loci (4q23, 6p25.1 and 21q22.3) were identified with p-values below genome wide significance level ($p<5 \times 10^{-8}$). The analysis further indicated that partially different SNPs were associated with sciatica in men and women. In women, 2q21.2 and 1q32.2 were the most significantly associated loci with the p-value level of 10^{-6}. After imputation, 29 additional SNPs on 4q28.3 and 6p25.1 showed associations with genome-wide significance ($p<5 \times 10^{-8}$). In men, locus 1q43 yielded significant p-values in genotyped SNP set; in total 21 additional SNPs in the same locus and 25 SNPs on the locus 16q23.2 showed association ($p<5 \times 10^{-8}$) after imputation. In summary, our preliminary GWAS results identified a number of novel candidate regions for sciatica that warrant further investigation. Replication of these new association signals in GWAS in another Finnish population is ongoing. The study received financial support from the Academy of Finland (project no 129364, MSDs@Lifecourse, SALVE).

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2121W

GWAS of Blood Cell Traits Identifies Novel Associated Loci and Epistatic Interactions in Caucasian and African American Children. *J. Li[1], J.T. Glessner[1], H. Zhang[1], C. Hou[1], Z. Wei[2], J.P. Bradfield[1], F.D. Mentch[1], Y. Guo[1], C. Kim[1], Q. Xia[3], R.M. Chiavacci[1], K.A. Thomas[1], H. Qiu[1], S. Grant[3], S.L. Furth[4], H. Hakonarson[1,3,4], P.M.A. Sleiman[1,3,4].* 1) Center for Applied Genomics, The Children's Hospital of Philadelphia, Philadelphia, PA; 2) Department of Computer Science, New Jersey Institute of Technology, Newark, NJ; 3) Division of Human Genetics, The Children's Hospital of Philadelphia, Philadelphia, PA; 4) Department of Pediatrics, University of Pennsylvania School of Medicine, Philadelphia, PA.

Hematological traits are important clinical indicators. Commonly measured hematological traits include white blood cell count (WBC), red blood cell count (RBC), hemoglobin concentration (Hgb), percentage of erythrocyte cells (Hct), mean corpuscular hemoglobin (MCH), mean corpuscular hemoglobin concentration (MCHC), mean corpuscular volume (MCV) and platelet count (PLT). More than 100 loci have been reported to be associated with hematological quantitative traits in populations of European ancestry and East Asian ancestry. However, the genetic determinants of hematological traits have not been fully investigated in African Americans. We carried out a GWA study of the eight common hematological traits in two pediatric cohorts which consist of 7943 African American and 6234 Caucasian children respectively, followed by meta-analysis of both cohorts. In African Americans, we replicated four previously reported loci at genome wide significance level. In addition, we identified five novel associations, including association of HBE1 variants with HCT and MCHC, the alpha globin gene cluster variants with RBC and MCHC, and an ARHGEF3 variant with PLT. In Caucasians, we report novel significant association between COPZ1 variants and PLT as well as replication of five previously reported loci at genome wide significance. In total, we replicated 95 out of 107 previously reported loci at nominal significance following meta-analysis of both cohorts. Finally, through detailed analysis of the association between MCH and variants at the alpha globin gene cluster in African Americans, we not only identified variants with independent effects on MCH at the locus but also show epistatic interaction between them. Individuals homozygous for the minor alleles of the SNPs that show epistasis had extremely low MCH values and were at significantly increased risk of iron deficiency anemia.

2122T

A genome-wide analysis of longitudinal forced expiratory volume in one second in the framingham heart study. *S.-Y. Liao, X. Lin, D. C. Christiani.* Harvard School of Public Health, Boston, MA.

Background: Chronic Obstructive Pulmonary Disease (COPD) has become a globally prevalent disease, with occurrence between 6% and 20% worldwide. The Forced Expiratory Volume in One Second (FEV1) is a stable measure of lung function and has been shown to predict clinical outcome, grade the severity, and follow up the natural history of disease. In addition to the well-known environmental risk factor, tobacco smoking, many studies have shown that genetic factors also play an important role in COPD. COPD is founded to aggregate in families, and the spirometric measurements of pulmonary function are also heritable. Previous studies always focused on cross-sectional FEV1, in the study, we explore the association between the quantitative trait (FEV1) and SNPs using a longitudinal genome-wide analysis. With longitudinal analysis, we are able to characterize aging and normal development, as well as growth or decline of the FEV1. In addition, we can observe individual patterns of change. Methods: The study population is the Offspring cohort from Framingham Heart Study. A total of 2,238 participants have both genotype and spirometry measurements with 9032 observations (around 4 measurements per person) and the mean follow-up year is 15.77 years. In the study, we used participants' 3, 5, 6, 7, 8 examination of FEV1 as the outcomes. Age, height (inches), pack-years, and smoking status at the point of the each examination were used as time-varying covariates. The longitudinal genome wide analysis used for individual SNPs was a linear mixed effect model for association between FEV1 and allele “dose” with adjustment for these covariates. We used family random intercept and individual random intercept/random age slope for the issue of both within and between subjects' correlation. Results: In our longitudinal genome wide analysis, we identified several genome-wide significant risk loci. The four top significant SNPs are located at chromosome 17q25 (rs3924256, P = 4.67 x 10-9), chromosome 14q32.11 (rs4904670, P=3.64 x 10-8), chromosome 6q25.1 (rs9397997, P=1.15 x 10-7), chromosome 9q34.3 (rs3125786, P=6.22 x 10-7) Conclusions: Longitudinal analysis is a more powerful way to identify disease risk SNPs. These SNPs are worthy of further investigation. We will attempt to replicate the analysis in other cohorts, though to date there are few longitudinal cohort studies with genetic data.

2123F

Hundreds of loci contribute to body fat distribution and central adiposity. *A.E. Locke[1], D. Shungin[2,3,4], T. Ferreira[5], T.W. Winkler[6], D.C. Croteau-Chonka[7], R. Magi[5,8], T. Workalemahu[9], K. Fischer[8], J. Wu[10], R.J. Strawbridge[11], A. Justice[12], F. Day[13], N. Heard-Costa[14,15], C.S. Fox[14], M.C. Zillikens[16], E.K. Speliotes[17,18], H. Völzke[19], L. Qi[9], I. Barroso[20,21], I.M. Heid[6], K.E. North[12], P.W. Franks[2,4,9], M.I. McCarthy[22], J.N. Hirschhorn[23], L.A. Cupples[10,14], E. Ingelsson[24], A.P. Morris[5], R.J.F. Loos[13,25], C.M. Lindgren[5], K.L. Mohlke[7], Genetic Investigation of ANthropometric Traits (GIANT) Consortium.* 1) Department of Biostatistics and Center for Statistical Genetics, University of Michigan, Ann Arbor, MI; 2) Genetic and Molecular Epidemiology Group, Department of Public Health and Clinical Medicine, Umeå University, Umeå, Sweden; 3) Department of Odontology, Umeå University, Umeå, Sweden; 4) Department of Clinical Sciences, Skåne University Hospital, Lund University, Malmö, Sweden; 5) Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK; 6) Regensburg University Medical Center, Department of Epidemiology and Preventive Medicine, Regensburg, Germany; 7) Department of Genetics, University of North Carolina, Chapel Hill, NC; 8) Estonian Genome Center, University of Tartu, Estonia; 9) Department of Nutrition, Harvard School of Public Health, Boston, MA; 10) Department of Biostatistics, School of Public Health, Boston University, Boston, MA; 11) Cardiovasvular Genetics and Genomics Group, Karolinska Institutet, Stockholm Sweden; 12) Department of Epidemiology and Carolina Center for Genome Sciences, University of North Carolina, Chapel Hill, NC; 13) MRC Epidemiology Unit, Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge, UK; 14) National Heart, Lung, and Blood Institute, Framingham, MA; 15) Department of Neurology, Boston University School of Medicine, Boston, MA; 16) Department of Internal Medicine, Erasmus MC Rotterdam, the Netherlands; 17) Department of Internal Medicine, Division of Gastroenterology, and Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI; 18) Broad Institute, Cambridge, MA; 19) Institute for Community Medicine, Ernst-Moritz-Arndt-University Greifswald, Greifswald, Germany; 20) Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, UK; 21) University of Cambridge Metabolic Research Labs, Institute of Metabolic Sciences,; 22) University of Oxford, Oxford, UK; 23) Department of Genetics, Harvard Medical School, Boston, MA; 24) Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden; 25) Charles R. Bronfman Institute of Personalized Medicine, Child Health and Development Institute, Department of Preventive Medicine, Mount Sinai School of Medicine, New York, NY.

Central adiposity and body fat distribution are risk factors for type 2 diabetes and cardiovascular disease and can be measured using waist circumference (WC), hip circumference (HIP), and waist-to-hip ratio (WHR). Adjusting for body mass index (BMI) differentiates effects from those for overall obesity. We performed fixed effects inverse variance meta-analysis for these traits with 72,919 individuals from 30 studies in a prior genome-wide association study (GWAS) meta-analysis, 71,139 individuals from 24 additional GWAS, and 67,163 individuals from 28 studies genotyped on Metabochip by the GIANT consortium. We identified 48 independent genome-wide significant ($p<5x10^{-8}$) associations for WHR adjusted for BMI, including all 14 previously published signals. Twelve signals are located near genes for transcription factors, including developmental homeobox-containing proteins. Among them, two are in the *HOXC* gene cluster near *HOXC8* and *miR-196a2*. *HOXC8* is expressed in white adipose tissue and is a regulator of brown adipogenesis, while miR-196a inhibits *Hoxc8* expression. Signals are located near*PPARG*, encoding a transcription factor known to regulate adipocyte differentiation, and near *HMGA1* and *CEPBA*, encoding transcription factors that act downstream of insulin receptor and leptin signaling, respectively. Further novel signals are located near genes involved in angiogenesis (*PLXND1, VEGFB,* and *MEIS1*). Among the other five traits, we estimate that a significant proportion of the genetic effects for WC and HIP adjusted for BMI are correlated with height (0.59, $p<5x10^{-79}$ and 0.83, $p<2x10^{-40}$, respectively). Despite this strong correlation, an appreciable proportion of the genetic contributions to these traits will be independent of height. Association meta-analysis for the five additional traits identified an additional 148 independent signals ($p<5x10^{-8}$), 32 of which have not been reported previously for an anthropometric trait. These novel signals suggest regulation of adipose gene expression (*KLF14*) and transcriptional control of cell patterning and differentiation in early development (*HLX, SOX11, ZNF423,* and *HMGXB4*) affect fat distribution. Meta-analyses for WHR, WC, and HIP, with and without adjustment for BMI, identified a total of 196 independent loci, 66 novel, affecting fat deposition and body shape, and implicating genes involved in development, adipose gene expression and tissue differentiation, response to metabolic signaling, and angiogenesis.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2124W

Significant evidence for association in genome-wide analyses of lipid levels in Pima Indians. *A. Malhotra, S. Kobes, W.C. Knowler, L.J. Baier, R.L. Hanson.* Phoenix Epidemiology and Clinical Research Branch, NIDDK, NIH, Phoenix, AZ.

Genome-wide association studies (GWAS) for lipid levels have been performed in multiple ethnic groups including Caucasians, African Americans, and Asians. However, to date, no GWAS in American Indians has been published. We performed GWAS analyses for total cholesterol (TC; N= 1135), low-density lipoprotein cholesterol (LDL; N=965), high-density lipoprotein cholesterol (HDL; N=976), and triglycerides (TG; N=970) using longitudinal Pima Indian data. Using SAS (Statistical Analysis Software), a mixed model was used which incorporates multiple tests per individual (ranging from 1 to 19 exams). We performed association analyses for 454,194 SNPs with these lipid traits. One SNP (rs1532625) reached genome-wide significance (estimated at 4.94×10^{-7}) for HDL (p-value: 1.53×10^{-7}). The *cholesterylester transfer protein* gene is located close to this SNP; the protein facilitates lipid transport. Analysis of TC (three SNPs) and LDL (two SNPs) identified a region on chromosome 1 (109,619,829–109,623,689bp) with p-values ranging from 2.48×10^{-9} to 4.72×10^{-8}. This region has been identified in other ethnic groups. Furthermore, a cluster of candidate genes is present at this locus, of these most notably, the *sortilin-1* gene. A previously published study of *sortilin-1* showed an association between increased expression levels in the liver and decreased LDL levels. For TG, seventeen SNPs on chromosome 11 (116,024,099–116,585,850bp) showed significant evidence for association, with p-values ranging from 5.86×10^{-11} to 4.35×10^{-7}. The *apolipoprotein A-V* gene is located in this region, which has been identified in previous TG association studies. The current GWAS in Pima Indians provides significant evidence for association that replicates previously identified regions with strong candidate genes that might affect lipid levels. Replication and fine-mapping studies are required to confirm these associations and to identify potential functional variants.

2125T

Investigation of Lipid Pathway Genes in Late Onset Alzheimer's Disease. *L.L. McClain[1], X. Wang[1], M.M. Barmada[1], F.Y. Demirci[1], O.L. Lopez[2,3], M.I. Kamboh[1,2].* 1) Department of Human Genetics, University of Pittsburgh, Pittsburgh, PA; 2) Alzheimer's Disease Research Center, University of Pittsburgh, Pittsburgh, PA; 3) Department of Neurology, University of Pittsburgh, Pittsburgh, PA.

Alzheimer's disease (AD) is the most common type of dementia and affects about 5.4 million Americans with symptoms of memory loss, depression, and profound neurological atrophy in the advanced stage. Recent genome-wide association studies have identified 10 loci for late-onset of AD, including *APOE, PICALM, CR1, BIN1, CLU, CD2AP, CD33, ABCA7, EPHA1, MS4A4*, and *MS4A6E*. Three of these genes are involved in cholesterol metabolism (*APOE*, *CLU*, and *ABCA7*), indicating that cholesterol-related genes may play an important role in the etiology of AD. In this study, we examined the association of 4,838 SNPs in 212 cholesterol-related genes in 1291 AD cases and 958 controls, adjusted for age, sex, and principal components. Analyses were carried out using PLINK software. In the single-locus analysis, 225 SNPs in 75 genes revealed significant associations (P<0.05) with *APOE* /rs429358 being the top hit (P=2.49E-53). The five other top hits were: *CETP*/ rs9930761 (P=5.57E-04), *PCTP*/ rs2912558 (P=6.02E-04), *PPP1R3B*/ rs6998837 (P=6.75E-04), *UBASH3B*/ rs1894084 (P=7.09E-04), and *ABCA7*/ rs4147920 (P=8.12E-04). Three of the 75 significant genes (*ABCA1, UBASH3B*, and *CUBN*) had 24% of the significant SNPs. In the haplotype analysis, our most significant result was in the *ABCA7* gene (P=1.16E-04). Our data reinforce the importance of the cholesterol-related genes in AD pathology and resequencing of these genes may identify rare and common functional SNPs.

2126F

Genome-wide association study identifies genetic loci associated with iron overload in HFE hemochromatosis. *J. Mosser[1], R. Bouvet[2], J. Morcet[3], M. Perrin[3], A.M. Jouanolle[4], M.P. Roth[5], E. Génin[6], Y. Deugnier[3], M. de Tayrac[1].* 1) CNRS, UMR 6290, Laboratoire de génomique Médicale, CHU Rennes, France; 2) Plateforme Génomique Santé Biogenouest Rennes Biogenouest, Rennes; 3) Centre d'investigation clinique (CIC INSERM 0203) Hôpital de Pontchaillou, Rennes; 4) Laboratoire de Génétique Médicale, Hôpital de Pontchaillou, Rennes; 5) Inserm U563, CHU Purpan, Toulouse, France; 6) INSERM U946, Fondation Jean Dausset - CEPH, Paris.

In Humans, hemochromatosis, the most common form of genetic iron loading disease, may result in severely damaging the liver, heart, endocrine organs, bones and articles. HFE hemochromatosis accounts for more than 95% of hemochromatotic phenotypes in populations of European descent and is chiefly related to the HFE C282Y/C282Y (p.Cys282Tyr/p.Cys282Tyr) genotype. C282Y homozygosity is highly frequent in Europe (up to 1‰ in Northern Europe). However, its expressivity is variable in terms of both iron burden and iron-related organ damage and, as a whole, its clinical penetrance is low, ranging from 1% in females to 25% in males. Previous studies suggest that, in addition to environmental factors, a genetic component could explain a substantial part of such a phenotypic variation. However, a few genetic factors have been identified so far. We thus search for such modifier genes studying associations between common variants and hemochromatosis penetrance defined using iron burden and/or iron-related organ damage indices. A cohort of 500 unrelated C282Y homozygotes were genotyped using a high-throughput process based upon tagSNP-CNV arrays (Human660W-Quad; Illumina). For these patients, four biological variables and two clinical variables were studied: amount of iron removed, serum ferritin, transferrin saturation, serum iron, disease staging and fibrosis. The GenABEL R package (http://www.r-project.org/) was used to perform genetic association analyses on common variants (MAF > 5%) considering additive models for both quantitative and qualitative traits. These models were adjusted for age, gender. Population stratification was also taken into account by adjusting models for the two principal components issued from the PCA of the whole genotype data. These analyses have identified a catalog of 30 SNPs. Among these markers, our results replicate the SNP rs3811647 within intron 11 of *TF* gene ($p=1.10^{-6}$) and identify new genes likely to modulate iron overload. Specifically, 3 loci (4p16.2, 10q26.12 et 16q21) have been identified associated with ferritin and amount of iron removed ($p<10^{-5}$). The replication phase is currently performed on an independent cohort of 700 patients and is will allow the selection of the polymorphisms most associated with hemochromatosis expressivity. These results may impact the development of new means for preventing, diagnosing and treating hemochromatosis and other iron-related disorders, notably iron deficiency.

2127W

Not a stretch: Variant near the dermal gene elastin is associated with stretch marks. *M.E. Mullins, N. Eriksson, A.K. Kiefer, J.Y. Tung.* 23andMe, Inc., Mountain View, CA.

Striae distensae, or stretch marks, are a common skin condition that appear as red or white lines on the skin. Though stretch marks are only harmful in extreme cases, even mild stretch marks can cause distress. The portions of cosmetic/pharmaceuticals and maternity-care industries that offer stretch mark prevention and treatment products are estimated to be worth $5–10 billion annually.

Despite the range of products claiming to prevent and treat stretch marks, the causes of stretch marks are not well understood. Skin distension, such as that which occurs during pregnancy or puberty; prolonged exposure to cortisols, such as in individuals with Cushing syndrome; and genetics have all been implicated. A small number of monogenic diseases, including Marfan syndrome and Ehlers-Danlos syndrome, count stretch marks amongst their characteristics. Both these syndromes are caused by mutations in genes that encode dermal proteins. To date, no genes are known to be associated with the stretch marks that afflict the general population.

We conducted a genome-wide association analysis of stretch marks using approximately 7 million SNPs imputed from about 1 million SNPs genotyped. Phenotypic data was self-reported by over 30,000 23andMe customers who responded to the following question on the 23andMe website: “Do you have stretch marks on your hips, thighs, or backs of your arms? Yes/No/I'm not sure”. We found a genome-wide significant association (p-value 1.28E-12) with rs6946106, which is upstream of elastin (ELN). Defects in ELN are known to affect the integrity of the skin and cause aortic defects.

We also detected a second association (p-value 3.86E-8) with rs35318931, a missense mutation in sushi-repeat containing protein, X-linked (SRPX). These results show a genetic component to stretch marks: defects in ELN may influence skin remodeling after stress. Many stretch mark sufferers spend hundreds of dollars on stretch mark treatments that offer minimal results. A better understanding of the etiology of stretch marks could help develop more targeted and effective treatment or prevention strategies.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2128T

Identification of three new genes associated to serum Butyrylcholinesterase activity by GWAS in two genetically distant isolated populations. *F. Murgia[1], L. Portas[1], S. Ulivi[3], N. Pirastu[2], S. Vaccargiu[1], D. Parracciani[4], P. Gasparini[2], M. Pirastu[1].* 1) Institute of Population Genetics, CNR, Sassari, Italy; 2) Institute for Maternal and Child Health - IRCCS "Burlo Garofolo" - Trieste, University of Trieste, Trieste, Italy, 34137; 3) Institute for Maternal and Child Health - IRCCS "Burlo Garofolo" - Trieste, Trieste, Italy, 34137; 4) Genetic Park of Ogliastra, Perdasdefogu, Italy.

Introduction: Great interest in the physiological role of BuChE has been given since it was demonstrated the correlation of its serum activity levels with Metabolic Syndrome (MS). Many variants in the BuChE locus have been described which influence the serum activity level and more recently other QTL were identified by GWAS in a large cohorts of outbreed populations. However it is not known the genetic link between this enzyme and the MS. **Objectives**: To investigate the phenotypic correlation between serum BuChE activity levels and MS endophenotypes and to identify genetic determinants of the trait in homogeneous populations. **Methods**: We analyzed 2,500 participants from two Italian isolated populations representing the Ogliastra (N=1272) and Friuli Venezia Giulia (N=1226) Genetic Parks. A comparison of continuous variables was performed by ANOVA. Factor Analyses to illustrate the relationship between BuChe activity and other variables was performed using Principal Components method. GWAS was carried out in each cohort with 2.5M genotyped or imputed SNPs using a mixed linear regression model taking into account sex, age, BMI and genomic kinship. In conditional analyses SNPs with the GWAS lowest P values were selected and included as covariates in the linear regression. Meta-analyses were conducted using the inverse variance weighted method. **Results**: A significant (P<0.05) phenotypic correlation was detected between BuChE levels and BMI, LDL and HDL cholesterol, fasting glucose, triglycerides, diastolic and systolic blood pressure, MS (ATPIII). Factor analysis confirmed these results. We identified five genome wide significant loci: two (rs1803274 BuChE K variant P=4.7e-68 on 3q26 and rs11686036 ABI2-RAPH1P=2.2e-9 on 2q33) described in recently GWAS and three (rs2911262 ANKRD11 P=6.1e-9 on 16q24, rs1244509 TAF3 P=3.4e-8 on 10p15 and rs4308478 SPOCK1 P=1.7e-8 on 5q31) novel. Some of the several suggestive loci reached a significant P value in meta-analysis with previously published GWAS data. We tested the effects of loci previously reported to affect MS components such as GALNT and LPIN (P<0.01 bonferroni corrected). **Conclusions**: We confirmed a strong correlation of BuChE activity with MS. We add new genetic data which may help to better understand the BuChE pathway complexity. Additional Next-gen sequencing and functional studies are required to elucidate the role of this increasingly interesting enzyme in physiology and disease.

2129F

Common genetic variants of primary open-angle glaucoma in Japanese population. *M. Nakano[1], Y. Ikeda[2], Y. Tokuda[1], M. Fuwa[1], R. Sato[1], N. Omi[1], H. Adachi[1], M. Ueno[2], K. Mori[2], S. Kinoshita[2], K. Tashiro[1].* 1) Dept Genomic Med Sci, Kyoto Prefectural Univ Med, Kyoto, Japan; 2) Dept Ophthalmol, Kyoto Prefectural Univ Med, Kyoto, Japan.

Purpose: Glaucoma is a neurodegenerative ocular disease and one of the leading causes of irreversible blindness worldwide. Although it has been suggested that genetic factors play an important role for the pathogenesis and it would be a great benefit if the risk of developing glaucoma could be predicted based on a simple blood test assessing the genetic markers, only a small portion of the genetic variation for glaucoma has been elucidated. Therefore, in order to identify authentic common variants for the major type of glaucoma, primary open-angle glaucoma (POAG), in Japanese, we performed two genome-wide association studies (GWAS), and additionally analyzed the variants of candidate genes, some of which were uncovered by the GWAS.

Methods: We performed a GWAS by analyzing 653,519 autosomal common variants in 833 POAG patients and 686 controls after the quality control. We also combined the data set with those from 411 POAG patients and 289 controls of our previous GWAS by Mantel-Haenszel test. We then constructed a custom DNA chip with the probes to detect 612 autosomal variants of 24 glaucoma related genes that have been reported previously, and genotyped the genomic DNA derived from 521 POAG patients and 519 controls for the association study.

Results: As for the GWAS, we identified 5 variants that passed the Bonferroni correction in *CDKN2B-AS1* on chromosome 9p21. All of the significant variants whose probes were designed in the microarrays of both GWAS showed stronger association with POAG after combining the two data sets. Moreover, some of the variants from candidate genes, including the four variants in *WDR36*, showed suggestive association (P<0.05) for POAG in our population.

Conclusions: In this study, we have successfully identified the reliable variants in 9p21 locus strongly associated with POAG by means of GWAS. Since glaucoma has been regarded as a typical multifactorial disease, it should be important to continue the comprehensive identification of genetic variants that could explain the disease etiology and apply them to the glaucoma diagnosis.

2130W

The associations of HLA-DP locus with chronic hepatitis B and viral clearance are widely replicated in East-Asian populations. *N. Nishida[1,2], Y. Tanaka[3], H. Sawai[2], Y. Mawatari[1,2], M. Yamaoka[2], A. Koike[4], K. Matsuura[3], M. Sugiyama[1], K. Murata[1], M. Korenaga[1], N. Masaki[1], K-H. Han[5], K. Tokunaga[2], M. Mizokami[1].* 1) Research Center for Hepatitis and Immunology, NCGM, Ichikawa, Chiba, Japan; 2) Department of Human Genetics, Graduate School of Medicine, University of Tokyo, Bunkyo-ku, Tokyo, Japan; 3) Department of Virology & Liver unit, Nagoya City University Graduate School of Medical Sciences, Nagoya, Japan; 4) Central Research Laboratory, Hitachi Ltd., Kokubunji, Japan; 5) Department of Internal Medicine, Yonsei University College of Medicine, Seoul, South Korea.

Overall, one-third of the world's population (2.2 billion) is infected with hepatitis B virus (HBV), and about 15% of these are chronic carriers. About 75% of the chronic carriers live in the east-south Asia and east pacific area, and there are 1.3–1.5 million chronic carriers living in Japan. Of chronic carriers, 10–15% develop liver cirrhosis (LC), liver failure and hepatocellular carcinoma (HCC), and the remaining individuals eventually achieve a state of nonreplicative infection, resulting in hepatitis B surface antigen (HBsAg) negative and hepatitis B core antibody (anti-HBc) positive, i.e. HBV-resolved individuals. We performed a GWAS using samples from Japanese HBV carriers, healthy controls and spontaneously HBV-resolved individuals in order to confirm or identify the host genetic factors related to CHB and viral clearance. In the subsequent replication analysis, we validated the associated SNPs in the GWAS using two independent sets of Japanese and Korean individuals. In our study, healthy controls were randomly selected with clinically no evidence of HBV exposure, therefore, HBV-resolved individuals were prepared to clearly identify the host genetic factors related with CHB or HBV clearance. Association analysis in the Japanese and Korean data identified the *HLA-DPA1* and *HLA-DPB1* genes with $Pmeta = 1.89 \times 10^{-12}$ for rs3077 and $Pmeta = 9.69 \times 10^{-10}$ for rs9277542. We also found that the *HLA-DPA1* and *HLA-DPB1* genes were significantly associated with protective effects against chronic hepatitis B (CHB) in Japanese, Korean and other Asian populations, including Chinese and Thai individuals ($Pmeta = 4.40 \times 10^{-19}$ for rs3077 and $Pmeta = 1.28 \times 10^{-15}$ for rs9277542). These results suggest that the associations between the HLA-DP locus and the protective effects against persistent HBV infection and with clearance of HBV were replicated widely in East Asian populations; however, there are no reports of GWAS in Caucasian or African populations. Based on the GWAS in this study, there were no significant SNPs associated with HCC development. To clarify the pathogenesis of CHB and the mechanisms of HBV clearance, further studies are necessary, including functional analyses of the HLA-DP molecule. URLs. The results of the present GWAS will be registered at a public database: https://gwas.lifesciencedb.jp/cgi-bin/gwasdb/gwas_top.cgi.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2131T

Three Novel Loci Identified with BMI in a Genome-wide Association Study of 47,098 Men and Women of African ancestry. *K.E. North[1], K.L. Monda[2], G.K. Chen[3], K.T. Taylor[4], L.A. Lange[5], C. Palmer[6], R.J.L. Loos[7], A.P. Reiner[8], D.R. Velez Edwards[9], M.C. Ng[10], T.L. Edwards[11], C. Kooperberg[12], B.E. Henderson[13], G.J. Papanicolaou[14], J.N. Hirschhorn[15], C. Haiman[13], African American BMI Consortium.* 1) Department of Epidemiology, University of North Carolina, Chapel Hill, NC; 2) The Center for Observational Research, Amgen, Inc. Thousand Oaks, CA and Department of Epidemiology, University of North Carolina, Chapel Hill, NC; 3) Department of Preventive Medicine, Keck School of Medicine and Norris Comprehensive Cancer Center, University of Southern California, Los Angeles, CA; 4) Department of Epidemiology and Population Health, University of Louisville, Louisville, KY; 5) Department of Genetics, University of North Carolina at Chapel Hill School of Medicine, Chapel Hill, NC; 6) Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA; Divisions of Genetics and Endocrinology, Children's Hospital, Boston, MA; 7) Charles R. Bronfman Institute for Personalized Medicine, Mount Sinai School of Medicine, New York, NY & MRC Epidemiology Unit, Institute of Metabolic Science, Addenbrooke's Hospital, CB2 OQQ, Cambridge, UK; 8) Public Health Science Division, Fred Hutchinson Cancer Research Center, Seattle, WA; 9) Center for Human Genetics Research, Vanderbilt Epidemiology Center, Department of Obstetrics and Gynecology, Vanderbilt University, Nashville, TN; 10) Center for Diabetes Research, Center for Genomics and Personalized Medicine Research, Wake Forest School of Medicine, Winston-Salem, NC; 11) The Center for Human Genetics Research, Vanderbilt Epidemiology Center, Department of Medicine, Vanderbilt University, Nashville, TN; 12) Division of Public Health Scienes, Fred Hutchinson Cancer Research Center, Seattle, WA; 13) Department of Preventive Medicine, Keck School of Medicine and Norris Comprehensive Cancer Center, University of Southern California, Los Angeles, CA; 14) Division of Prevention and Population Sciences, National Heart, Lung and Blood Institute, National Institutes of Health, Bethesda,MD; 15) Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA; Divisions of Genetics and Endocrinology, Children's Hospital, Boston, Massachusetts; and Departments of Genetics, Harvard Medical School, Boston, MA.

Obesity is a well-established risk factor for several chronic diseases and presents a substantial public health challenge. Thirty six loci influencing body mass index (BMI) have been discovered through genome-wide association studies (GWAS) in individuals of European and Asian descent. Far fewer GWAS have been conducted in individuals of other ancestries, and have so far not identified new BMI loci. Our study addresses two key issues: are loci identified in populations of European and Asian ancestry important in those of African ancestry, and can novel associations be detected by studying individuals of African ancestry? Our study included 39,144 African ancestry adults from 19 studies that each tested ~2.5 million SNPs for association with BMI; individual study results were meta-analyzed (Stage 1). All analyses were adjusted for global ancestry, and imputation was performed using a combination of the HapMap YRI and CEU panels. To further validate associations, we selected the 1,500 most significantly associated SNPs from Stage 1 and tested them in a second stage (Stage 2) of 7,954 men and women of African ancestry from eight additional studies and a third stage (Stage 3) in ~250,000 European descent individuals in the GIANT consortium. We identified seven loci associated with BMI at genome-wide significance ($p<5.0x10-8$); two novel loci at 5q33 (GALNT10, rs7708584, $p=3.2x10-11$) and 6q16 (KLHL32, rs974417, $p=2.0x10-8$) and five signals at previously established loci in or near BDNF, RP11/GNPDA2, FTO, MC4R, and PCSK1/CAST. In Stage 3, after meta-analysis with data from the GIANT consortium, a third locus at 7p15 (rs10261878, $p=4.5x10-8$) was identified. We also confirmed that 32 of 36 of loci previously associated with BMI in European and Asian descent populations showed directional consistency in our Stage 1 data (p-value for binomial test of direction=4.05x10-7), with similar effects on BMI. At eight loci (PTBP2, TMEM18, RBJ, FLJ35779, NUDT3, BDNF, FTO, MC4R) we identified SNPs that are more strongly associated than are the previously reported index SNPs (to at least one order of magnitude), which may eventually aid in fine mapping and identification of underlying functional variants. Taken together, we identified novel BMI-associated variants by studying individuals of African ancestry, demonstrating the utility of studying populations with distinct ancestries and environmental influences to better understand obesity risk factors.

2132F

Genetic Determinants of IL-1 Receptor Antagonist in the Circulation. *ML. Nuotio[1,2], K. Kristiansson[1], M. Perola[1,2,3], V. Salomaa[1].* 1) National Institute for Health and Welfare Helsinki, Finland; 2) Institute for Molecular Medicine Finland; 3) University of Tartu Estonia.

Interleukins play an essential role in human immunity system acting as important mediators of inflammation and tissue damage in various organs. The interleukin-1 superfamily consists of two agonists, IL-1alpha and IL-1beta, two receptors, IL-1RI and IL-1RII, and a specific receptor antagonist, IL-1Ra. IL-1Ra gene (IL1RN) maps into the chromosome 2 and has long been under a careful research, since the level of IL-1Ra concentration in circulation may have clinical significance and possibly associates for example with diabetes and cardiovascular diseases. Systemic administration of recombinant human IL-1Ra has been shown to enable a better control of type 2 diabetes mellitus. Furthermore, IL-1Ra concentrations predict the progression of metabolic syndrome to clinically incident diabetes. Recombinant form of IL-1-receptor antagonist (anakinra) is also a commonly used drug in the treatment of rheumatoid arthritis and gout. In this study we wanted to test genome-wide association with IL-1Ra concentrations in circulation measured in several independent Finnish population- and case-control-cohorts, and meta-analyze the results further in collaboration with European population cohorts. Currently we seek to replicate variants remaining genome-wide significant in various European population cohorts. We also wanted to test the association of top SNPs with leukocyte transcriptome data. We performed the primary GWAS in a Finnish sample consisting in total of 7200 adults. Independent association studies were combined using inverse-variance weighted meta-analysis. As a result of the discovery phase meta-analysis, the best 17 SNPs were associated with IL-1Ra at $p < 10E-12$, when adjusting for age, sex, BMI and waist/hip-ratio. Most of these variants were located in genes belonging to the interleukin cluster in chr 2. While testing the association of the Finnish top SNPs with leukocyte transcriptome data, one polymorphism in IL1F10 was shown to act as cis-eQTL to the gene PAX8 and trans-eQTL to ALDH2 in chromosome 12 and LILRB4 in chromosome 19. Interleukin-1-receptor antagonist has shown a clinical significance and an association with a number of medical conditions. Identifying genetic variants associated with IL-1-Ra level in circulation provides further information about related pathophysiological pathways and enables the evaluation of its potential effects on cardiovascular disease and diabetes.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2133W

Multi-ethnic GWA meta-analysis identifies new endometriosis risk loci. *D.R. Nyholt[1,14], S-K. Low[2,14], C.A. Anderson[3], J.P. Painter[1], S. Uno[2,4], A.P. Morris[5], S. MacGregor[1], S.D. Gordon[1], A.K. Henders[1], N.G. Martin[1], J. Attia[6], E.G. Holliday[7], M. McEvoy[8], R.J. Scott[9], S.H. Kennedy[10], S.A. Treloar[11], S.A. Missmer[12], S. Adachi[13], K. Tanaka[13], Y. Nakamura[2], K.T. Zondervan[5,10,15], H. Zembutsu[2,15], G.W. Montgomery[1,15].* 1) Queensland Institute of Medical Research, Brisbane, QLD 4029, Australia; 2) Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan; 3) Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, CB10 2HH, UK; 4) First Department of Surgery, Sapporo Medical University, School of Medicine, Sapporo, Japan; 5) Genetic and Genomic Epidemiology Unit, Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, OX3 7BN, UK; 6) Centre for Clinical Epidemiology and Biostatistics, School of Medicine and Public Health, The University of Newcastle, Newcastle, NSW 2305, and Hunter Medical Research Institute, John Hunter Hospital, Newcastle, NSW 2305, Australia; 7) Centre for Information Based Medicine and the School of Medicine and Public Health, The University of Newcastle, Newcastle, NSW 2308, and Hunter Medical Research Institute, John Hunter Hospital, NSW 2305, Australia; 8) Centre for Clinical Epidemiology and Biostatistics, School of Medicine and Public Health, The University of Newcastle, Newcastle, NSW 2308, Australia; 9) Centre for Information Based Medicine and the School of Biomedical Science and Pharmacy, The University of Newcastle, Newcastle, NSW 2308, and Hunter Medical Research Institute, and Hunter Area Pathology Service, John Hunter Hospital, Newcastle, NSW 2305,; 10) Nuffield Department of Obstetrics and Gynaecology, University of Oxford, John Radcliffe Hospital, Oxford, OX3 9DU, UK; 11) Centre for Military and Veterans' Health, The University of Queensland, Mayne Medical School, 288 Herston Road, QLD 4006, Australia; 12) Department of Obstetrics, Gynecology and Reproductive Biology, Brigham and Women's Hospital and Harvard Medical School, 75 Francis Street, Boston, MA 02115, USA; 13) Department of Obstetrics and Gynecology, Niigata University Graduate School of Medical and Dental Sciences, 1-757 Asahimachi-dori, Niigata, Japan; 14) These authors contributed equally to this work; 15) These authors jointly directed this work.

Endometriosis (MIM131200) is a common gynecological disease associated with severe pelvic pain, affecting 6–10% of women in their reproductive years and 20–50% of women with infertility. Endometriosis risk is influenced by genetic factors and has an estimated heritability of around 51%. We conducted a multi-ethnic genome-wide association (GWA) meta-analysis of 4,604 endometriosis cases and 9,393 controls of European[1] and Japanese[2] ancestry. In addition to showing that rs12700667 on chromosome 7p15.2 previously found in Europeans replicates in Japanese ($P = 3.6 \times 10^{-3}$), we confirmed association of rs7521902 on 1p36.12 near *WNT4*, established association of a SNP on 2p25.1, and identified a novel locus on 12q22. Excluding endometriosis cases with minimal (rAFS stage I-II) or unknown severity, we identified additional novel loci on 2p14, 6p22.3 and 9p21.3. All seven SNP effects were replicated in an independent cohort and produced $P < 5 \times 10^{-8}$ in a combined analysis of 5,648 endometriosis cases and 13,410 controls. A gene-based GWA analysis using VEGAS[3] revealed two genome-wide significant genes and implicated other genes associated with endometriosis. Finally, we found a significant overlap in polygenic risk for endometriosis between the European and Japanese GWA cohorts ($P = 8.8 \times 10^{-11}$), indicating that many weakly associated SNPs represent true endometriosis risk loci. To our knowledge, this is the first report confirming a shared polygenic risk for a complex trait such as endometriosis across different ethnic populations, signifying risk prediction and future targeted disease therapy may be transferred across these populations.

1. Painter, J.N. et al. Genome-wide association study identifies a locus at 7p15.2 associated with endometriosis. Nat Genet 43, 51-4 (2011)

2. Uno, S. et al. A genome-wide association study identifies genetic variants in the CDKN2BAS locus associated with endometriosis in Japanese. Nat Genet 42, 707-10 (2010)

3. Liu, J.Z. et al. A versatile gene-based test for genome-wide association studies. Am J Hum Genet 87, 139-45 (2010).

2134T

Genome-wide association study reveals new candidate loci for hand osteoarthritis. *A. Näkki[1,2,3], J. Eriksson[4,5,6], E. Widen[1], A. Palotie[1,7], U.M. Kujala[8,9], J. Saarela[1].* 1) Institute for Molecular Medicine Finland FIMM, Helsinki, Finland; 2) Unit of Public Health Genomics, National Institute for Health and Welfare, Helsinki, Finland; 3) Department of Medical Genetics, University of Helsinki, Helsinki, Finland; 4) Department of General Practice and Primary Health Care, University of Helsinki, Helsinki, Finland; 5) Helsinki University Central Hospital, Unit of General Practice, Helsinki, Finland; 6) Folkhalsan Research Centre, Helsinki, Finland; 7) Wellcome Trust Sanger Institute, Hinxton, Cambridge, United Kingdom; 8) ORTON Orthopedic Hospital, Invalid Foundation, Helsinki, Finland; 9) Department of Health Sciences, University of Jyväskylä, Jyväskylä, Finland.

Objective. Osteoarthritis (OA) is a complex disease common in the elderly. Our aim was to study single nucleotide polymorphisms (SNPs) using a genome-wide association (GWA) data in a setting of hand OA cases versus healthy controls. Methods. The study subjects were over 56 years of age and were part of Finnish Helsinki Birth Cohort Study sample genotyped using Illumina HumanHap 610 GWA chip. Hand joints of the study subjects were visually evaluated and individuals with Heberden's nodes in at least one DIP joint were graded as affected (n=524). Individuals graded as healthy (n=970) had visually healthy symptomless finger joints. Association with individual SNPs was monitored using the Plink program. Age, sex and two principal components differed between the cases and controls and were used as covariates. Results. In our hand OA GWA analysis the most significant results were for SNP flanking the CD28 gene in 2q33 (p=4.37x10-6, OR=1.68, 95% CI 1.35-2.10); SNP in the SUCLG2 gene in 3p14 (p=6.15x10-6, OR=0.62, 95% CI 0.50–0.76) and SNP in the SLC22A3 gene flanking lipoprotein genes LPAL2 and LPA in 6q25 (p=1.61x10-5, OR=1.41, 95% CI 1.20–1.64). Conclusion. Variants in regions harbouring CD28, SUCLG2, SLC22A3, LPAL2 and LPA genes showed suggestive evidence for association with hand OA. CD28 has been shown to affect T-cell functions and IL2 production. CD28 and LPA genes have been shown to play role in rheumatoid arthritis. Results did not reach genome-wide significance in the initial screening phase.

2135F

Genome-wide Association Study and Meta-Analysis of Intraocular Pressure in Primary Open-Angle Glaucoma Cases and Controls in the NEIGHBOR, GLAUGEN and AMD-MMAP MI Consortia. *A.B. Ozel[1], S. Moroi[2], D. Reed[2], E. Trager[2], K. Scott[2], J. Gilbert[2], S. Garnai[2], S. Akbari[2], M. Nika[2], C. Schmidt[2], B. Yaspan[3], K. Branham[2], W. Chen[4], J. Heckenlively[2], A. Swaroop[5], G. Abecasis[6], A. Ashley-Koch[7], M. Ulmer[7], M. Hauser[7], J. Haines[3], L. Pasquale[8], J. Wiggs[8], J. Richards[2], J.Z. Li[1], NEIGHBOR Consortium.* 1) Department of Human Genetics, University of Michigan, Ann Arbor, MI; 2) Department of Ophthalmology and Visual Sciences, University of Michigan, Ann Arbor, MI; 3) Center for Human Genetics Research, Vanderbilt University School of Medicine, Nashville, TN; 4) Children's Hospital of Pittsburgh, University of Pittsburgh Medical School, Pittsburgh, PA; 5) National Eye Institute, National Institutes of Health, Bethesda, MD; 6) Department of Biostatistics, University of Michigan, Ann Arbor, MI; 7) Center for Human Genetics, Duke University School of Medicine, Durham, NC; 8) Department of Ophthalmology, Harvard Medical School, Massachusetts Eye and Ear Infirmary, Boston, MA.

Intraocular pressure(IOP) is determined by intraocular fluid production, fluid drainage and venous pressure of the eye. The IOP range of 10–20mmHg is regulated by physiological pathways and influenced by environmental factors that are not fully understood. In many forms of glaucoma, including the most common form, primary open angle glaucoma (POAG), there is relative obstruction to intraocular fluid drainage that causes elevated IOP. In susceptible individuals, this leads to glaucomatous nerve damage with vision loss and even blindness. We performed genome-wide association studies of IOP from healthy individuals (controls, n=3,386) and glaucoma patients (cases, n=805) of European ancestry from three studies: (1) NEIGHBOR, with 350 untreated cases and 2,189 controls genotyped on Illumina Human660W_Quad_v1 array (~510K SNPs); (2) GLAUGEN, with 325 untreated cases and 325 controls genotyped on Illumina Human660W_Quad_v1 array (~484K SNPs after QC); and (3) AMD-MMAP (the MI subset), with 130 untreated cases and 872 controls genotyped on Illumina HumanCNV370v1_C platform and ~2.5 million imputed SNPs based on the HapMap2 CEU data. IOP was defined as the maximum value from one of the two eyes. The effect of variables including age, gender, central cornea thickness (CCT), and vertical cup-to-disc ratio (vCDR), on IOP was assessed by linear regression. Significant variables were included as covariates in association tests. We performed meta-analyses using cases and controls or only controls from the three studies, and a two-study meta-analysis using only cases from NEIGHBOR and GLAUGEN. No signal of genome-wide significance (p<5E-8) was revealed. A candidate gene analysis (~100K SNPs) was undertaken to specifically investigate previously reported IOP or glaucoma genes or those expressed in ciliary body or trabecular meshwork. This analysis identified a SNP (rs10037649,p=5.6E-07) at 5q23.2, which is within GLC1M, a novel locus for autosomal dominant juvenile-onset open angle glaucoma (JOAG) previously identified from family-based linkage studies.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2136W

Haplotype Association Mapping In 33 Inbred Mouse Strains Identifies Genetic Regions Contributing To Chronic Hypoxia-Induced Pulmonary Hypertension. *M.W. Pauciulo[1], D. Koller[2], P. Hale[1], T.D. Le Cras[3,5], P. Pastura[3], B. Aronow[4,5], C. Tolentino[1,3], D. Li[2], T. Foroud[2], W.C. Nichols[1,5].* 1) Division of Human Genetics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH; 2) Department of Medical and Molecular Genetics, Indiana University Medical Center, Indianapolis, IN; 3) Division of Pulmonary Biology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH; 4) Division of Biomedical Informatics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH; 5) Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH.

While mutations primarily in BMPR2 have been identified in patients with both heritable and idiopathic forms of pulmonary arterial hypertension (PAH), data suggest that additional genetic factors contribute to the disease. Short of the necessary large patient cohorts to power genomewide studies that could lead to the identification of additional genetic factors, researchers are left with little alternatives for genetic studies of PAH. To circumvent the paucity of PAH patient cohorts available, we are performing genomewide studies in readily available strains of inbred mice. Mice from 33 different strains were housed under hypoxic conditions (10% O_2) for four weeks with equivalent numbers of age and sex-matched mice maintained under normoxic conditions for all strains. After the hypoxic exposure, all mice underwent right heart catheterization to measure right ventricular systolic pressure (RVSP). Hearts were dissected at necropsy to assess the degree of right ventricular hypertrophy (RVH) as determined by the ratio of the weight of the right ventricle (RV) to the weight of the left ventricle and septum (LV+S). Hematocrit was measured via retro-orbital puncture. Due to the small size of some mice, RVSP was only determined for 31 strains. Haplotype association mapping (HAM) was performed using efficient mixed-model analysis (EMMA) and a set of 132,000 SNPs to identify regions in the genome associated with any of the quantitative traits measured. The HAM included separate analyses for hypoxic and normoxic mice as well as a "response to hypoxia" analysis taking into account the difference in RVSP (or RV/LV+S or hematocrit) after hypoxia as compared to normoxic controls. Associations reaching genomewide significance ($p=1x10^{-5}$) were identified for both RVSP and RV/LV+S. The most significant association for RVSP was identified on chromosome 8 ($p=2.34x10^{-11}$) while that for RV/LV+S was on chromosome 6 ($p=9.37x10^{-11}$). Additional significant associations were identified on other chromosomes. These data suggest that genetic factors contributing to RVH and increases in RVSP can be identified using genomewide methods in inbred mice. Efforts are underway to identify the specific loci involved with the goal of translating these findings to patients with PAH. Identification of additional genetic factors contributing to PAH can ideally reveal novel pathways involved in disease pathogenesis that will lead to the development of novel therapeutics for patient treatment.

2137T

Large-scale genome-wide association meta-analysis of fasting glycemic traits using imputation from 2188-haplotype 1000 Genomes reference panel within ENGAGE consortium. *I. Prokopenko[1,2], R. Mägi[3], M. Horikoshi[1,2], I. Surakka[4,5], S. Wiltshire[1,2], A.-P. Sarin[4,5], A. Mahajan[1], L. Marullo[6,1], T. Ferreira[1], S. Hägg[7], J. S. Ried[8], G. Thorleifsson[9], T. Esko[3,10], S. M. Willems[11], A. P. Morris[1], M. I. McCarthy[1,2,12], S. Ripatti[4,5,13], European Network for Genetic and Genomic Epidemiology Consortium.* 1) Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford OX3 7BN, United Kingdom; 2) Oxford Centre for Diabetes, Endocrinology and Metabolism,, University of Oxford, Oxford, OX3 7LJ, United Kingdom; 3) Estonian Genome Center, University of Tartu, Tartu, Estonia; 4) Institute for Molecular Medicine Finland FIMM, University of Helsinki, Helsinki, Finland; 5) National Institute for Health and Welfare, Helsinki, Finland; 6) Department of Evolutionary Biology, Genetic Section, University of Ferrara, Ferrara, 44121, Italy; 7) Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden; 8) Institute of Genetic Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany; 9) deCODE Genetics, Reykjavik, Iceland; 10) The Institute of Molecular and Cellular Biology of the University of Tartu, Tartu, Estonia; 11) Department of Genetic Epidemiology, Erasmus MC, Rotterdam, The Netherlands; 12) Oxford NIHR Biomedical Research Centre, Churchill Hospital, Oxford OX3 7LJ, UK; 13) Wellcome Trust Sanger Institute, Hinxton, UK.

To date ~50 loci have been described for fasting glycemic traits in non-diabetic individuals; however these explain only ~5% of phenotypic variance and are represented by common variants with minor allele frequency (MAF)>5%. We aimed to identify potential causal variants at known fasting glucose (FG) and fasting insulin (FI) loci, to uncover novel associated variants and to identify low allele frequency sites influencing glycemic trait variability. To that end, we performed imputation from the largest to date 1000 Genomes reference panel (2188 haplotypes, June 2011 release) within 10 European cohorts with genome-wide association (GWA) data from ENGAGE consortium. Following FG and FI association analysis, we performed a fixed-effects inverse variance meta-analysis for 38,326 and 16,831 individuals, respectively. 14.9M SNPs were present in at least 5 studies and passed quality control. We observed genome-wide significant (p-value$<5x10^{-8}$) associations at 8 known FG loci, but no such associations were seen for FI. Our strategy highlighted several novel characteristics in comparison to previous imputation efforts based on the HapMap2 reference panel: (a) we increased the number of SNPs tested by ~6 fold; (b) for FG we identified a novel signal at common variant within *RMST*gene (rs17331697, MAF=0.10, β=0.05 (se=0.01), $p=2.0x10^{-10}$); (c) a known common missense variant within *GCKR* (rs1260326 [P446L]) on imputation showed more significant association for both, FG and FI, (MAF=0.36, FG $p=7.6x10^{-10}$; FI $p=1.1x10^{-7}$) than its proxy variant previously described as a lead SNP for fasting glycemic traits (rs780094, MAF=0.35 FG $p=5.6x10^{-9}$; FI $p=1.2x10^{-6}$), which could be exampled as a positive control given the strong functional data pointing to P446L as causal variant; (d) in the vicinity of 5 known FG loci we identified novel associations at low MAF variants with larger effect sizes than previously described GWAs signals (e.g. within *G6PC2*, a missense variant rs138726309 [MAF=0.011, β=0.18(se=0.022), $p=9.4x10^{-17}$] 114bp from the known signal at rs560887 [MAF=0.31, $p=1.1x10^{-70}$]), for which conditional analyses will help to verify, whether low frequency variants represent independent associations. Imputation from successively larger and more complete reference panels generated by genome resequencing in existing GWAs underscores the prospective for identification of novel associated variants and for fine-mapping within already known loci.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2138F

BRC Allergene study: A TwinsUK GWAS of a metabolite associated with allergic response to nickel. *L. Quaye[1], L. Bevan[1], T. Tsakok[2], C. Menni[1], A. Moayyeri[1], I. Erte[1], T.D. Spector[1], C.J. Hammond[1], F.O. Nestle[2].* 1) Twin Research & Genetic Epidemiology, King's College London, London, United Kingdom; 2) St Johns Institute of Dermatology, King's College London, London, United Kingdom.

Allergic conditions affect 20–30% of the population and are responsible for considerable morbidity. However, little is known about the genetic aetiology of immediate and delayed type immunity underpinning allergic reactions. Allergic response to nickel is common, affecting approximately 15% of women. The identification of associated genes would result in greater understanding of disease mechanisms and potential therapeutic targets. The Biomedical Research Centre allergy project is the first twin study to investigate genetic determinants of type I and IV allergic responses. It encompasses a combination of genetic, expression, methylation and metabolomics data on up to 5,500 individuals from the TwinsUK registry and allergy data using questionnaires, skin and blood. 784 females, comprising 120 monozygotic (MZ) and 270 dizygotic (DZ) twin pairs were tested for 48 hour response to nickel patch testing. A positive reaction was observed in at least 1 individual of 34 MZ and 102 DZ twin pairs. The case-wise concordance was ~50% for MZs and 30% for DZs, which suggests an underlying genetic component to nickel allergy. Associations between nickel response and the abundance of 401 metabolites were evaluated in order to identify a potential biomarker of the allergic response. The strongest association was with isovalerate (P= 1.97E-03), a component of fatty acid metabolism. A genome-wide association study of the abundance of isovalerate was performed to ascertain correlations with common genetic variants. The strongest association was with a SNP of ANO6 (12q12), rs11183014 (P=1.8E-06). We also found suggestive associations with 4 additional SNPs from ANO6 and a variant of CAMK1D (P<9. 5E-06). ANO6 encodes a multi-pass transmembrane protein, which is important for calcium-dependent exposure of phospholipids on the cell surface. CAMK1D encodes a calcium/calmodulin-dependent protein kinase which been proposed to regulate granulocyte function through chemokine signal transduction. The roles of ANO6 and CAMK1D in delayed type hypersensitivity have not previously been described, however loci from ANO6 were identified in Ankylosing Spondylitis and variants of CAMK1D have been implicated in type II diabetes. The results from this study will be replicated with additional samples. To conclude, the combined use of metabolomic and genome-wide genotyping data offer great promise in the discovery of novel genetic variants for nickel allergy.

2139W

Experience with Illumina Infinium FFPE DNA Restoration Solution. *J. Romm, S. Pottinger, L. Aker, C. Hilton, C. Ongaco, K. Doheny.* CIDR, Baltimore, MD.

The Center for Inherited Disease Research (CIDR) provides high quality next-generation sequencing (NGS), genotyping and statistical genetics consultation to investigators working to discover genes that contribute to disease. Over the last year, CIDR has gained experience with the Illumina Infinium FFPE DNA Restoration Solution. This solution is intended for formalin-fixed, paraffin embedded (FFPE) samples; however, our current experiences with the kit have been to restore severely compromised, highly-degraded DNA samples. CIDR used the Restoration Solution for severely degraded genomic DNA samples (Project 1–493 samples) as well as for DNA extracted from serum (Project 2–4,126 samples). Without restoration, the predicted overall failure rate for samples and SNPs was 50% and 17% respectively for Project 1 based on the first 3 plates run in the lab and the predicted sample and SNP failure rate was 10% and 15% respectively for Project 2 based on pilot study projections. Project 1 was comprised of both restored and unrestored samples and run using the IlluminaHumanOmniExpress-FFPE BeadChip. Due to the CIDR workflow, all samples were restored in Project 2, including HapMap controls, and run on the Illumina HumanOmni1Quad BeadChip with a reagent substitute from Illumina. Samples for each project were reclustered using GenomeStudio 2011.1, Genotyping module 1.9.4 and Gentrain version 1. Overall quality statistics were very high for the restored samples. Project 1 has an overall missing rate of 1% for the restored samples and a duplicate sample reproducibility rate of 99.97% for the restored samples. Average call rate for the restored samples was 98.8%. Project 2 had an overall missing rate of 0.97%, compared with 0.23% for HapMap samples included with the project. The overall blind duplicate reproducibility rate was 99.91%. The mean call rate of the restored samples was 99.03%. One noticeable observation in the data was the elevated LogRRatio for the restored samples. Average LogRRatio for the restored samples was 0.36 for Project 1 and 0.26 for Project 2, compared with an average LogRRatio of 0.16 for projects with no restored samples.

2140T

The Kaiser Permanente/UCSF Genetic Epidemiology Research Study on Adult Health and Aging: A genome-wide association analysis of type 2 diabetes. *L.C. Sakoda[1], L. Shen[1], S. Sciortino[1], D.K. Ranatunga[1], R.W. Grant[1], T.J. Hoffmann[2], M.N. Kvale[2], S. Hesselson[2], Y. Banda[2], P. Kwok[2], N. Risch[1,2], C.A. Schaefer[1], A. Ferrara[1].* 1) Division of Research, Kaiser Permanente, Oakland, CA; 2) Institute for Human Genetics, UCSF, San Francisco, CA.

Linking clinical data from electronic medical records (EMRs) to comprehensive genetic data from biological samples offers a highly efficient approach for examining genetic associations with a number of disease-related phenotypes. Conducting such analyses using a large cohort of patients within an integrated health care system has transformative potential to advance clinical care. To assess the utility of this approach, we performed a genome-wide association (GWA) analysis of type 2 diabetes (T2D) in the Kaiser Permanente (KP)/UCSF Genetic Epidemiology Research Study on Adult Health and Aging cohort. Of the ~100,000 participants whose DNA has been genotyped with custom Affymetrix Axiom genome-wide arrays, those with and without diabetes were identified through linkage to the KP Diabetes Registry, which uses algorithms from multiple sources including diagnosis, prescription, and laboratory data. Persons with T2D were distinguished based on self-report, age at diagnosis of ≥30 years, or age at diagnosis of <30 years and no insulin prescriptions. Persons with other or undetermined diabetes type, based on medical diagnosis, prescription records, or self-reported history, were excluded. GWA analyses were limited to non-Hispanic whites: 7,068 with and 62,582 without T2D. Based on additive models adjusted for age (at DNA collection), sex, Ashkenazi/European ancestry, and body mass index (calculated from EMR data), we detected SNP associations in multiple, previously reported T2D susceptibility loci, including *TCF7L2* ($p = 1.4 \times 10^{-83}$), *PPARG* ($p = 1.3 \times 10^{-16}$), *IGF2BP2* ($p = 5.5 \times 10^{-14}$), an intergenic region at 9p21.3 near *CDKN2B-AS1* ($p = 1.3 \times 10^{-13}$), *SLC30A8* ($p = 7.0 \times 10^{-12}$), *JAZF1* ($p = 9.7 \times 10^{-11}$), *SYN2* ($p = 1.9 \times 10^{-8}$), an intergenic region at 10q23.33 near *HHEX* ($p = 2.5 \times 10^{-8}$), and an intergenic region at 2q36 near *IRS1* ($p = 3.6 \times 10^{-8}$). We also found suggestive evidence of novel associations in regions of chromosomes 1 and 5 ($p < 1 \times 10^{-7}$). GWA analyses of T2D in ethnic minorities are ongoing. Using T2D as an example, we demonstrate in this cohort that patient data routinely captured in EMRs are a valid and valuable resource for characterizing phenotypes to discover and replicate gene-disease associations.

2141F

A Genome-wide association study of *Helicobacter pylori* infection susceptibility. *C. Schurmann[1], J. Mayerle[2], L. Stolk[3,4], M.J. Peters[3,4], G. Homuth[1], A.G. Uitterlinden[3,4,5], M.M. Lerch[2], E. Kuipers[3,6].* 1) Department of Functional Genomics, Interfaculty Institute for Genetics and Functional Genomics, University of Greifswald, Greifswald, Germany; 2) Department of Medicine A, University Medicine Greifswald, Greifswald, Germany; 3) Department of Internal Medicine, Erasmus University Medical Center, Rotterdam, The Netherlands; 4) Netherlands Consortium for Healthy Aging, Rotterdam, Leiden, The Netherlands; 5) Department of Epidemiology Erasmus University Medical Center, Rotterdam, The Netherlands; 6) Department of Gastroenterology and Hepatology, Erasmus University Medical Center, Rotterdam, The Netherlands.

Helicobacter pylori (*H. pylori*) is a spiral-shaped, Gram-negative, microaerophilic bacterium that inhabits the stomach. It is the major cause of gastritis, gastroduodenal ulcer disease and the only bacterial pathogen that causes cancer. The infection prevalence rises up to 90% in some developing countries but 10% of subjects are never colonized with *H. pylori* regardless of exposure. Genetic factors are thought to confer *H. pylori* susceptibility. Therefore, the aim of our study was to identify genetic risk factors for *H. pylori* infection susceptibility.

We performed genome-wide association studies (GWAS) in Europeans from three population based cohorts (RS-I, RS-II and SHIP; N=10,485) to identify genetic variants associated with*H. pylori* colonization as defined by the anti-*H. pylori* immunglobulin G (IgG) titer. Furthermore, we analyzed *cis*-eQTL effects of the top GWAS SNPs based on whole-blood gene expression data of two population based cohorts (RS-III and SHIP-TREND; N= 1,753) and investigated the correlation of genome-wide mRNA expression profiles and *H. pylori* colonization (N=961).

Our GWAS meta-analysis identified two loci to be associated with the defined *H. pylori* infection susceptibility phenotype. At the first locus, we identified three genes coding for proteins involved in the innate immune system, of which only one was differentially expressed with respect to the GWAS top SNP genotype. Furthermore, its expression was upregulated in individuals with prominent *H. pylori* colonization measured in stool samples. The relative risk for *H. pylori* infection was reduced by 41% in homozygotes of the protective minor allele. The second locus also contains genes coding for proteins involved in the immune system. For both loci, we identified significant *cis*-eQTL effects.

In summary, we identified two loci to be associated with *H. pylori* susceptibility and propose that variation affecting the corresponding genes might explain why some individuals have a lower risk for *H. pylori* infection.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2142W

GWAS needs a good phenotype: clustering tooth surfaces into biologically-informative dental caries outcomes. *J.R. Shaffer[1], E. Feingold[1,2], X. Wang[3], M. Lee[3], K. T.Cuenco[3], D.E. Weeks[1,2], R.J. Weyant[4], R. Crout[5], D.W. McNiel[6], M.L. Marazita[1,3,7].* 1) Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA; 2) Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA; 3) Center for Craniofacial & Dental Genetics, Department of Oral Biology, School of Dental Medicine, Pittsburgh, PA; 4) Department of Dental Public Health and Information Management, School of Dental Medicine, University of Pittsburgh, Pittsburgh, PA; 5) Department of Periodontics, School of Dentistry, West Virginia University, Morgantown, WV; 6) Dental Practice and Rural Health, West Virginia University, Morgantown, WV; 7) Clinical and Translational Science, School of Medicine, University of Pittsburgh, Pittsburgh, PA.

Dental caries (tooth decay) affects most adults worldwide, and the importance of susceptibility genes in risk of dental caries has been clearly established. While many candidate caries genes have been proposed to date, few of them have been rigorously validated through observational and experimental studies. Moreover, most genetic epidemiological studies have analyzed global caries phenotypes that ignore the possibility that genes may exert differential effects across tooth surfaces of the dentition. However, risk factors for dental caries do not necessarily exert their effects uniformly across all tooth surfaces. Instead, the actions of some risk factors, including genetic factors, may be limited to a subset of teeth/surfaces. Therefore, caries phenotypes that capture the patterns of tooth decay are needed. Toward this end, we used hierarchical clustering on tooth surface-level caries data for 920 Appalachian adults (ages 18–75) to group surfaces based on co-occurrence of caries. Cluster analysis yielded evidence of five distinct groups of tooth surfaces that differ with respect to caries: (1) pit and fissure molar surfaces, (2) mandibular anterior surfaces, (3) posterior non-pit and fissure surfaces, (4) maxillary anterior surfaces, and (5) mid-dentition surfaces. We created new caries outcomes defined as the number of carious tooth surfaces within each cluster. We showed that some cluster-based caries outcomes are heritable (i.e., under genetic regulation; p-value<0.05), and performed genome-wide association studies (GWAS) of each cluster-based phenotype in order to nominate new candidate caries genes. We identified a significant genetic association between dental caries of the anterior mandibular teeth and LYZL2 (p-value=9e-9), a bacteriolytic agent thought to be involved in host defense. We also identified a significant genetic association between caries of the mid-dentition tooth surfaces and AJAP1 (p-value=2e-8), a gene possibly involved in tooth development. Numerous suggestive genetic associations in plausible caries genes were also observed. These results suggest that the permanent dentition can be subdivided into groups of tooth surfaces that are useful for understanding the genetic factors influencing cariogenesis. Funding support was provided by NIH grants U01-DE018903, R01-DE014899, and R03-DE021425.

2143T

X Chromosome SNPs Not Strongly Associated with Juvenile Idiopathic Arthritis. *M. Sudman[1], M. Marion[2], T. Howard[2], J. Haas[3], S. Prahalad[4], J. Bohnsack[5], C. Wise[6], M. Punaro[7], C. Rose[8], M. Ryan[1], M. Wagner[9], C. Langefeld[2], D. Glass[1], S. Thompson[1].* 1) Division of Rheumatology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH; 2) Department of Biostatistical Sciences, Wake Forest University School of Medicine, Winston-Salem, NC; 3) German Centre for Rheumatology in Children and Young People, Garmisch-Partenkirchen, Germany; 4) Division of Rheumatology, Emory Children's Center, Atlanta, GA; 5) Department of Pediatrics, University of Utah Health Sciences Center, Salt Lake City, UT; 6) Sarah M. and Charles E. Seay Research Center, Texas Scottish Rite Hospital for Children, Dallas, TX; 7) Arthritis Clinic, Texas Scottish Rite Hospital for Children, Dallas, TX; 8) Division of Rheumatology, duPont Children's Hospital, Wilmington, DE; 9) Division of Biomedical Informatics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH.

Juvenile Idiopathic Arthritis (JIA) is the most common chronic rheumatic disease in the western world. JIA is a complex genetic disease with an estimated sibling risk ratio of ~15, and is a clinically heterogeneous group of diseases that can be classified into seven subtypes according to the International League of Associations for Rheumatology (ILAR). The two most common subtypes of JIA, oligoarthritis (both persistent and extended) and rheumatoid factor (RF) negative polyarthritis, exhibit a strong female bias with a girl to boy ratio of ~4:1. This gender bias is also common to other rheumatic autoimmune diseases. Given this gender disparity, it is important to investigate whether the X chromosome might hold JIA-predisposing loci. We have completed a meta-analysis of three genome-wide association studies (GWAS) comprising 1902 JIA cases limited to oligoarthritis and RF negative polyarthritis subtypes (1464 female, 438 male) and 12,450 controls (6408 female, 6042 male), all of European descent and genotyped on the Affymetrix Genome-Wide Human SNP Array 6.0. GWAS-1 and GWAS-2 were USA-based cohorts (813 cases/3057 controls, 609 cases/7601 controls, respectively) whereas GWAS-3 was a German-based cohort (480 cases/1792 controls). A total of 21,342 SNPs on the X chromosome passed standard GWAS QC (including principal component analysis), suggesting a Bonferroni adjustment for statistical significance of 2.3x10-6. Analysis was stratified by gender and cohort, and a meta-analysis was computed (weighting by sample size). No SNPs approached chromosome-wide statistical significance. The strongest and most consistent association was on Xq28 for rs6627174 within PASD1 (P=0.00048), a gene whose protein is thought to function as a transcription factor, is expressed in a range of hematopoietic malignancies and stimulates autologous T-cell responses. Although rs6627174 is located within the systemic lupus erythematous (SLE) SLEB15 genetic locus (X:147,100,000–155,270,560), it remains over 2MB from the genes IRAK1 and MECP2, implicated in this region for SLE. A modest association was also observed on Xp21 for rs12008257 (P=0.00065) and rs2404506 (P=0.00097) within the gene DMD. In addition, no SNP had P<0.05 in all three cohorts. Thus, even in these large Affymetrix SNP 6.0 GWAS, we find little evidence that X chromosome SNP variants are associated with JIA, a female-biased disease.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2144F

In quest of ageing and longevity associated genetic markers. *R. Tamm[1, 2], M. Alver[2], K. Fischer[1], R. Magi[1], A. Metspalu[1, 2].* 1) Estonian Genome Center, Tartu, Estonia; 2) University of Tartu, Institute of Molecular and Cell Biology, Tartu, Estonia.

Backround Due to longer life expectancy and decline in fertility rates, the proportion of people over 65 years of age is growing faster than any other age group in developed countries. However, healthy life expectancy is not keeping the pace. Age associated chronic diseases account for an overwhelmingly large fraction of mortality and morbidity in the elderly. This has imposed a socio-economic burden on today's societal challenges. Regardless, there are individuals among the elderly endowed with a peculiar capacity to retain basic functions with age, while escaping major age-related diseases and disabilities. The large variation in healthy lifespan together with the importance in providing the well-being of the ageing population has prompted research in understanding the differences between healthy and diseased ageing mechanisms. As these complex phenotypes are influenced by a variety of genes and molecular pathways, environment and stochastic factors, a wider approach is needed to determine the underlying cause. Whole genome SNP analysis using Illumina arrays and a data-driven model in identifying genetic variants is more promising than scrutinizing candidate genes alone.

Methods The aim of the study was to examine allele frequency dependent discrepancies by conducting a genome-wide association study with with >8.9 million HumanOmniExpress genotyped and 1000 Genome Project Imputed and Quality Controlled SNP markers on 4919 Estonians from different age groups (18–39, 40–59, 60–79, 80+). The allele frequencies in all 4 age groups were compared using the Cochran-Armitage test for trend. Genome-wide significance was concluded for p-values less than 5×10^{-8}.

Results and conclusions Results suggest association between allele frequency differences in several loci, including 2p12, 3p24.1, 12q13.2, 14q32.12, of different age groups. These findings indicate the loci, which include the alleles that are increased and/or decreased in the oldest and could putatively represent new ageing related molecular markers. Our results together with other GWAS data would clarify the mechanisms that contribute to healthy and diseased ageing. New insights can provide new targets to prevent, diagnose and treat the disabilities and diseases that occur with increasing age.

2145W

Genome-Wide Association Scan of Osteoporotic Fracture: A Meta-Analysis of 10,012 African-American Women. *K.C. Taylor[1], D. Edwards[2], T. Edwards[2], D.S. Evans[3], G. Li[4], K.E. North[5], N. Franceschini[5], R. Jackson[6], M. Donneyong[1], A.Z. LaCroix[7], J.A. Robbins[8], B. Lewis[9], M.L. Stefanick[10], Y. Liu[11], M. Garcia[12], T.B. Harris[12], J.A. Cauley[13].* 1) Sch Public Health & Information Sci, Univ Louisville, Louisville, KY; 2) Vanderbilt University, Nashville, TN; 3) California Pacific Medical Center Research Institute, San Francisco, CA; 4) University of Washington, Seattle, WA; 5) University of North Carolina at Chapel Hill, Chapel Hill, NC; 6) The Ohio State University, Columbus, OH; 7) Fred Hutchinson Cancer Research Center, Seattle, WA; 8) University of California at Davis Medical Center, Sacramento, CA; 9) University of Alabama, Tuscaloosa, AL; 10) Stanford Prevention Research Center, Stanford University, Stanford, CA; 11) Wake Forest University, Winston-Salem, NC; 12) National Institute on Aging, Bethesda, MD; 13) University of Pittsburgh, Pittsburgh, PA.

Osteoporosis is a major public health problem associated with excess disability and mortality. For postmenopausal women, the lifetime risk of fracture exceeds 40% and it has been estimated that by 2025, 21% of all fractures will occur in non-white women. Osteoporosis and risk for fracture have been shown to be heritable, with 25–85% of the variation in osteoporosis-related traits being attributable to genetic factors. Recently, nine loci have been associated with fracture-related phenotypes in Caucasian populations. The purpose of this study is to leverage GWAS data from several cohorts to identify genetic determinants of fracture among African American (AA) women. Methods: Data on incident fractures, except fingers, toes, face, skull or sternum, were analyzed for unrelated AA female participants in the Women's Health Initiative (N=8155; 313 cases), Cardiovascular Health Study (N= 504; 50 cases), BioVU (N=704; 43 cases), and Health ABC (N= 651; 88 cases). The Affymetrix 6.0 (WHI, CHS, and BioVU) and Illumina 1M (Health ABC) GWAS panels were used for genotyping, and a 1:1 ratio of YRI:CEU HapMap haplotypes was used as an imputation reference panel. We used Cox proportional hazard models to evaluate the effect of ~2.5 million SNPs on incident fracture rate, adjusting for ancestry, age, and region where applicable. Fractures were all adjudicated except in CHS, which used self-reported fractures. SNPs were filtered on minor allele count (N>10) for all studies except BioVU, which filtered data based on minor allele frequency (>15%). We conducted a fixed-effects, inverse variance-weighted meta-analysis. Genome-wide significance was set at P<5x10-8. Results: One SNP, rs12775980 in SVIL on chromosome 10p11.2, reached genome-wide significance (P=3.5x10-8). Although the minor allele frequency is small (0.03), the direction of effect and effect size were consistent across all studies and this SNP had a P-value of 1.2x10-5 in the largest study, WHI. This locus was not reported in any previous osteoporosis-related GWA studies. Conclusions: This large fracture GWAS meta-analysis has identified a novel locus involving supervillin, a platelet-associated factor that has been previously reported to play a role in thrombus formation. As injection of platelets at a fracture site has been reported to aid in fracture healing, this SNP might predict a new pathway involved in fracture. Replication is needed to determine whether the novel finding is robust.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2146T

Genome-wide association analysis of percent body fat in adults from the Fels Longitudinal Study. *B. Towne[1], J. Blangero[2], A.C. Choh[1], J.E. Curran[2], C. Bellis[2], T.D. Dyer[2], E.W. Demerath[3], M. Lee[1], R.M. Siervogel[1], S.A. Czerwinski[1].* 1) Wright State University School of Medicine, Dayton, OH; 2) Texas Biomedical Research Institute, San Antonio, TX; 3) University of Minnesota, Minneapolis, MN.

Elucidating the genetic underpinnings of overweight and obesity continues to be a challenge. Part of the problem has been an over-reliance on crude measures such as weight or BMI. Using more direct and delineated phenotypic measures, however, has provided new insights into the genetics of body composition. We present here results from genome-wide association analysis of percent body fat (PBF) in 916 healthy adults (430 males; 486 females) aged 18 to 96 years (mean 48.9 years) in the Fels Longitudinal Study. PBF data were obtained from total body dual energy x-ray absorptiometry (DXA) scans taken using a Hologic instrument. Mean PBF was 23.1% in males and 34.9% in females. Subjects were all SNP genotyped using the Illumina Human 610-Quad BeadChip. To facilitate comparison of findings with those from other studies, we used these SNP data to impute HapMap2 data set genotypes. Association analyses were conducted on the imputed HapMap2 SNP data using measured genotype analysis implemented in SOLAR (Almasy and Blangero, 1998) accounting for residual non-independence among relatives. SNP rs7170121, on 15q25.3 at 86,794,649bp in the ATP/GTP binding protein-like 1 gene (AGBL1) is significantly associated with PBF (χ^2 = 29.21; p = $6.5x10^{-8}$). The minor (effect) allele (A) has a frequency of 0.0482, is associated with a 0.651 SDU decrease in PBF, and accounts for 3.6% of the phenotypic variation in PBF. Two other SNPs in the same region of AGBL1, rs17595649 and rs8033536, provide corroborative suggestive association with PBF, with similar effect sizes and proportion of phenotypic variance explained. The physiologic functions of AGBL1 are little known, and there have been few linkage or association studies implicating AGBL1. Of particular interest, however, is that Lewis et al. (2005) reported linkage of PBF to 15q25.3 (nearest marker D15S655), ~300kb from the region of our association signals for PBF. Results of that study, and the results presented here, suggest that AGBL1, or an as yet unknown nearby gene, may contribute to variation observed in percent body fat in healthy adults. Thus, further typing of markers in this region and testing their association with percent body fat appear warranted. Supported by NIH grants R01HD012252, R01DK064391, R01DK064870, R01AR052147, and R37MH59490.

2147F

Genome-wide association meta-analyses in over 210,000 individuals identify 20 sexually dimorphic genetic variants for body fat distribution. *T.W. Winkler[1], D.C. Croteau-Chonka[2], T. Ferreira[3], K. Fischer[4], A.E. Locke[5], R. Mägi[3,4], D. Shungin[6,7,8], T. Workalemahu[9], J. Wu[10], F. Day[11], A.U. Jackson[5], A. Justice[12], R. Strawbridge[13], H. Völzke[14], L. Qi[9], M.C. Zillikens[15], C.S. Fox[16], E.K. Speliotes[17,18], I. Barroso[19,20], E. Ingelsson[21], J.N. Hirschhorn[22], M.I. McCarthy[23], P.W. Franks[6,8,9], A.P. Morris[3], L.A. Cupples[10,24], K.E. North[12], K.L. Mohlke[2], R.J.F. Loos[11,25], I.M. Heid[1], C.M. Lindgren[3], GIANT Consortium.* 1) Public Health and Gender Studies, Institute of Epidemiology and Preventive Medicine, Regensburg University Medical Center, Regensburg, Germany; 2) Department of Genetics, University of North Carolina, Chapel Hill, NC; 3) Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK; 4) Estonian Genome Center, University of Tartu, Tartu, Estonia; 5) Department of Biostatistics, University of Michigan, Ann Arbor, MI; 6) Department of Clinical Sciences, Skåne University Hospital, Lund University, Malmö, Sweden; 7) Department of Odontology, Umeå University, Umeå, Sweden; 8) Genetic and Molecular Epidemiology Group, Department of Public Health and Clinical Medicine, Section for Medicine, Umeå University, Umeå, Sweden; 9) Department of Nutrition, Harvard School of Public Health, Boston, MA; 10) Department of Biostatistics, School of Public Health, Boston University, Boston, MA; 11) MRC Epidemiology Unit, Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge, UK; 12) Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, NC; 13) Cardiovascular Genetics and Genomics Group, Karolinska Institute, Stockholm, Sweden; 14) Institute for Community Medicine, Ernst-Moritz-Arndt-University Greifswald, Greifswald, Germany; 15) Department of Internal Medicine, Erasmus MC Rotterdam, the Netherlands; 16) National Heart, Lung, and Blood Institute, Framingham, MA; 17) Broad Institute, Cambridge, MA; 18) Department of Internal Medicine, Division of Gastroenterology, and Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI; 19) University of Cambridge Metabolic Research Labs, Institute of Metabolic Science Addenbrooke's Hospital, Cambridge, UK; 20) Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, UK; 21) Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden; 22) Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA; 23) University of Oxford, Oxford, UK; 24) Framingham Heart Study, Framingham, MA; 25) Charles R. Bronfman Institute of Personalized Medicine, Child Health and Development Institute, Department of Preventive medicine, Mount Sinai School of Medicine, New York, NY 10029, USA.

It is well-known that body fat distribution differs between men and women, a circumstance that may be due to innate, genetic differences between sexes. Previously, we performed a large-scale meta-analysis of GWAS of waist-to-hip ratio adjusted for BMI (WHR), a measure of body fat distribution independent of overall adiposity and found that of the 14 loci established in men and women combined, seven showed a significant sex-difference. In a subsequent genome-wide analysis that was specifically tailored to detect sex-differential genetic effects for WHR, we identified two additional loci with significant sex-difference. Despite these findings, the genetic basis affecting the sexual dimorphism of WHR as well as the genetic architecture of WHR in general are still poorly understood. We therefore conducted sex-combined and sex-stratified meta-analyses comprising >210,000 individuals (>116,000 women; >94,000 men) of European ancestry from 57 GWAS studies and 28 studies genotyped on the MetaboChip within the GIANT consortium. The sex-combined analysis yielded 39 loci with genome-wide significant association (P<5x10-8), of which 11 loci showed significant sex-difference (Bonferroni-corrected P<0.05/39). Six of these loci influence WHR in women only without any effect in men (near*COBLL1, LYPLAL1, PPARG, PLXND1, MACROD1, FAM13A*); four loci have an effect in women and a less pronounced effect in men (near *VEGFA, ADAMTS9, HOXC13, RSPO3*); and one locus has only an effect in men (near *GDF5*). The sex-stratified analyses identified nine additional female-specific loci that had been missed in the sex-combined analysis due to the lack of effect in men (near *MAP3K1, BCL2, TNFAIP8, CMIP, NKX3-1, NMU, SFXN2, HMGA1, KCNJ2*). No additional loci were identified in the male-specific analysis. We confirmed all previously established sexually dimorphic variants for WHR. Of particular interest is the *PPARG* region that is a well-known target in type 2 diabetes treatments and shows a female-specific association with WHR. The enrichment of female-specific associations, i.e. 19 of the 20 sexually dimorphic loci, is consistent with the heritability of WHR as estimated in the Framingham Heart study; we found that WHR is more heritable in women (h2~46%) compared to men (h2~19%). Our results highlight the importance of sex-stratified analyses and can help to better understand the genetics underpinning the sex-differences of body fat distribution.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2148W

1000 Genomes imputation identifies low frequency-large effect circulating biomarker associations undetected by HapMap based imputation. *A.R. Wood[1], J. Tyrrell[1,2], D.G. Hernandez[3,4], M.A. Nalls[3], H. Yaghootkar[1], J.R. Gibbs[3,4], L.W. Harries[5], S. Chong[3], M. Moore[3], J.M. Guralnik[6], S. Bandinelli[7], A. Murray[1], L. Ferrucci[8], A.B. Singleton[3], D. Melzer[5], M.N. Weedon[1], T.M. Frayling[1].* 1) Genetics of Complex Traits, Peninsula College of Medicine and Dentistry, Exeter, United Kingdom; 2) European Centre for Environment and Human Health (ECEHH), Peninsula College of Medicine and Dentistry, Truro, United Kingdom; 3) Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA; 4) Department of Molecular Neuroscience and Reta Lila Laboratories, Institute of Neurology, UCL, London, United Kingdom; 5) Institute of Biomedical and Clinical Sciences, Peninsula College of Medicine and Dentistry, Exeter, United Kingdom; 6) Department of Epidemiology and Public Health, University of Maryland School of Medicine, Baltimore, MD, USA; 7) Geriatric Unit, Azienda Sanitaria di Firenze, Florence, Italy; 8) Clinical Research Branch, National Institute on Aging NIA-ASTRA Unit, Harbor Hospital, MD, USA.

Imputation from the sequence-based 1000 Genomes reference panel can potentially identify signals of association not detected by imputation from the genotype-based HapMap reference panel. We tested the hypothesis that imputation of variants identified by the 1000 Genomes Project (including low frequency and rare variants) can enhance previously identified associations and identify novel loci.

We selected 93 traits measured in ~1200 individuals from the InCHIANTI population study, a longitudinal study of aging men and women from Tuscany, Italy. All traits were continuous measures of circulating factors. These phenotypes included circulating lipids, proteins, ions and vitamins and were selected for two reasons. First, strong genetic associations have previously been observed in relatively small sample sizes, and so we potentially had good power to detect novel effects. Second, identifying genetic associations with circulating factors is potentially important to help understand their associations with disease.

We observed 21 signals (3 novel) and 33 signals (8 novel) at $P<5x10^{-8}$ based on HapMap and 1000 Genomes imputation respectively. Of the 13 signals specific to 1000 Genomes imputation, 6 had a HapMap SNP in modest linkage disequilibrium ($r^2 \geq 0.26$) close to GWAS significance ($P<5x10^{-5}$) Of the remaining 7 signals, 6 represented low-frequency (MAF < 2%) large-effect variants with corresponding effect sizes (0.51–2.1 SDs) and were poorly captured by the leading HapMap-imputed SNPs at the loci ($r^2 < 0.034$). One of these associations was between rs28929474 in the *SERPINA1* gene and Alpha-1-antitrypsin deficiency (MAF 0.007) and was previously detected in a candidate gene study, providing validation of our result. Of the 14 signals detected in both HapMap and 1000 Genomes imputation, 10 were marginally stronger in the 1000 Genomes analysis.

We show that 1000 Genomes based imputation can find more putative loci with variants of lower frequency and larger effects compared to HapMap based imputation.

2149T

So fresh or so clean? A genetic variant near olfactory receptor genes associates with cilantro taste preference. *S. Wu, J.Y. Tung, A.B. Chowdry, A.K. Kiefer, J.L. Mountain, N. Eriksson.* 23andMe, Inc., Mountain View, CA.

The leaves of the *Coriandrum sativum* plant, known as cilantro or coriander, are widely used in many cuisines around the world. However, far from being a benign culinary herb, cilantro can be highly polarizing — many people love the taste while others claim that it tastes foul, like soap or dirt. Cilantro aroma is due largely to the presence of several aldehydes; similar aldehydes are found in soaps, lotions, and certain insects. Cilantro taste preference is suspected to have a genetic component, yet to date very little has been discovered about specific mechanisms.

Here we present the results of a genome-wide association study of cilantro preference using genotype and survey response data from customers of 23andMe, Inc., a personal genomics and genetic research company. Across this cohort, we find ethnic differences in cilantro taste preference. South Asians and Latinos were both significantly ($p < 1x10^{-8}$) more likely to enjoy the taste of cilantro (OR=2.32 and 1.75 respectively, controlled for age and sex). Women are less likely to enjoy cilantro (OR=0.84, $p=4x10^{-9}$). In addition, within a European subset, we find that the SNP rs4036310 is significantly associated with both disliking cilantro and detecting a soapy taste ($p<5x10^{-8}$ for each, OR=1.14 and 1.25 respectively). This SNP lies within a cluster of olfactory receptor genes on chromosome 11, notably OR6A2 (also known as OR-I7), which has a high binding specificity for a number of aldehydes that provide the characteristic odor in cilantro, including E-2-dodecenal, E-2-Decenal, decanal, and dodecanal. We propose that OR6A2/OR-I7 contributes to the detection of a soapy taste from cilantro in a subset of people.

2150F

Genome-wide association study for Behçet's disease in an Iranian dataset using a DNA pooling approach. *J. Xavier[1,2], M. Martins[1,2], J. Sobral[1,2], F. Shahram[3], F. Davatchi[3], B.S. Abdollahi[3], A. Nadji[3], N.M. Shafiee[3], F. Ghaderibarim[3], S.A. Oliveira[1,2].* 1) Instituto de Medicina Molecular, Lisbon, Portugal, Portugal; 2) Instituto Gulbenkian de Ciência, Oeiras, Portugal; 3) Rheumatology Research Center, Tehran University of Medical Sciences, Tehran, Iran.

Behçet's disease (BD) is a complex multifactorial disorder with genetic and environmental risk factors implicated in its etiology, however BD pathology is poorly understood. Genome-wide association studies (GWAS) represent the gold-standard approach to study the genetic basis of complex diseases. Since genotyping individually thousands of polymorphisms in the large datasets required for the study of complex traits is time-consuming, labor-intensive, and cost-prohibitive, we used a recently developed and validated strategy that combines DNA pooling and microarray genotyping.

300 Iranian cases and 300 Iranian controls age and gender-matched were used in the first phase of the GWAS. Two pools of cases and two pools of controls were genotyped in quadruplicate using the Affymetrix Human SNP Array 6.0. Two different statistical approaches were applied to select the SNPs for technical validation: a) the RAS (relative allele score) difference between the average of cases and controls and b) a two sample, equal variance Student's t-test between the 8 arrays of cases and controls. Our results were compared to the results obtained from two previously published GWAS.

We identified 28 autosomal SNPs with RAS difference ≥ 10% and 45 SNPs with a *P*-value ≤5.0E-8. For both strategies, the top SNPs are located in the 6p21.33 locus, confirming the known HLA-B51 association with BD and supporting both strategies. For the RAS difference and t-test approaches there were, respectively, 7 and 5 out of 50 SNPs, previously found associated in the published GWASs. Additionally, we identified significant independent evidence of association with BD at loci 19q13.42 and 1q23.2 using both strategies. These SNPs are currently being individually genotyped in the 600 samples for technical validation. The validated SNPs will be further genotyped in a second independent Iranian sample consisting of 676 BD cases and 539 controls and in a Portuguese replication dataset with 75 cases and 120 controls. We have therefore confirmed well established associations with BD and possibly identified new susceptibility genes.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2151W

A genome-wide association study identifies susceptibility loci for non-syndromic sagittal craniosynostosis on chromosomes 20 and 7. *G. Yagnik[1, 2], C.M. Justice[3, 2], Y. Kim[3], I. Peter[4], E.W. Jabs[4], X. Ye[4], L. Shi[4], M.L. Cunningham[5], V. Kimonis[6], T. Roscioli[7], S.A. Wall[8], A.O.M Wilkie[8, 9], J. Stoler[10], J.T. Richtsmeier[11], Y. Heuzé[11], P.A. Sanchez-Lara[12], M.F. Buckley[13], C.M. Druschel[14], J.L. Mills[15], M. Caggana[16], P.A. Romitti[17], D.M. Kay[16], C. Senders[18], P.J. Taub[19], O.D. Klein[20], J. Boggan[21], C. Naydenov[22], J. Kim[1], A.F. Wilson[3], S.A. Boyadjiev[1].* 1) Department of Pediatrics, Section of Genetics, University of California, Davis, Sacramento, CA; 2) Authors with equal contribution; 3) Genometrics Section, IDRB, Division of Intramural Research, NHGRI, NIH, Baltimore, MD; 4) Department of Genetics and Genomic Sciences, Mount Sinai School of Medicine, New York, NY; 5) Department of Pediatrics, Division of Craniofacial Medicine, University of Washington and Seattle Children's Research Institute, Seattle, WA; 6) Division of Genetics, Department of Pediatrics, University of California Irvine, Irvine, CA; 7) School of Women's and Children's Health, Sydney Children's Hospital, University of New South Wales, Sydney, Australia; 8) Craniofacial Unit, Oxford University Hospitals NHS Trust, John Radcliffe Hospital, Oxford OX3 9DU, UK; 9) Weatherall Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Oxford OX3 9DS, UK; 10) Division of Genetics, Children's Hospital Boston, Harvard University, Boston, MA; 11) Department of Anthropology, Pennsylvania State University, University Park, PA; 12) Division of Genetics, Department of Pediatrics, University of South California, Los Angeles, CA; 13) Department of Haematology and Genetics, SEALS, Sydney, Australia; 14) Congenital Malformations Registry, New York State Department of Health, Albany, NY; 15) Division of Epidemiology, Statistics, and Prevention Research, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Department of Health and Human Services, Bethesda, MD; 16) Division of Genetics, Wadsworth Center, New York State Department of Health, Albany, NY; 17) Department of Epidemiology, College of Public Health, The University of Iowa, Iowa City, IA; 18) Department of Otolaryngology, University of California Davis, Sacramento, CA; 19) Division of Plastic and Reconstructive Surgery, Kravis Children's Hospital, Mount Sinai Medical Center, New York, NY; 20) Departments of Orofacial Sciences and Pediatrics and Program in Craniofacial and Mesenchymal Biology, University of California San Francisco, San Francisco, CA; 21) Department of Neurological Surgery, University of California Davis, Sacramento, CA; 22) Department of Chemistry and Biochemistry, Medical University, Sofia, Bulgaria.

Sagittal craniosynostosis is a common congenital malformation, affecting approximately one out of 5,000 newborns. We conducted the first genome-wide association study (GWAS) for non-syndromic sagittal craniosynostosis (sNSC) using 130 non-Hispanic white (NHW) case-parent trios. Robust associations were observed in a 120 kb region downstream of *BMP2* on chromosome 20p.12.3, flanked by rs1884302 ($P \leq 1.13 \times 10^{-14}$; odds ratio [OR] = 4.58) and rs6140226 ($P \leq 3.4 \times 10^{-11}$; OR = 0.24) and within a 167 kb region of *BBS9* on chromosome 7p14.3 between rs10262453 ($P \leq 1.61 \times 10^{-10}$; OR=0.19) and rs17724206 ($P \leq 1.50 \times 10^{-8}$; OR = 0.22). We replicated these associations for rs1884302 ($P \leq 4.39 \times 10^{-31}$) and rs10262453 ($P \leq 3.50 \times 10^{-14}$) in an independent NHW population of 172 unrelated sNSC probands and 548 unaffected controls. Both *BMP2* and *BBS9* implicated by these associations are biologically plausible genes with a role in skeletal development warranting functional studies to further understand the etiology of sNSC.

2152T

Finding genome-transcriptome-phenome association with structured association mapping and visualization in GenAMap. *J. Yin[1], R.E. Curtis[2], E.P. Xing[1,3].* 1) Lane Center for Computational Biology, Carnegie Mellon University, Pittsburgh, PA; 2) Joint Carnegie Mellon-University of Pittsburgh PhD Program in Computational Biology, Pittsburgh, PA; 3) Machine Learning Department, Carnegie Mellon University, Pittsburgh, PA.

Despite the success of genome-wide association studies in detecting novel disease variants, we are still far from a complete understanding of the mechanisms through which variants cause disease. Most of previous studies have considered only genome-phenome associations. However, the integration of transcriptome data may help further elucidate the mechanisms through which genetic mutations lead to disease and uncover potential pathways to target for treatment. We present a novel structured association mapping strategy for finding genome-transcriptome-phenome associations when SNP, gene-expression, and phenotype data are available for the same cohort. We do so via a two-step procedure where genome-transcriptome associations are identified by GFlasso, a sparse regression technique presented previously. Transcriptome-phenome associations are then found by a novel proposed method called gGFlasso, which leverages structure inherent in the genes and phenotypic traits. Due to the complex nature of three-way association results, visualization tools can aid in the discovery of causal SNPs and regulatory mechanisms affecting diseases. Using well- grounded visualization techniques, we have designed new visualizations that filter through large three-way association results to detect interesting SNPs and associated genes and traits. The two-step GFlasso-gGFlasso algorithmic approach and new visualizations are integrated into GenAMap, a visual analytics system for structured association mapping. Results on simulated datasets show that our approach has the potential to increase the sensitivity and specificity of association studies, compared to existing procedures that do not exploit the full structural information of the data. We report results from an analysis on a publically available mouse dataset, showing that identified SNP-gene-trait associations are compatible with known biology.

2153F

A genome-wide association study of central corneal thickness in Latinos. *X. Gao[1,2], W.J. Gauderman[2], P. Marjoram[2], M. Torres[1,2], T. Haritunians[3], J.Z. Kuo[3], Y.I. Chen[3], K.D. Taylor[3], J.I. Rotter[3], R. Varma[1,2].* 1) Department of Ophthalmology, Keck School of Medicine, University of Southern California, Los Angeles, CA; 2) Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA; 3) Medical Genetics Institute, Cedars-Sinai Medical Center, Los Angeles, CA.

Central corneal thickness (CCT) is a risk factor for primary open angle glaucoma and keratoconus. Genetic factors controlling CCT in Latinos, the most populous minority population in the US, are unclear. Here we describe the first genome-wide association study (GWAS) report of CCT in Latinos. We performed a population-based GWAS for CCT on 1,779 Latinos recruited in the Los Angeles Latino Eye Study (LALES) using the Illumina HumanOmniExpress Beadchip (~730K markers). To further discover additional associated single nucleotide polymorphisms (SNPs), we imputed SNPs based on the 1,000 Genomes Project reference panels. All subjects were age 40 years and older. We used linear regression with adjustment for age, sex and principal components of genetic ancestry. Statistical analyses were performed using PLINK, R and SAS. We replicated the involvement of several previously reported genes, e.g. COL5A1, FOXO1 and ZNF469, for CCT in Latinos ($P < 0.005$). Moreover, we discovered the involvement of novel SNPs that reached GWAS significance in the uncharacterized LOC100506532 (gene type: miscRNA) for CCT in Latinos ($P = 2.49 \times 10^{-9}$). Genetic analysis in Latinos may lead to the discovery of population-specific loci that regulate CCT, offering insights in the molecular regulatory mechanism of CCT. These results help to further describe the genetic architecture of CCT, demonstrating the utility of using Latinos for GWAS analyses.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2154W

CDKN2BAS Genotype - Primary Open Angle Glaucoma Feature Correlations in the GLAUGEN Study and the NEIGHBOR Consortium. *L.R. Pasquale[1, 2], S.J. Loomis[1, 2], J.H. Kang[2], B.L. Yaspan[3], D.L. Budenz[4], D.S. Friedman[5], D. Gaasterland[6], T. Gaasterland[7], R.K. Lee[8, 18], P.R. Lichter[9], Y. Liu[10, 16], C.A. McCarty[11], S. Moroi[9], L.M. Olson[3], T. Realini[12], J.S. Schuman[13], K. Singh[14], D. Vollrath[14], G. Wollstein[13], D.J. Zack[5], W. Abdrabou[1], M.A. Hauser[10, 16], E. Delbono[1], R.R. Allingham[10], M.A. Pericack-Vance[8, 18], R.N. Weinreb[15], K. Zhang[15], J.E. Richards[9,17], J.L. Haines[3], J.L. Wiggs[1].*
1) Glaucoma Service, Massachusetts Eye and Ear Infirmary, Boston, MA; 2) Channing Division of Network Medicine, Brigham and Women's Hospital, Boston, MA; 3) Center for Human Genetics Research, Vanderbilt University, Nashville, TN; 4) Department of Ophthalmology, University of North Carolina, Chapel Hill, NC; 5) Wilmer Eye Institute, Johns Hopkins University Hospital, Baltimore, MD; 6) Eye Doctors of Washington, Chevy Chase, MD; 7) Scripps Genome Center, University of California at San Diego, San Diego, CA; 8) Bascom Palmer Eye Institute, University of Miami Miller School of Medicine, Miami, FL; 9) Department of Ophthalmology and Visual Sciences, University of Michigan, Ann Arbor, MI; 10) Department of Ophthalmology, Duke University Medical Center, Durham, NC; 11) Essentia Institute of Rural Health, Duluth, MN; 12) Department of Ophthalmology, West Virginia University Eye Institute, Morgantown, WV; 13) Department of Ophthalmology, UPMC Eye Center, University of Pittsburgh, Pittsburgh, PA; 14) Department of Ophthalmology, Stanford University, Palo Alto, CA; 15) Department of Ophthalmology, University of California at San Diego, San Diego, CA; 16) Department of Medicine, Duke University Medical Center, Durham, NC; 17) Department of Epidemiology, University of Michigan, Ann Arbor, MI; 18) Center for Human Genomics, University of Miami, Miami, FL.

PURPOSE: Loci at *CDKN2BAS* modify the risk of primary open-angle glaucoma (POAG). We assessed the association between *CDKN2BAS* single nucleotide polymorphisms (SNPs) and glaucoma features like cup-disc ratio (CDR). **DESIGN**: Observational case series. **METHODS**: We studied associations between ten *CDKN2BAS* SNPs and various glaucoma features among 976 POAG cases from the Glaucoma Genes and Environment (**GLAUGEN**) study and 1971 cases from the NEI Glaucoma Human Genetics Collaboration (**NEIGHBOR**) consortium. For each patient, we chose the feature from the eye with the higher value. We created multivariable models for glaucoma features adjusting for appropriate covariates in each study separately, and then pooled results with meta-analytical techniques. **RESULTS:** Cases with GA or GG genotypes at rs2157719 (associated with reduced POAG risk) had smaller CDR at diagnosis (0.05 units smaller per G allele; 95% CI: -0.08, -0.03; p=6.23E-05) and higher intraocular pressure (IOP) at the time of DNA collection (+0.70 mm Hg per G allele; 95% CI: 0.40, 1.00; p=5.45E-06). The other protective *CDKN2BAS* SNPs showed similar trends with CDR and IOP. Cases with AG or AA genotypes at rs3217992 (associated with increased POAG risk) had larger CDR at diagnosis (0.05 units larger per A allele; 95% CI: 0.02, 0.07; p=4.74E-04) and lower IOP at DNA collection (-0.57 mm Hg per A allele; 95% CI: -0.84, -0.29; p=6.55E-05). There were also other significant relations between *CDKN2BAS* SNPs and some visual field features. **CONCLUSION:** Among POAG subjects, *CDKN2BAS* alleles are associated with CDR and IOP but not with most other glaucoma features.

2155T

Ethnicity and ancestry information from genome-wide association studies: The NHGRI GWAS Catalog. *H.A. Junkins[1], J.A.L. MacArthur[2], P. Hall[1], K.A. Harvey[3], T.A. Manolio[1], L.A. Hindorff[1].* 1) Office of Population Genomics, NHGRI, National Institutes of Health, Bethesda, MD; 2) EMBL-EBI, Wellcome Trust Genome Campus, Hinxton, UK; 3) Centers for Disease Control and Prevention, Atlanta, GA.

The impact of genome-wide association studies (GWAS) on the knowledge base of disease-associated variants has been vast, with over 1,300 publications and 6,400 disease- or trait-associated genetic variants identified to date (http://www.genome.gov/gwastudies). However, much of this knowledge is based upon studies of predominantly European-descent individuals, posing the question of whether these variants will also generalize to non-European populations and, importantly, to populations with disproportionate disease burdens. As part of the curation efforts for the National Human Genome Research Institute GWAS Catalog, we developed a framework for systematically and consistently extracting ethnicity and ancestry information at multiple levels of granularity. Of 443 papers from January, 2011 through April, 2012 (35% of all papers), 56% exclusively report populations of European descent; 11% include both European and non-European populations, and 28% report only non-European populations. On a per-participant level, most participants were also classified as European descent. Of studies that included participants of non-European descent, the most frequently represented countries were China, Japan, Korea, and Singapore. Representation of non-European participants was most pronounced for the following traits: blood pressure, body mass index, HDL cholesterol and type 2 diabetes. Papers including non-European participants were more likely to be published in high impact journals (impact factor >7; OR 1.64, 95% CI 1.04–2.60), adjusted for study year, sample size, and previous GWAS publications. Preliminary data from a small study of 153 papers published from 2005–2008 suggest that a variety of statistical methods were used to account for population stratification in the analysis of GWAS data; an increased number of such methods was associated with use of more than one ethnic group in the discovery set, later publication date and publication in high-impact journals. These data need to be extended to more recent studies, which through meta-analyses of multiple cohorts have increased in complexity. Although these data are limited to GWAS published in 2011–2012, our work suggests that non-European participants are still greatly underrepresented in GWAS and that additional studies are needed to fully evaluate whether potentially functional variants identified in GWAS generalize to populations most impacted by chronic disease.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2156F

OFCD syndrome gene BCOR is associated with dental caries. *X. Wang[1], Z. Zeng[2,3], J.R. Shaffer[2], E. Feingold[2,3], D.E. Weeks[2,3], M. Lee[1], K. T.Cuenco[1,2], R.J. Weyant[4], R. Crout[5], D.W. McNeil[6], M.L. Marazita[1,2,7].* 1) Dept Oral Biology, Univ Pittsburgh, Pittsburgh, PA; 2) Dept Human Genetics, Univ Pittsburgh, Pittsburgh, PA; 3) Dept Oral Biol, Biostatistics, Pittsburgh, PA; 4) Dept Dental Public Health and Information Management, Univ Pittsburgh, Pittsburgh, PA; 5) Dept Periodontics, West Virginia Univ, Morgantown, WV; 6) Dental Practice and Rural Health, West Virginia Univ, Morgantown, WV; 7) Clinical and Translational Science, Univ Pittsburgh, Pittsburgh, PA.

Dental caries (i.e. tooth decay) is one of the most common diseases worldwide. In high-risk populations, the caries prevalence in adults can reach 90%. While treatment of caries costs tremendous amounts each year, untreated caries lesions can lead to pain, tooth loss, and oral infection. Genetic contributions to the etiology of dental caries are well-established. However, few specific caries genes have been successfully discovered and validated. Recent family-based analyses suggested differential genetic factors for primary dentition (i.e., baby teeth) caries and permanent dentition (i.e., adult teeth) caries, as well as for pit and fissure surface and smooth surface caries. Here we sought to find genes preferentially responsible for caries on these two surface types in the permanent dentition. We performed separate genome-wide association studies on caries scores for the above two surface types in the permanent dentition. We used a sample of 1,063 individuals of European ancestry (ages 14 to 56 years), adjusting for the effects of age, sex, presence of Streptococcus mutans (an oral bacteria), and home water fluoride level. More than 1.2 million SNPs were either successfully genotyped or imputed and were tested for association. The BCOR gene, mutations in which are known to cause oculofaciocardiodental syndrome (a condition involving multiple dental anomalies), was suggestively associated with pit and fissure surface caries (p-value = 3.89E-7), while the highly homologous gene, BCORL1, was suggestively associated with smooth surface caries (p-value = 5.54E-6). Associations of other genes with functional plausibility in cariogenesis were also observed for pit and fissure surface caries (INHBA, p-value = 5.14E-6) and for smooth surface caries (CXCR1 and CXCR2, p-value = 1.48E-6). Not only did our study nominate susceptibility genes for dental caries on the two types of surfaces in the permanent dentition but demonstrated that surface-specific phenotypes have potential to increase the power for detecting genes with surface-specific effects despite the presence of a rather high genetic correlation between the two phenotypes. Our findings contribute to a better understanding of genetic pathogenic mechanisms of cariogenesis and eventually, may lead to better prevention and prognosis of dental caries. Supported by NIH/NIDCR U01-DE018903, R01-DE014899, R03-DE021425.

2157W

Sickle-cell locus is a major modifier of human susceptibility to invasive bacterial infection - a genome-wide association study of bacteremia susceptibility. *A. Rautanen[1], M. Pirinen[1], C. Spencer[1], T. Mills[1], S. Chapman[1], K. Rockett[1], J.A. Scott[2,3], T.N. Williams[3,4], P. Donnelly[1], A.V. Hill[1].* 1) WTCHG, University of Oxford, Oxford, United Kingdom; 2) KEMRI-Wellcome Trust Programme, Kilifi, Kenya; 3) Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom; 4) INDEPTH Nettwork, Accra, Ghana.

Bacteremia, a bacterial infection of the bloodstream, is a leading cause of death among children in sub-Saharan Africa. As part of the Wellcome Trust Case Control Consortium 2, we have conducted a genome-wide association study of 1536 microbiologically confirmed bacteremia cases and 2677 controls, recruited from Kilifi, Kenya. This is probably the largest DNA collection from patients with bacteremia world-wide, the first GWAS of bacteremia susceptibility that we know of, and one of the few large-scale GWAS conducted in African populations. None of the genotyped SNPs reached genome-wide significance but imputation against the 1000 Genomes data revealed 4 loci with a $p<5x10^{-8}$. The strongest association was found on chromosome 11 (genotypic $p=4.8x10^{-13}$) in the HBB (beta hemoglobin) gene region and the sickle cell disease causing mutation rs334 was among the most strongly associated SNPs. As the reference panel does not perfectly reflect the Kenyan populations, we confirmed the imputation result by direct genotyping. We replicated the HBB finding in 769 cases and 4337 controls from the same population in Kenya, resulting in a highly significant combined genotypic p-value of $1.1x10^{-21}$. Heterozygotes (sickle-cell trait, HbAS) were protected against bacteremia ($p=8.3x10^{-7}$, OR=0.68, compared to HbAA), whereas minor homozygote individuals (sickle-cell disease, HbSS) were highly predisposed to bacteremia ($p=3.07x10^{-16}$, OR=3.94, compared to HbAA). In addition to bacteremia overall, sub-groups stratified by causative bacteria were analysed: gram-positive, gram-negative, *S. pneumoniae*, *Non-typhoidal salmonella*, *S. aureus*, *H. influenzae*, and *E. coli*being the largest sub-groups. These analyses showed that HbSS most strongly increases the risk of invasive infection caused by encapsulated bacteria whereas protection by HbAS is strongest against E. coli infection. Replication genotyping of the other significantly associated loci is ongoing. This study demonstrates the importance of imputation in African GWAS, even when the reference panel does not perfectly represent the ethnicities under study. HbAS is highly protective against malaria and malaria may predispose to bacteremia in the malaria-endemic region, which could partly explain the heterozygote protection. However, the strong risk of bacteremia conferred by the HbSS genotype is likely independent of malaria as the risk does not change over time, even when the incidence of malaria has dramatically decreased.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2158T

Using genomics to decipher causation in severe edematous childhood malnutrition: a pilot GWAS. *N.A. Hanchard[1,2], S. Howell[3], K. Marshall[3], T. Forrester[3], M. Reid[4], X. Wang[1,2], L.M. Franco[1,2,5], J.W. Belmont[1,2], C.A. McKenzie[3].* 1) Molecular & Human Genetics, Baylor College of Medicine, Houston, TX; 2) USDA/ARS Children's Nutrition Research Center, Baylor College of Medicine, Houston, TX; 3) Tropical Metabolism Research Unit, Tropical Medicine Research Institute, University of the West Indies, Kingston, Jamaica; 4) Sickle Cell Unit, Tropical Medicine Research Institute, University of the West Indies, Kingston, Jamaica; 5) Department of Medicine, Baylor College of Medicine, Houston, TX.

Background: Severe childhood malnutrition (SCM) is an important global health problem that contributes to more than two million deaths worldwide each year. SCM exists in two well-described and clinically distinct forms - edematous SCM (ESCM), which has a higher mortality and is more difficult to treat, and non-edematous SCM (NESCM). The appropriate development of prevention and treatment protocols for SCM requires an understanding of the reasons why some children develop ESCM while others develop NESCM, even though both groups are exposed to the same apparent nutritional stress. The approach to understanding this clinical dichotomy has largely focused on physiological observations of biochemical, inflammatory, and nutritional markers during and after nutritional recovery. We have advocated the application of genome-wide technologies to SCM as a means of circumventing the problem of reverse causation inherent in such physiological studies, and to provide a global integrated view of the physiological response associated with severe nutrient and caloric restriction. Methods and Results: We genotyped ~1.3 million SNPs in 101 ESCM cases and 103 NESCM controls using the Illumina HumanOmni1-Quad BeadChip array. Patterns of linkage disequilibrium, minor allele frequency, and multi-dimensional scaling indicated a close ethnic relationship between the Jamaican samples and the Yoruba of West Africa. After imposition of data quality control filters, 91 ESCM cases and 88 NESCM controls were used for genome-wide association tests under an allelic model. There was no evidence of population stratification (GIF=1.0). No markers met genome-wide significance for association (P<5 x 10-8); however, at a suggestive statistical significance level (P<1 x 10-5) we observed a cluster of SNPs at the Glucan (1,4-alpha-) branching enzyme (GBE1) gene (maximum OR 2.64, 95%CI 1.72–4.05) that were associated with ESCM status. Pathway analysis implicated genes involved in carbohydrate and inositol phosphate metabolism, ErbB signaling, and immune cytotoxicity in ESCM. Conclusions: We have undertaken the first genome-wide association study of SCM. The results from this pilot study illustrate the potential of this approach to further inform the pathophysiology of ESCM and suggest a more prominent role for glycogen and carbohydrate metabolism in ESCM. Larger studies are required to confirm and validate these findings.

2159F

Genomewide Association Study of Body Mass Index in Samoans. *R.L. Minster[1], G. Sun[2], S.R. Indugula[2], H. Cheng[2], N.L. Hawley[3], S. Viali[4], R. Deka[2], D.E. Weeks[1], S.T. McGarvey[3].*1) Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA; 2) Department of Environmental Health, School of Medicine, University of Cincinnati, Cincinnati, OH; 3) Department of Epidemiology, International Health Institute, Program in Public Health, Brown University, Providence, RI; 4) Ministry of Health, Apia, Samoa.

Compared to most populations worldwide, obesity is considerably more prevalent in Polynesian populations, and it has been speculated that this may be due to greater genetic susceptibility, especially upon exposure to the Western diet and decreased activity. Genomewide association studies (GWAS) of obesity-related traits have been conducted in many populations, including East Asians, but none, to date, have been completed in Polynesian populations. Such studies might highlight obesity-related variants unique to Oceanic populations. Here we report the results of a GWAS of body mass index (BMI) in 3,122 adults from Samoa. A total of 871,188 autosomal single-nucleotide polymorphisms (SNPs) were genotyped and extensive quality control of the genotypes conducted. One region was associated with BMI at genomewide significance ($p < 5 \times 10^{-8}$) on chromosome 5q35.1 (most significant SNP $p = 2.54 \times 10^{-14}$). Thirteen SNPs in this region were associated with BMI at suggestive levels of significance ($p < 10^{-5}$). None of the nearby genes have been described as associated with BMI before. SNPs in five other regions were also suggestively associated with BMI: 10q22, 10q25.2, 13q13.2, 13q14 and 15q26.2. Additional studies of Polynesian populations will be needed to confirm these findings.

2160W

Variants in *RUNX3* contribute to susceptibility to psoriatic arthritis exhibiting further common ground with ankylosing spondylitis. *A. Reis[1], M. Apel[1], S. Uebe[1], J. Bowes[2], E. Giardina[3], E. Korendowych[4], K. Juneblad[5], F. Pasutto[1], A.B. Ekici[1], R. McManus[6], P. Ho[2], I.N. Bruce[2], A.W. Ryan[6], F. Behrens[7], B. Böhm[7], H. Traupe[8], J. Lohmann[9], C. Gieger[10], H.E. Wichmann[10,11,12], L. Padyukov[13], O. FitzGerald[14], G.M. Alenius[5], N. McHugh[4,15], G. Novelli[3,16], H. Burkhardt[7], A. Barton[2], U. Hüffmeier[1].* 1) Institute of Human Genetics, Erlangen, Germany; 2) Arthritis Research UK Epidemiology Unit, Manchester Academic Health Science Centre, The University of Manchester, UK; 3) Department of Biopathology, Centre of Excellence for Genomic risk Assessment in Multifactorial and Complex Diseases, School of Medicine, University of Rome “Tor Vergata” and Fondazione PTV “Policlinico Tor Vergata”, Italy; 4) Royal National Hospital for Rheumatic Diseases, Bath, UK; 5) Department of Public Health and Clinical Medicine/ Rheumatology, University Hospital, Umeå, Sweden; 6) Department of Clinical Medicine, Institute of Molecular Medicine, Trinity College Dublin, Ireland; 7) Division of Rheumatology, Department of Internal Medicine II, Johann Wolfgang Goethe University, Frankfurt am Main, Germany; 8) Department of Dermatology, University of Münster, Germany; 9) Rehabilitation Hospital, Bad Bentheim, Germany; 10) Institute of Genetic Epidemiology, Helmholtz Center Munich, Germany; 11) Institute of Medical Informatics, Biometry and Epidemiology, Chair of Epidemiology, Ludwig-Maximilians-Universität, Munich, Germany; 12) Klinikum Grosshadern, Munich, Germany; 13) Rheumatology Unit, Department of Medicine, Karolinska Institute, Stockholm, Sweden; 14) Department of Rheumatology, St. Vincent's University Hospital, UCD School of Medicine and Medical Science and Conway Institute of Biomolecular and Biomedical Research, University College Dublin, Ireland; 15) Department of Pharmacy and Pharmacology, University of Bath, UK; 16) National Agency for Evaluation of Universities and Research, Rome, Italy.

Psoriatic arthritis (PsA) is a common inflammatory joint disease distinct from other chronic arthritides and frequently accompanied by psoriasis vulgaris (PsV). In a first genome-wide association study (GWAS), we were able to identify several genetic risk factors. However, even combined with previously identified factors, the genetic contribution to disease is not fully explained. Therefore we investigated further 17 loci from our GWAS which did not reach genome-wide significance levels of association in the initial analysis. 19 of 20 single nucleotide polymorphisms (SNPs) were successfully genotyped in independent cohorts of 1,398 PsA cases and 6,389 controls, and a group of 964 German PsV patients. Association to a *RUNX3*variant, rs4649038, was replicated in independent cases and controls and resulted in a combined p-value of 1.40E-08 in a Cochran-Mantel-Haenszel test and an OR of 1.24 (1.15–1.33). Further analyses based on linkage disequilibrium at *RUNX3* refined the most significant association to a linkage disequilibrium (LD) block of ~30kb located in the first intron of one isoform (p=3.40E-08). Weaker evidence for association was detected in German PsV patients (p-value 5.89E-02; OR 1.13 (1.00–1.28)), indicating a possible role in the skin manifestations of psoriasis. Our analyses identified variants in *RUNX3* as susceptibility factors for PsA. *RUNX3* has already been implicated in susceptibility to ankylosing spondylitis, another spondylarthropathy, although its risk allele seems to be independent from that in PsA. RUNX3 is involved in CD8 lymphocyte differentiation and therefore a good candidate for PsA and PsV as T-cell mediated diseases.

2161T

Association analysis between X-chromosome variants and asthma in 50,000 individuals. *C.S. Tang[1], D. Hinds[2], M.A.R. Ferreira[1].* 1) Asthma Genetics, Queensland Institute of Medical Research, Brisbane, QLD, Australia; 2) 23 and Me, Mountain View, California, United States.

Gender differences in asthma development and progression have been widely reported. Recent studies have indicated that more males are affected during early childhood but a gender shift has been observed at puberty, with a higher prevalence of asthma in female than male at adulthood. Factors contributing to such sex difference remain unclear. It has been suggested that hormonal changes together with genetic predisposition might be involved. Despite genes on the sex chromosomes contributing to many sexually dimorphic traits, associations on the X-chromosome were often overlooked in previous genome-wide association scans (GWAS) on asthma. To determine if genetic variants on the X-chromosome contribute to disease risk, we performed a meta-analysis of association studies comprising >50,000 individuals and interrogating >150,000 genotyped or imputed SNPs. Mixed-gender analysis revealed putative novel common variants associated with asthma, mapping to Xq22.3 and Xq27.3. Variants in *IL1RAPL1* were associated with asthma in females only. SNPs from these loci have been selected for replication in independent cohorts. Results from these analyses will presented.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2162F

Early adiposity phenotypes associate with common variants at LEPR, FTO, TFAP2B and GNPDA2. *M-R. Jarvelin[1,12,20,21], S. Das[1], H-R. Taal[2,3,4], U. Sovio[5], D. Mook-Kanamori[2,3,4], N.M. Warrington[6,7], N.J. Timpson[8], D.L. Cousminer[9], E. Thiering[10], I.Y. Millwood[1,24], J.L. Buxton[11], S. Sebert[1], M. Kaakinen[12,13], A.I.F. Blakemore[11], J. Eriksson[14,15,16], P. Froguel[22,23], M. McCarthy[17,18,19], L. Coin[1], P. O'Reilly[1], V.W.V. Jaddoe[2,3,4], Early Growth Genetics Consortium (EGG).* 1) Epidemiology & Public Health, Imperial Col London, London, United Kingdom; 2) Department of Epidemiology, Eramus Medical Center, Rotterdam, The Netherlands; 3) Department of Paediatrics, Erasmus Medical Center, Rotterdam, The Netherlands; 4) The Generation R Study Group, Erasmus Medical Center, Rotterdam, The Netherlands; 5) Department of Medical Statistics, London School of Hygiene and Tropical Medicine, London, UK; 6) School of Women's and Infants' Health, The University of Western Australia, Perth, Australia; 7) Samuel Lunenfeld Research Institute, University of Toronto, Toronto, Canada; 8) MRC Centre for Causal Analyses in Translational Epidemiology, School of Social and Community Medicine, University of Bristol, Bristol, UK; 9) Institute for Molecular Medicine Finland, University of Helsinki, Helsinki, Finland; 10) Institute of Epidemiology I, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany; 11) Section of Investigative Medicine, Division of Diabetes, Endocrinology and Metabolism,Imperial College London; 12) Institute of Health Sciences, University of Oulu, Finland; 13) Biocenter Oulu, University of Oulu, Finland; 14) National Institute for Health and Welfare, Helsinki, Finland; 15) Department of General Practice and Primary Health Care, University of Helsinki, Helsinki, Finland; 16) Folkhalsan Research Centre, Biomedicum Helsinki, University of Helsinki, Helsinki, Finland; 17) Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Churchill Hospital, Oxford, UK; 18) Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK; 19) Oxford NIHR Biomedical Research Centre, Churchill Hospital, Oxford, UK; 20) National Institute for Health and Welfare, Oulu, Finland, Biocenter Oulu, University of Oulu, Finland; 21) MRC Health Protection Agency (HPA) Centre for Environment and Health, Imperial College London, UK; 22) Department of Genomics of Common Disease, Imperial College London,UK; 23) CNRS8199, University of Lille North of France; 24) 24. Clinical Trial Service Unit and Epidemiological Studies Unit, University of Oxford, Oxford, UK.

Variations in early growth patterns are associated with obesity, type 2 diabetes and cardiovascular diseases but the genetic basis of this relationship is unknown. We conducted a Genome Wide Association Study (GWAS) of early adiposity parameters (age and body mass index (BMI) at adiposity peak (AP) and at adiposity rebound (AR)) in up to 7,215 children from 4 European birth cohorts with replication in up to 16,017 children from 12 independent studies. Early growth phenotypes were derived using sex-specific mixed effects models fitted on BMI measurements collected longitudinally from 2 weeks to 13 years. These phenotypes were regressed on ~2.5M single nucleotide polymorphisms (SNPs), genotyped/imputed, assuming an additive genetic model and adjusting for gestational age and sex. We identified four loci at p<5x10-8 and two suggestive at p<5x10-6 that associated with BMI at AP, age at AR and/or BMI at AR in overall meta-analyses. Two variants, one near the leptin receptor gene (LEPR; p= 1.23x10-9) and another, suggestive, near proprotein convertase subtilisin/kexin type I (PCSK1; p=1x10-6), associated with BMI at AP, occurring at ~9 months. LEPR encodes the cellular receptor for leptin, a hormone involved in regulating appetite and energy balance. PCSK1encodes an enzyme which is a key in the production of insulin and other hormones involved in regulating energy metabolism, and has been associated with BMI and obesity in adults. Three loci (FTO; p=3.12 x10-30, TFAP2B; p=1.17x10-9, GNPDA2; p= 2.92x10-8) associated with earlier age at AR and/or BMI at AR (~5.5y). Polymorphisms in all of these genes are associated with adult BMI and obesity in large-scale studies, with FTO also involved in BMI regulation in childhood. Another suggestive variant, near discs large homolog 2 (DLG2; p= 5.11x10-6), associated with BMI at AR. Variants in the DLG2 locus are associated with maintenance of glucose homeostasis through insulin sensitivity and β-cell function. We show that the loci identified associated with BMI in an age-dependent manner in childhood. We provide insight into the mechanisms linking early growth measures with later metabolic health and further evidence that the genetic variants associated with adult phenotypes may impact on disease development from early in life. Our methodological novelty is the use of phenotypes from longitudinal growth models. These methods are valuable for understanding bio-mechanisms for preventive purposes.

2163W

Multiple human genetic variants associated with Epstein-Barr virus load in the 1000 Genomes and HapMap lymphoblastoid cell lines. *C. Houldcroft[1], A. Gall[1], J.Z. Liu[1], A.L. Palser[1], S.J. Watson[1], C.A. Anderson[1], P. Kellam[1, 2].* 1) Virus Genomics, Wellcome Trust Sanger Institute, Cambridge, Cambs, United Kingdom; 2) Division of Infection & Immunity, University College London, London, United Kingdom.

Epstein-Barr virus (EBV) is a ubiquitous human gammaherpes virus, infecting around 99% of the adult population. It is the aetiological agent of infectious mononucleosis, and associated with a number of cancers such as Burkitt's lymphoma and nasopharyngeal carcinoma. EBV is also an important tool in molecular biology, used to transform white blood cells to derive lymphoblastoid cell lines (LCLs), permanent sources of DNA, RNA and protein. The virus has a two-stage life cycle, a lytic phase where the full repertoire of virus genes are expressed and primary infection of cells occurs; and a latent phase in which very few genes are expressed and virus immunogenicity is low (seen in latent natural infection). It is unclear what role host-genetic variants play in the switch from the latent to the lytic phase in LCLs. Epstein-Barr virus load also varies between LCLs and may be a confounding factor in studies of gene-expression performed on these cell lines. Here we characterise variation in virus load between cell lines and aim to understand the host genetics underlying Epstein-Barr virus load and reactivation from latency. This will allow future studies to take account of this variability and correct for it. Methods: 749 LCLs from 8 populations in the HapMap and 1000 Genomes projects were characterised for intra- and inter-sample EBV load variations. Virus loads were assayed by qPCR using absolute and relative quantification of genome copy numbers. A genome-wide association study was then conducted on 639 individuals genotyped for 500,000 SNPs to identify host genetic variants associated with EBV load using PLINK and EMMAX. Results: Absolute EBV loads varied from 2–350 copies per cell. Cell lines with the highest EBV loads showed evidence of viral lytic reactivation, with mRNA expression of immediate early lytic genes BZLF1 and BRLF1. Mixed-effect modelling using EMMAX showed a number of suggestive associations of host genetic variants with relative EBV loads in the HapMap LCLs. Top SNPs include rs6879787 (p = 4.12x10-8) and a number of associations on chromosome 6. We are currently increasing the sample size to boost our power to detect variants associated with the EBV load trait. Future work will include replication of associations found in this initial screen in an independent sample of LCLs. Microarray data is available for 270 HapMap cell lines and will be used to prioritise candidate genes identified by GWAS for functional validation.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2164T

Genome-wide association study on expressive language development in children from the general population links early word production to common variation at 3p12.3. *B. StPourcain[1], R. Cents[2], W. Ang[3], P.F. O'Reilly[4], F. Velders[2], N. Warrington[3], C.M.A. Haworth[5], O.S.P. Davis[5], A. Whitehouse[6], N. Timpson[1], J. Kemp[1], D.M. Evans[1], M. Marinelli[7], Y. Wren[8], S. Roulstone[8], R. Plomin[5], V. Jaddoe[2], M.-R. Jarvelin[4,9,10,11], C. Pennell[3], H. Tiemeier[2], G. Davey Smith[1], EAGLE.* 1) MRC CAiTE, University of Bristol, Bristol, England, United Kingdom; 2) Departments of Epidemiology and Child and Adolescent Psychiatry, Erasmus Medical Center, Rotterdam, The Netherlands; 3) School of Women's and Infants' Health, University of Western Australia, Perth, Australia; 4) Department of Epidemiology and Biostatistics, School of Public Health, MRC-HPA Centre for Environment and Health, Faculty of Medicine, Imperial College London, UK; 5) MRC Social, Genetic & Developmental Psychiatry Centre, Institute of Psychiatry, King's College London, London, UK; 6) Telethon Institute for Child Health Research, Centre for Child Health Research, University of Western Australia, Perth, Australia; 7) Center for Research in Environmental Epidemiology, Hospital del Mar Research Institute, Barcelona, Spain; 8) Bristol Speech and Language Therapy, Research Unit, North Bristol NHS Trust and The University of the West of England; 9) Institute of Health Sciences, University of Oulu, Oulu, Finland; 10) Biocenter Oulu, University of Oulu, Oulu, Finland; 11) Department of Lifecourse and Services, National Institute for Health and Welfare, Oulu, Finland.

The genetic architecture underlying expressive language (EL) development is largely unknown and so far no common genetic variants have been established. EL is a developmental phenotype that shows a dramatic change in growth, passing from an initial stage with few spoken words (<18 months, mean N of words ≤50) to a stage of rapid word acquisition (>18 months, mean N of words >50), which results in more complex language use thereafter. To identify common variation associated with EL development we focused on an early (ETW: 15–18 months) and a late time window (LTW: 24–30 months), using vocabulary lists of the Communicative Development Inventories and the Language Development Survey. A genome-wide (GW) analysis on EL in European descent children from the general population was conducted using up to 7280 children from ALSPAC or ALSPAC+RAINE cohorts, with variants from 19 (ETW) and 14 (LTW) loci followed up in up to 2038 independent children from the GENR cohort. In each study, we tested the association between the ranktransformed EL score, adjusted for sex, age and age^2, and single SNPs using linear regression. We performed fixed-effects inverse-variance meta-analyses to combine results from all studies. For the ETW, the analyses identified one novel locus on 3p12.3 near the roundabout axon guidance receptor homolog 2 (Drosophila) gene (*ROBO2*), which was associated with EL in ALSPAC (lead SNP: *P*=9.8E-07) and replicated in GENR (*P*=0.0044) reaching GW significance (Total N=8889, meta-*P*=1.3E-08). Variants at two further ETW loci on 11p15.2 and 12q15 were also replicated approaching suggestive levels of GW significance (lead SNPs: meta-*P*=2.0E-07; meta-*P*=7.2E-07). Investigation of the 3 ETW signals in 4772 12-month old infants from the NFBC1966 cohort, i.e. during a time where one-word utterances start, resulted in no such evidence for association. Likewise, a sensitivity analysis of these signals using LTW EL measures markedly attenuated or abolished the association, suggesting that the observed effects might be time-sensitive. We were not able to replicate signals for the LTW part of the GW analysis yet. Estimation of the proportion of additive phenotypic variation in ALSPAC, which is accounted for by all genotyped SNPs together, however supported the assumption that common genetic variation explains a significant proportion of variation in EL development both during early (ETW: h^2(SE)=0.13(0.05), *P*=0.008) and later stages (LTW: h^2(SE)=0.17(0.06), *P*=0.002).

2165F

A genome-wide association study of early menopause and the predictive impact of identified variants on risk. *J.R.B. Perry[1], T. Corre[2], T. Esko[3], S. Chanock[4], D. Chasman[5], E. Demerath[6], M. Garcia[7], S. Hankinson[8], F. Hu[8], D. Hunter[8], K. Lunetta[9], A. Metspalu[3], G. Montgomery[10], J. Murabito[11], A. Newman[12], K. Ong[13], T. Spector[14], K. Stefansson[15], A. Swerdlow[16], U. Thorsteinsdottir[15], R. Van Dam[17], A. Uitterlinden[18], J. Visser[18], P. Vollenweider[19], D. Toniolo[2], A. Murray[1] on behalf of the Reprogen Consortium.* 1) Gen Complex Traits, Peninsula Med Sch, Exeter, United Kingdom; 2) Division of Genetics and Cell Biology, San Raffaele Scientific Institute, Milan, Italy; 3) Estonian Genome Center, University of Tartu, Tartu, Estonia; 4) Division of Cancer Epidemiology and Genetics, National Cancer Institute, US National Institutes of Health (NIH), Bethesda, Maryland, USA; 5) Division of Preventive Medicine, Brigham and Women's Hospital, Boston, Massachusetts, USA; 6) Division of Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis, Minnesota, USA; 7) Laboratory of Epidemiology, Demography and Biometry, NIA, NIH, Bethesda, Maryland, USA; 8) Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts, USA; 9) Department of Biostatistics, Boston University School of Public Health, Boston, Massachusetts, USA; 10) Queensland Institute of Medical Research, Brisbane, Queensland, Australia; 11) Sections of General Internal Medicine, Preventive Medicine and Epidemiology, Department of Medicine, Boston University School of Medicine, Boston, Massachusetts, USA; 12) Departments of Epidemiology and Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania; 13) Medical Research Council (MRC) Epidemiology Unit, Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge, UK; 14) Department of Twin Research and Genetic Epidemiology, King's College London, London, UK; 15) deCODE Genetics, Reykjavik, Iceland; 16) The Institute of Cancer Research, London, UK; 17) Department of Nutrition, Harvard School of Public Health, Boston, Massachusetts, USA; 18) Department of Internal Medicine, Erasmus MC, Rotterdam, The Netherlands; 19) Division of Internal Medicine, Lausanne University Hospital, CHUV, Lausanne, Switzerland.

Early menopause affects up to 10% of the female population, reducing reproductive lifespan and thus influencing fertility. The genetic aetiology of early menopause is largely unknown in the majority of cases. We have undertaken a meta-analysis of genome-wide association studies in 3493 early menopause cases and 13598 controls from 10 independent studies. No novel genetic variants were discovered, but 17 variants previously associated with normal menopause age as a quantitative trait were associated with early menopause and POI (Primary Ovarian Insufficiency). Thus early menopause has a genetic aetiology which overlaps normal menopause variation and is at least partly caused by the additive effect of the same polygenic variants. The combined effect of the 17 variants on the risk of early menopause was greater than the best validated non-genetic risk factor, smoking.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2166W

Genome wide association study of nocturnal awakening in asthma. *D. Chhabra[1], A. Dahlin[1], Q. Duan[1], B. Himes[1,2], S.T. Weiss[1,2], K.G. Tantisira[1,3].* 1) The Channing Laboratory, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School., Boston, MA; 2) Partners Center for Personalized Genetic Medicine, Partners Health Care, Boston, MA; 3) Pulmonary Division, Brigham and Women's Hospital, Harvard Medical School., Boston, MA.

Introduction: Nocturnal awakening continues to be a major morbidity affecting quality of life in asthmatic subjects and has been shown to affect school attendance and performance. Genetic variation has been previously associated with nocturnal asthma in candidate gene studies. However, despite numerous reported GWAS in asthmatic subjects, none have focused on nocturnal awakening. We believe that identifying the genetic factors underlying severity of symptoms in asthma patients will improve our ability to personalize treatment and alter management protocols in this population.We hypothesized that a genome wide association study(GWAS) of nocturnal symptoms in asthmatic symptoms will identify novel genetic markers. Methods:For the initial analysis, we used the CAMP (Childhood Asthma Management Program) dataset, which enrolled 1,041 children, initially aged 5–12 years, with mild to moderate persistent asthma. This cohort has information on nocturnal awakening available in the form of interview questionnaires performed at baseline and diary card data. Based on 3 questions from baseline data: night time awakening in the last 6 months, night time awakening in the last month and night time awakening in the last week, we identified 223 cases with at least once weekly nocturnal awakenings and 351 controls. GWAS analysis was then carried out on genotype data obtained on Illumina Human Hap550v3 Bead Chip and Illumina 610 quad chip. Statistical analysis was performed in PLINK using age, sex and height as covariates and was limited to Caucasian subjects. The top 30 SNPs were then carried forward for replication in the Leukotriene Modifier or Corticosteroid or Corticosteroid-Salmeterol Trial (LOCCS). This dataset has nocturnal awakening information in the form of symptom severity in the last week. Results:The SNP (rs13430119) was noted to be the most significant at a p value of 2.09x10-7. This SNP lies in the intronic region of the EML6 gene (Echinoderm microtubule associated protein like 6). Of the top 30 SNPs carried forward for replication rs509237 was found to be significant at a p value of 0.002. Conclusion: Our study has identified novel genetic variants associated with nocturnal awakening in asthma subjects and replication in other populations is currently ongoing. Funding: R01 NR013391,U01 HL065899 and T32 HL007427.

2167T

Preliminary Meta-Analysis of Genome-wide Association Studies of Lung Function in Children. *M. Kowgier[1,2], M. Bustamante[3,4,5,6], J. Sunyer[3,4,6], J. Heinrich[11], C. Flexeder[11], G. Koppelman[10], M. Kerkhof[10], E.S. Schultz[7], E. Melen[7], S. Birch[8], R. Granell[9], E. Kreiner-Møller[8], H. Bisgaard[8], J. Henderson[9], L.J. Palmer[1,2].* 1) Ontario Institute for Cancer Research, Toronto, ON; 2) Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, ON; 3) Center for Research in Environmental Epidemiology (CREAL) - Hospital del Mar Research Institute (IMIM), Barcelona, Catalonia, Spain; 4) CIBER Epidemiologia y Salud Pública (CIBERESP), Spain; 5) Genes and Disease Program, Center for Genomic Regulation (CRG) and UPF, Barcelona, Catalonia, Spain; 6) Pompeu Fabra University (UPF), Barcelona, Catalonia, Spain; 7) Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden; 8) Copenhagen Prospective Studies on Asthma in Childhood, Health Sciences, University of Copenhagen, Copenhagen University Hospital, Gentofte, Denmark; 9) School of Social and Community Medicine, University of Bristol, UK; 10) Department of Pediatric Pulmonology and Pediatric Allergology, University of Groningen; 11) Institute of Epidemiology I, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany.

Objective/Purpose: Lung function measures, such as forced expiratory volume in one second (FEV1) and forced vital capacity (FVC), are complex traits that are important factors associated with respiratory and other health outcomes (such as asthma). To date, however, there have been no published genome-wide association (GWA) study results for lung function in children. We performed a preliminary meta-analysis of 4 GWAS using two lung function measures: FEV1 and the ratio of FEV1 to FVC. Methods: SNP associations for both lung function measures were computed for ~2.5 million genotyped and imputed single nucleotide polymorphisms (SNPs) in 4 birth cohort studies: ALSPAC, COPSAC, BAMSE and PIAMA. The combined analysis included up to 5,813 children, and stratified analysis in boys and girls included up to 2,838 and 2,788 children, respectively. Each cohort estimated residuals by regressing lung function measures on height, age, sex, passive smoking (if available), and ancestry principal components. The residuals were then used as the primary quantitative trait in the GWA. Regression coefficients for SNP effects were then meta-analyzed across cohorts using the sample-size weighting method, after applying genomic control. Results: Among the 4 cohorts, the age of children ranged from 5 to 10 years. Although no SNPs showed genome-wide significant associations, in the combined analysis of the ratio of FEV1 to FVC a SNP located just upstream of ADAM metallopeptidase with thrombospondin type 1 motif, 9 (ADAMTS9) on chromosome 3 showed near genome-wide level significance with p=1.1x10-7. Three other SNPs in this gene region had p < 1x10-5. Several other SNPs indicated suggestive associations. In girls, for the ratio of FEV1 to FVC, a SNP located in the intronic region of gene KCNMA1 on chromosome 10 showed suggestive association with p=8.96x10-7. Three other SNPs in this region had p < 1x10-5. Conclusions: A novel locus associated with pediatric lung function was identified. This locus belongs to the ADAM gene family whose other members have been previously implicated in the genetics of pulmonary function in adults. This locus, and others, will be further investigated in a larger meta-analysis which will include the addition of two other birth cohorts. The identification of novel genetic determinants of pediatric lung function has the potential to improve our understanding of the biological basis of normal physiology as well as the pathophysiology of asthma.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2168F

A genome-wide meta-analysis identifies common variants in *LOC201229* and *NOS2A* associated with fractional exhaled nitric oxide in childhood. *R.J.P. van der Valk[1], L. Duijts[2], M.T. Salam[3], J. Sunyer[4], E. Melen[5], G. Koppelman[6], A. Custovich[7], J. Heinrich[8], H. Bisgaard[9], A.J. Henderson[2], V.W.V. Jaddoe[1], J.C. de Jongste[1] on behalf of GABRIEL and the EArly Genetics and Lifecourse Epidemiology Consortia.* 1) The Generation R Study Group, Erasmus University Medical Center - Sophia Children's Hospital, Rotterdam, the Netherlands; 2) Department of Community Based Medicine, University of Bristol, Bristol, UK; 3) Department of Preventive Medicine, University of Southern California, Los Angeles, California, United States; 4) Center for Research in Environmental Epidemiology, Pompeu Fabra University, Barcelona, Spain; 5) Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden; 6) Department of Pediatric Pulmonology and Pediatric Allergology, Beatrix Children's Hospital, University Medical Center Groningen, GRIAC Research Institute, Groningen, the Netherlands; 7) Respiratory Research Group, University of Manchester, Manchester, UK; 8) Institute of Epidemiology I, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany; 9) Copenhagen Prospective Studies on Asthma in Childhood (COPSAC), Health Sciences, University of Copenhagen and Copenhagen University Hospital, Gentofte, Denmark.

Single nucleotide polymorphisms (SNPs) have been associated with childhood asthma. Limited progress has been made in understanding the genetics of specific endotypes of asthma. Fractional exhaled nitric oxide (FeNO) is a non-invasive biomarker of the degree of eosinophilic airway inflammation and is associated with childhood asthma symptoms and exacerbations, doctor diagnosis of asthma and atopy. The genetic factors associated with susceptibility to eosinophilic airway inflammation remain largely unknown. To identify common genetic variants associated with FeNO in childhood, we meta-analyzed 14 genome-wide association (GWA) samples from birth cohorts, including 8,062 subjects. Results showed genome-wide significant signals located at chromosome 17q11.2-q12, in novel locus *LOC201229* and in the candidate locus *NOS2A* (P = 1.28E-09, P = 4.38E-09, respectively). A conditional analysis showed that the signals mapping *LOC201299* and *NOS2A* are independent (P = 1.63E-09, P = 6.93E-08, respectively). In addition we found a further 13 suggestive signals associated with FeNO (P <= 1.00E-05), including an interesting signal at chromosome 4q21 near *SLC4A4* (P = 8.93E-07), best known to be a modifier gene of meconium ileus in cystic fibrosis. We are currently pursuing replication of all fifteen suggestive associated loci in two independent large cohorts with GWA and FeNO data in adults available. The present findings shed light on a novel independent signal in *LOC201229* next to the neighboring candidate *NOS2A* gene, and represents the first multi-site GWA meta-analysis of SNPs and FeNO in children.

2169W

Phenotypic dissection of bone mineral density facilitates the identification of skeletal site specificity on the genetic regulation of bone. *J.P. Kemp[1,2], C. Medina-Gomez C[3,4,5], K. Estrada[3], D.H.M. Heppe[5], M.C. Zillikens[3], N.J. Timpson[1,2], B. St Pourcain[1,2], S.M. Ring[1,2], A. Hofman[3], V.W.V. Jaddoe[5], G. Davey Smith[1,2], A.G. Uitterlinden[3,4,5], J.H. Tobias[6], F. Rivadeneira[3,4,5], D.M. Evans[1,2].* 1) MRC Centre for Causal Analyses in Translational Epidemiology, University of Bristol, Bristol, United Kingdom; 2) School of Social and Community Medicine, University of Bristol, Bristol, United Kingdom; 3) Department of Internal Medicine, Erasmus Medical Center, Rotterdam, The Netherlands; 4) Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands; 5) The Generation R Study Group, Erasmus Medical Center, Rotterdam, The Netherlands; 6) School of Clinical Science at North Bristol, University of Bristol, Bristol, UK.

Heritability of bone mineral density (BMD) varies at skeletal sites, ranging from 30% in the extremities to 90% in the skull. These differences may be attributed to varying environmental (i.e. mechanical loading and muscular activity) and genetic (i.e. bone mechanosensing, accrual, modelling and remodelling) components. The skull is least prone to loading and muscular activity, whereas arms experience moderate loading and increased activity. Exposure to both is highest in the legs. To identify genetic variants associated with site specificity and bone mass attainment, we performed a genome-wide association meta-analysis using DXA BMD measures at the skull (S), arms (A), legs (L), and total-body less head (TB). The study (n=9395) combined data from the Avon Longitudinal Study of Parents and their Children (n=5299, mean age=9.9, SD=0.32 years) and the Generation R study (n=4096, mean age=6.2, SD=0.5 years). Association analyses were performed using MACH2QTL, fitting a linear regression model based on expected allelic dosage model for SNPs, whilst adjusting for sex, age, weight or height and population stratification. After controlling for genomic inflation, we combined association data for ~2.5 million imputed SNPs into an inverse variance fixed-effects additive model meta-analysis, using METAL. Variants in 7 loci reached genome wide significance with TB-BMD and several displayed varying degrees of site specificity. SNPs mapping to *WNT4* and *WNT16* were associated with all BMD sites, whereas variants proximal to *TNFRSF11A*, *TNFRSF11B*, *LIN7C*, *LRP5* and *EYA4* were specific to S-BMD. *RSPO3* was associated with unloaded sites (A and S), while *TNFSF11* was associated with A and TB. With the exception of *GALNT3* (solely associated with TB-BMD), the remaining loci (*FUBP3*, *KLHDC5*, *PTHLH* and *RIN3*) were associated with L-BMD. Aside from a variant upstream of the Eyes Absent 4 (*EYA4*) locus (MAF=0.35; B=-0.1267; P=8.3x10-17), all other loci are known to influence BMD in elderly adults. In conclusion: BMD measures of axial and appendicular skeletal sites appear to aid genetic investigations of bone accrual, allowing for the dissection between effects on mechanosensing, muscle bone interactions and mass attainment. Furthermore, loss of function mutations in *EYA4* are associated with sensorineural hearing loss; which has previously been identified in sclerosteosis, where skull base overgrowth occurs in the context of high bone mass, leading to cranial nerve compression.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2170T

A multiethnic genome-wide association study of HIV-1 viral load among injection drug users. *D.B. Hancock[1], J.L. Levy[2], G.P. Page[3], S.P. Novak[1], C. Glasheen[1], N.C. Gaddis[2], N.L. Saccone[4], J.P. Rice[5], Q. Wang[6], M. Moreau[6], K. Doheny[7], J. Romm[7], A.I. Brooks[6], L.J. Bierut[5], A.H. Kral[8], E.O. Johnson[1].* 1) Behavioral Health Epidemiology, Research Triangle Institute International, Research Triangle Park, NC; 2) Research Computing Division, Research Triangle Institute International, Research Triangle Park, NC; 3) Genomics, Statistical Genetics, and Environmental Research Program, Research Triangle Institute International, Atlanta, GA; 4) Department of Genetics, Washington University in St. Louis, MO; 5) Department of Psychiatry, Washington University in St. Louis, MO; 6) Rutgers University Cell and DNA Repository (RUCDR), Riscataway, NJ; 7) Center for Inherited Disease Research (CIDR), Johns Hopkins University; 8) Urban Health Program, Research Triangle Institute International, San Francisco, CA.

Among those infected with HIV-1, there is considerable variation in the amount of circulating virus during the asymptomatic phase after acute infection and prior to development of AIDS. Genome-wide association (GWA) studies have discovered genetic risk factors of HIV-1 viral load, most notably single nucleotide polymorphisms (SNPs) in the human leukocyte antigen (HLA) region on chromosome 6. However, much of the population variability in disease progression remains unexplained. The previous GWA studies have generally not focused on viral load among active injection drug users, who may have additional biological challenges controlling the virus due to their drug use. To identify novel genetic risk factors of HIV-1 viral load, we conducted GWA studies in African American and European American HIV-1-positive injection drug users from the Urban Health Study. DNA was extracted from stored serum at RUCDR and treated with Illumina FFPE restoration prior to genotyping at CIDR on the Illumina Omni1-Quad array. Following quality control procedures for genotyped SNPs and participants, the analysis data set included ~700,000 autosomal SNPs in 569 African Americans and 314 European Americans. In the separate ethnic groups, additive SNP genotypes were tested for association with log-transformed viral load using linear regression models adjusted for age, gender, behavioral risk class (based on a latent class analysis of several important behavioral and demographic risk factors), recruitment pre- or post-availability of antiretroviral therapy, and 10 principal component eigenvalues for population stratification. The ethnic-specific GWA results were then combined in a fixed-effects meta-analysis. The lowest meta-analysis P-value ($P=5x10^{-7}$) was observed for an intergenic SNP located between two genes on chromosome 3p14 not previously associated with HIV-1 viral load. However, one of the novel genes has been directly linked to HIV protease inhibitor treatment responsiveness, providing further evidence to support its role in regulation of viral load. Several SNPs on chromosome 6 in the previously implicated HLA region had nominally significant meta-analysis P-values (lowest $P=4x10^{-5}$). Further analyses based on 1000 Genomes-imputed SNPs and genotyping of Hispanic American samples are underway to characterize these and other loci that may offer insights into the control of HIV-1 viral load and clinical progression towards AIDS.

2171F

Genetic variants near the *NTRK2* gene are associated with weight at birth in twins. *S.J. Metrustry, T. Spector, A.M. Valdes.* Department of Twin Research and Genetic Epidemiology, King's College London, St Thomas' Hospital, Westminster Bridge Road, London, United Kingdom. SE1 7EH.

Background. The Developmental origins of health and disease hypothesis suggests that pre and peri natal factors have a large impact on adult health. Birth weight has been shown to be associated with an increased risk of developing diseases such as type 2 diabetes, cardiovascular disease, obesity, osteoporosis, thus potentially, osteoarthritis later in life. Studies on the genetics of birth weight, to date, have shown that birth weight has an important genetic component but association studies have, thus far, focused only on singleton births to avoid confounding due to (1) the much lower weight at birth in twins and (2) the large difference in birth weight between twins in a pair which is on average 10% of weight at birth. In this study we investigated the genetic component of birth weight in a large twin cohort from the UK (TwinsUK). **Methods**. Genome-wide association study (GWAS) in 4,775 study participants with birth weight data available. **Findings.** We identified 23 genome wide significant hits in chromosome 9 (top SNP has beta=-0.106, SE=0.0186 and p-value 1.01E-08) close to *NTRK2*, associated with birth weight. *NTRK2*(neurotrophic tyrosine kinase, receptor, type 2) is a membrane-bound receptor. This gene encodes a member of the neurotrophic tyrosine receptor kinase (NTRK) family. This kinase, upon neurotrophin binding, phosphorylates itself and members of the MAPK pathway. Signalling through this kinase leads to cell differentiation and mutations in this gene have been associated with obesity and mood disorders. We ran further association analyses and found that the four top SNPs associated in both the high-weight twin and the lower birth twin although effect sizes were slightly larger in higher birth weight twins. Results were also tested in a cohort of 3000 singleton individuals from the UK. **Conclusions.** Our findings indicate that genetic variation plays an important role in determining birth weight in twins and provides insights into the pathways likely to be involved in affecting birth weight in this group. The *NTRK2* gene has previously been associated with obesity and so these findings reinforce previous work stating that birth weight can be linked to adulthood age-related diseases with an obesity component, such as osteoarthritis.

2172W

A Systematic Meta-analysis of Associations of Genetic Variants Influencing Protein Phosphorylation with Bone Mineral Density Phenotypes. *T. Niu[1], L. Zhang[1,2], F.-Y. Deng[1], J. Li[1], Y.-F. Pei[1,2], Y. Liu[1], H. Shen[1], Y.-Z. Liu[1], H.-W. Deng[1*,2,3].* 1) Center for Bioinformatics and Genomics, Department of Biostatistics and Bioinformatics, Tulane University School of Public Health and Tropical Medicine, New Orleans, LA. 70112; 2) Center of System Biomedical Sciences, University of Shanghai for Science and Technology, Shanghai 200093, China; 3) College of Life Sciences, Hunan Normal University, Changsha, Hunan 410081, China.

Aim: Protein phosphorylation is a ubiquitous regulatory mechanism and an important type of post-translational modification. Our objective is to identify phosphorylation-related polymorphisms in coding sequences that affect bone mineral density (BMD) phenotypes. **Methods:** We performed a systematic, three-stage meta-analysis of genome-wide association (GWA) studies for lumbar spine, hip, and femoral neck BMD phenotypes. After analyzing 64,035 phosphorylation-related single nucleotide polymorphisms (SNPs) in 11,125 individuals of various ancestries (Stage I), top 60 phosphorylation-related SNPs were selected for in silico replication in 3 GWA studies in Australian, Dutch, and Korean populations respectively (Stage II). Five SNPs, i.e., **rs1827293** (Tyr114Cys; *NBPF3*, 1p36.12), **rs6831280** (Ala361Thr; *IDUA*, 4p16.3), **rs3755955** (Arg105Gln; *IDUA*, 4p16.3), **rs2707466** (Thr253/263Ile; *WNT16*, 4p16.3), and **rs56358776** (Arg1561Gln; *ESPL1*, 12q13.13) were further examined in *de novo* genotyping replication in a cohort comprising a sample of 3,923 unrelated European individuals (CEU3) and a sample of 2,822 unrelated Chinese Han individuals (CHI2) (Stage III). **Results:** In combined meta-analysis across all stages, at $P < 10^{-5}$ level, **rs6831280**, **rs3755955**, and rs2707466 were significantly associated with femoral neck BMD ($P = 3.42x10^{-10}$, $5.18x10^{-10}$, and $5.61x10^{-7}$, respectively) and hip BMD ($P = 1.32x10^{-6}$, $2.18x10^{-6}$, and $1.80x10^{-9}$, respectively) phenotypes, and **rs56358776** was significantly associated with spine BMD ($P = 2.21x10^{-6}$). **Conclusion:** Phosphorylation-related non-synonymous SNPs located in *IDUA*, *WNT16* and *ESPL1* genes could have critical roles in BMD variation, and these results should be corroborated in future studies of osteoporosis. *Corresponding Author: Dr. Hong-Wen Deng (Tulane University).

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2173T

Chromosome X revisited - Variants in Xq21.1 associate with adult stature in a meta-analysis of 14,700 Finns. *T. Tukiainen[1], J. Kettunen[1,2], A.-P. Sarin[1,2], J.G. Eriksson[3,4,5,6,7], A. Jula[8], V. Salomaa[3], O.T. Raitakari[9,10], M.-R. Järvelin[11,12], S. Ripatti[1,2,13].* 1) Institute for Molecular Medicine Finland FIMM, University of Helsinki, Finland; 2) Unit of Public Health Genomics, National Institute for Health and Welfare, Helsinki, Finland; 3) Department of Chronic Disease Prevention, National Institute for Health and Welfare, Finland; 4) Department of General Practice and Primary Healthcare, University of Helsinki, Finland; 5) Unit of General Practice, Helsinki University Central Hospital, Finland; 6) Folkhälsan Research Center, Helsinki, Finland; 7) Vaasa Central Hospital, Vaasa, Finland; 8) Population Studies Unit, Department of Chronic Disease Prevention, National Institute for Health and Welfare, Turku, Finland; 9) Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Finland; 10) Department of Clinical Physiology, Turku University Hospital, Finland; 11) Department of Epidemiology and Biostatistics, Faculty of Medicine, Imperial College London, United Kingdom; 12) Institute of Health Sciences, Biocenter Oulu, University of Oulu, Finland; 13) Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK.

Genome-wide association studies (GWAS) provide a powerful tool to assess genetic associations between common marker alleles and complex traits in large numbers of individuals. Typically these studies have focused on testing the markers in the 22 autosomal chromosomes while the X-chromosome has been omitted from the analyses. The chromosome X, however, constitutes approximately 5% of genomic DNA encoding for more than 1000 genes, and thus also likely contains genetic variation contributing to common traits and disorders.

We set to test associations between 560,000 genotyped and imputed SNP markers and eight anthropometric (BMI, stature, WHR) and biochemical (CRP, HDL, LDL, TC, TG) traits in 14,710 individuals (7468 males, 7242 females) from five Finnish cohorts.

A region in chromosome Xq21.1 was associated with adult stature (meta-analysis p-value = 3.32×10^{-10}). The lead SNP in the locus explained up to 0.55% of the variance in height in 31-year-old women corresponding to 1.09 cm difference between minor and major allele homozygotes. The associated lead variant (MAF = 0.31) is located upstream of *ITM2A*, a gene encoding for a membrane protein that plays a role in osteo- and chondrogenic differentiation. As this is among the first studies using the X chromosome reference haplotypes from the 1000 Genomes project, we are currently validating the imputation with genotyping methods.

The findings pinpoint the value of including chromosome X in the GWAS of complex traits to identify further relevant gene regions that also account for some of the missing heritability. The study illustrates that the 1000 Genomes reference haplotypes allow for high-resolution investigations of the genetic variants in chromosome X even with a relative modest sample sizes compared to the current-day GWAS meta-analyses. Our finding demonstrates that the same analysis strategy is also likely to be useful in the meta-analyses of the large consortia with complex traits.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2174F

A genome-wide association study identifies two susceptibility loci for Crohn's disease in a Japanese population. *K. Yamazaki[1], J. Umeno[1,2], A. Takahashi[3], A. Hirano[1,2], T. Johnson[4], N. Kumasaka[3], T. Morizono[4], N. Hosono[1], T. Kawaguchi[5], M. Takazoe[5], T. Yamada[6], Y. Suzuki[6], H. Tanaka[7], S. Motoya[7,8], M. Hosokawa[8], Y. Arimura[8], Y. Shinomura[8], T. Matsui[9], T. Matsumoto[2], M. Iida[2], T. Tsunoda[4], Y. Nakamura[10], N. Kamatani[3], M. Kubo[1].* 1) Center for Genomic Medicine, RIKEN, Yokohama, Japan; 2) Department of Medicine and Clinical Science, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan; 3) Laboratory for Statistical Analysis, Center for Genomic Medicine, RIKEN, Tokyo, Japan; 4) Laboratory for Medical Informatics, Center for Genomic Medicine, RIKEN, Yokohama, Japan; 5) Department of Medicine, Division of Gastroenterology, Social Insurance Chuo General Hospital, Tokyo, Japan; 6) Department of Internal Medicine, Faculty of Medicine, Toho University, Chiba; 7) Department of Gastroenterology, Sapporo Kosei Hospital, Sapporo, Japan; 8) First Department of Internal Medicine, Sapporo Medical University School of Medicine, Sapporo, Japan; 9) Department of Gastroenterology, Fukuoka University Chikushi Hospital, Fukuoka, Japan; 10) Laboratory of Molecular Medicine, Institute of Medical Science, The University of Tokyo, Tokyo, Japan.

Crohn's disease is (CD) an inflammatory bowel disease induced by multiple genetic and environmental factors. Genome-wide association studies (GWAS) have identified many genetic factors for CD in the European population, but information in other ethnic groups is scarce. Though there are no significant differences in the clinical characteristics and natural history, previous studies have showed apparent ethnic differences in the CD susceptibility loci between European and East-Asian populations such as *NOD2*, *IL23R*, and *ATG16L1*. To clarify the genetic factors for CD in Japanese population, we performed a GWAS comprising 372 affected individuals and 3,389 controls. After excluding SNPs located at previously reported loci, we carried out genotyping for the replication study in an independent panel of 1,151 cases and 15,800 controls. After combining the data from the GWAS and the replication study, we confirmed the association in the MHC (rs7765379, P = 2.11 x 10-59), *TNFSF15* (rs6478106, P = 3.87 x 10^{-45}) and *STAT3* (rs9891119, P = 2.24 x 10^{-14}) regions. In addition, we identified two novel susceptibility loci on chromosome 4p14 (rs1487630, P = 2.40 x 10^{-11}, odds ratio = 1.33) and in the *SLC25A15-ELF1-WBP4* region on 13q14 (rs7329174 in *ELF1*, P = 5.12 x 10^{-9}, odds ratio = 1.27). Our analysis suggested that the associated SNPs on 13q14 had eQTL activity. In conclusion, we identified two new susceptibility loci for CD at 4p14 (spanning from 38.32 to 38.37 MB) and 13q14 (spanning from 41.416 to 41.59 MB) in a Japanese population. Further functional studies of these loci will provide a better understanding of the pathogenesis of CD.

2175W

Large scale GWAS-meta analysis identifies novel variants associated with mean leukocyte telomere length. *V. Codd[1], C.P. Nelson[1], E. Albrecht[2], P. van der Harst[3], M. Mangino[4], J. Deelen[5], J.L. Buxton[6], J.J. Hottenga[7], K. Fischer[8], I. Surakka[9], L. Broer[10], D.R. Nyholt[11], P. Salo[12], S. Hagg[13], P.J. Talmud[14], N.L Pedersen[13], M. Perola[12], J. Kaprio[15], N.G. Martin[11], C.M. van Duijn[10], I. Hovatta[16], C. Gieger[2], A. Metspalu[8], D.I. Boomsma[7], M.R. Jarvelin[17], W.H. van Gilst[3], P.E. Slagboom[5], T.D. Spector[4], N.J. Samani[1], ENGAGE consortium.* 1) Cardiovascular Sciences, University of Leicester, Leicester, United Kingdom; 2) Institute of Genetic Epidemiology, Helmholtz Zentrum München, Germany; 3) Department of Cardiology, University of Groningen, University Medical Center Groningen, The Netherlands; 4) Department of Twin Research and Genetic Epidemiology, King's College London, London, UK; 5) Section of Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands; 6) Division of Diabetes, Endocrinology and Metabolism, Imperial College, UK; 7) Netherlands Twin Register, Dept Biological Psychology, VU Univ Amsterdam, The Netherlands; 8) Estonian Genome Center, University of Tartu, Tartu, Estonia; 9) Institute for Molecular Medicine Finland, University of Helsinki, Helsinki, Finland; 10) Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands; 11) Queensland Institute of Medical Research, Brisbane, Queensland, Australia; 12) Public Health Genomics Unit, Department of Chronic Disease Prevention, National Institute for Health and Welfare, Helsinki, Finland; 13) Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Sweden; 14) Institute of Cardiovascular Science, Univerisity College London, London, UK; 15) University of Helsinki, Hjelt Institute, Dept of Public Health, Helsinki, Finland; 16) Research Programs Unit, Molecular Neurology, Biomedicum Helsinki, University of Helsinki, Finland; 17) Department of Epidemiology and Biostatistics, Imperial College London, UK.

Telomeres are the protein bound DNA repeat structures at the ends of chromosomes which have an important role in maintaining genomic stability and in regulating cellular lifespan. Telomere length (TL) is an important determinant of telomere function. Mean leukocyte TL shows considerable inter-individual variability and family-based studies have shown that it is a highly heritable trait with estimates varying between 44-80%. Inter-individual variation in mean leukocyte TL has been associated with cancer and several age-associated diseases. To date, in genome-wide association studies (GWAS), only common variants at one locus on Chr3q26 (*TERC*), explaining <1% of the variance in TL, have shown an association with mean TL in more than one study. Variants at two other loci, Chr10q24.33 (*OBFC1*) and Chr2q21 (*CXCR4*) have shown association at genome-wide significance in a single study. To identify further genetic determinants of TL we conducted a large scale GWAS meta-analysis of 37,684 individuals from 15 European cohorts, followed by replication of selected variants in a further ~11,000 individuals from 6 cohorts. We identified seven loci, including five novel loci, harbouring common variants affecting mean leukocyte telomere length at a genome-wide significant level ($P<5x10^{-8}$). Five of the loci include genes (*TERC*, *TERT*, *NAF1*, *OBCF1*, *RTEL1*) known to be involved in telomere biology, including 3 (*TERC*, *TERT*, *NAF1*) that are involved in the formation and activity of telomerase. The individual impact of the variants on telomere length ranged from ~ 57-117 base pairs per allele. Interrogation of genetic association databases showed that the lead SNPs at two of the loci (*TERC* and *TERT*) are also associated with a several cancers as well as other diseases, including idiopathic pulmonary fibrosis. We report five novel and confirm two previously reported variants associated with TL in humans. Our findings provide a framework for investigating the causal role of telomere length in both telomere- and age-related diseases.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

† Trainee Award Finalist

2176T

Suggestive loci for osteoporosis: a variance components linkage analysis of extended pedigrees. *S.C. Nguyen, N.D. Nguyen, J.C. Center, J.A. Eisman, T.V. Nguyen.*Osteoporosis and Bone Biology, Garvan Institute of Medical Research, Darlinghurst, NSW, Australia.

Genome-wide association studies to date have been successful at identifying many genes and common variants associated with diseases such as osteoporosis, however they face limitations such as numbers needed to detect, missing heritability, and the prevalance of rare variants. Linkage analysis using large pedigrees represents an attractive alternative as they avoid these issues, potentially identifying genes not detectable in association studies. The primary aim of this study was to identify genetic loci that are linked to bone mineral density (BMD) and quantitative ultrasound parameters (QUS), both measures of osteoporosis.

The Dubbo Osteoporosis Genetics Study is a large multi-generational family study involving 141 pedigrees. Individuals are recruited through probands with high BMD (Z-score greater than +1.28, corresponding to the top 10% of an age- and sex-matched population). BMD (g/cm^2) was measured at the femoral neck, lumbar spine and whole body by DXA (GE-LUNAR), and QUS measurements were obtained using a Sunlight Omnisense. Subjects were genotyped using 530 microsatellite markers. Following adjustment for age, sex and other covariates, variance components linkage analyses were performed using SOLAR.

In this study we analysed 1007 individuals (636 women and 371 men) from 141 pedigrees, with a mean age of 58 ± 19 (range 18-98). We found suggestive linkage for BMD at the whole body on chromosome 6 (LOD score 2.02) and for QUS at the tibia on chromosome 18 (LOD score 2.77), though not for BMD at the femoral neck (highest LOD score 1.50 on chromosome 6) or lumbar spine (highest LOD score 1.38 on chromosome 12). We performed further analysis following stratification in order to identify any age-specific linkage, and found suggestive signals for BMD at both the femoral neck (LOD score 2.88 on chromosome 18) and lumbar spine (LOD score 1.82 on chromosome 18) for individuals aged 50-80 years.

These data suggest the presence of multiple loci regulating bone mineral density and quantitative ultrasound measurements, both in general and age-specific manners. The identification of specific markers or genes may lead to better prediction of disease and fracture risk, and greater individualised osteoporosis treatments.

2177F

Linkage analysis for identification of rare, high-penetrance celiac disease risk variants in the Finnish and Hungarian populations. *E. Einarsdottir[1], K. Kurppa[2], K. Kaukinen[3], I. Korponay-Szabo[4], K. Mustalahti[2], M. Balogh[5], M. Mäki[2], P. Saavalainen[1].* 1) Research Program Unit, Molecular Medicine, University of Helsinki, Helsinki, Finland and Department of Medical Genetics, Haartman Institute, University of Helsinki, Helsinki, Finland; 2) Paediatric Research Centre, School of Medicine, University of Tampere, Tampere, Finland and Department of Paediatrics, Tampere University Hospital, University of Tampere, Tampere, Finland; 3) Department of Gastroenterology and Alimentary Tract Surgery, Tampere University Hospital, Tampere, Finland and School of Medicine, University of Tampere, Tampere, Finland; 4) Heim Pal Children's Hospital, Budapest, Hungary and Research Centre for Molecular Medicine, University of Debrecen, Debrecen, Hungary; 5) Markusovsky Teaching Hospital, Szombathely, Hungary.

Background: Celiac disease is an immune disease of the gut, triggered by gluten ingestion in genetically susceptible individuals. GWAS studies have identified a number of risk variants, but a large proportion of the risk is still unexplained and may be due to rare variants of high risk-effect. **Methods:** Our approach to identifying rare celiac disease variants of high risk-effect involved selecting 31 families with Mendelian-like inheritance pattern (minimum of five affected individuals), and performing whole-genome allele-sharing linkage analysis. The custom Illumina Immunochip was used for the analysis, allowing for direct fine-mapping of previously known celiac disease /autoimmune disease loci. **Results:** We obtained non-parametric LOD-scores >2 on chromosomes 6 (the HLA locus), 11, 15, 19 and 21 in the Finnish families (N=28) and in the combined Finnish and Hungarian families (N=31). In the Hungarian families alone (N=3), we had LODs over >2 on chromosomes 2, 9, 17, 19, and 20. A subset of these loci overlap with previously described linkage peaks/association signals. Another subset of loci constitutes novel putative celiac disease loci, some of which have been implicated in other immune disease previously. **Conclusions:** Our preliminary results indicate that using the Immunochip for linkage analysis on large pedigrees can be a viable approach to identifying high-risk variants in complex diseases such as celiac disease. The further follow-up of the most promising loci with whole-locus or exome sequencing will help identify putative high-risk disease variants involved in celiac disease.

2178W

Whole Genome Linkage Analysis to Identify Genes for Childhood Dental Caries. *N. Mukhopadhyay[1], M. Govil[1], Z. Zheng[2,3], E. Feingold[2,3], D.E. Weeks[2,3], J.R. Shaffer[2], X. Wang[1], R.J. Weyant[5,6], R. Crout[5,7], D.W. McNeill[8], M.L. Marazita[1,2,4].* 1) Department of Oral Biology, University of Pittsburgh, School of Dental Medicine, Pittsburgh, PA; 2) Department of Human Genetics, University of Pittsburgh, Graduate School of Public Health, Pittsburgh, PA; 3) Department of Biostatistics, University of Pittsburgh, Graduate School of Public Health, Pittsburgh, PA; 4) Clinical and Translational Science, School of Medicine, University of Pittsburgh, Pittsburgh, PA; 5) Center for Oral Health Research in Appalachia, University of Pittsburgh, Pittsburgh, PA and West Virginia University, Morgantown, WV; 6) Department of Dental Public Health and Information Management, School of Dental Medicine, University of Pittsburgh, Pittsburgh, PA; 7) Department of Periodontics, West Virginia University School of Dentistry, Morgantown, WV; 8) Department of Dental Practice and Rural Health, West Virginia University School of Dentistry, Morgantown, WV.

Childhood dental caries is one of the most common chronic diseases among children. Approximately 40% of children have a cavity or filling by 5 years of age. Both environmental (including social-economic) and genetic factors play a role in the observed prevalence of the disease, with an increased prevalence of childhood caries in recent years, particularly among children in the 2-11 years age group. Earlier studies have demonstrated the role of multiple genes in the development of caries, although few of the candidate genes proposed to date have been validated. To identify genetic regions with loci that may influence susceptibility to childhood caries, we performed a whole-genome linkage scan for 327 families from the Appalachia region of West Virginia and Pennsylvania. Family sizes ranged from 4-17 individuals. All individuals aged 1 to 43 years with one or more primary teeth having caries were designated as affected. Genotyping for the Illumina HumanHap 550 panel (550,000 evenly spaced SNPs) was conducted by the Center for Inherited Disease Research (CIDR). We carried out a multipoint genome-wide nonparametric linkage analysis as implemented in the software package Merlin. To allow for linkage disequilibrium (LD), the SNPs were clustered based on a 0.1 r2 threshold prior to conducting the linkage analysis. To improve detection of LD clusters among markers, we included an additional 526 genotyped three-member families, although they dropped out of the subsequent analysis. Our preliminary results show a peak LOD score of 3.04 in the 1q24-25 region. A suggestive peak with a LOD score of 2.6 was also obtained for the 3p23-24 region. While we continue to refine the complex childhood caries phenotype as we explore the role of genes implicated in the development of the disease, our linkage results support prior association findings (p-value > 10 E-05) for SNPs on chromosomes 1, 7, 8, 10, 11, 12, 15 and 17 by Shaffer et al. (JDR, 2011). This analysis was supported under NIDCR grants R00-DE018085, U01-DE018903, R01-DE 014899 and 1R03-DE021425.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2179T

simQTL: Software for simulation of a complex etiological model under a set of epidemiological model parameters. *J.D. Terwilliger[1, 2, 3, 4, 5], K.M. Weiss[6], J.H. Lee[7, 8, 9], B. Lambert[6].* 1) Genetics and Development, Columbia University, New York, NY; 2) Medical Genetics, New York State Psychiatric Institute, New York, NY; 3) Chronic Disease Prevention, Public Health Genomics Unit, National Institute for Health and Welfare, Helsinki, Finland; 4) Psychiatry, Columbia University, New York, NY; 5) Columbia Genome Center, Columbia University, New York, NY; 6) Anthropology, Pennsylvania State University, State College, PA; 7) Epidemiology, Columbia University, New York, NY; 8) Sergievsky Center, Columbia University, New York, NY; 9) Taub Institute, Columbia University, New York, NY.

Biotechnology has advanced at an absolutely amazing pace over the past two decades, fueled largely by the desire to understand the relationships between genetic variation and the complex common diseases which contribute the largest amount to human morbidity and mortality in Western societies. However, this accomplishment has not brought us significantly closer to the widely promised medical benefits of early diagnosis and treatments, largely because scientists grossly underestimated the complexity of the genotype-phenotype relationships underlying such diseases. We have developed a software package, simQTL, to simulate complex etiological models consistent with user-specified epidemiological and population genetic parameters. The present version of simQTL generates an etiological architecture model involving up to thousands of functional genetic variants, together with environmental exposures (both discrete and continuous, familially shared and unshared) from the infinite number of possible etiological models consistent with a large set of user-defined parameters describing the population distribution and familial correlations seen in a given trait, along with locus-specific effect size estimates from which a complex architecture involving potentially numerous functional and nonfunctional variants per gene is simulated. A population of individuals in families (if desired) is then simulated with a genome-spanning set of marker (and functional variant) data and precise chromosomal segregation patterns, along with environmental exposures and phenotypes (both quantitative and qualitative as requested). Summary statistical analysis is provided as well of both the linkage information (pointwise and multipoint) and the association information (i.e. detectance distributions) in the dataset requested by the user under the requested ascertainment criteria. A whole population is simulated, but a user-described dataset is ascertained and output therefrom. The current version of the software is available freely and runs under various flavors of UNIX with an easy to use GUI guiding the user through the numerous questions which must be answered to define the desired genetic epidemiological characteristics of the resulting etiological model from which the dataset is to be generated. simQTL is freely available from the authors.

2180F

A hybrid likelihood model for sequence-based disease association studies. *Y. Chen[1], H. Carter[1], J. Parla[2], M. Kramer[2], F.S. Goes[3], M. Pirooznia[3], P.P. Zandi[3], W.R. McCombie[2], J.B. Potash[4], R. Karchin[1].* 1) Biomedical Engineering, Johns Hopkins University, Baltimore, MD; 2) Stanley Institute for Cognitive Genomics, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY; 3) Psychiatry and Behavioral Sciences, Johns Hopkins School of Medicine, Baltimore, MD; 4) Psychiatry, University of Iowa, Iowa City, Iowa.

In the past few years, case-control studies of common diseases have shifted their focus from single genes to whole exomes. New sequencing technologies now routinely detect hundreds of thousands of sequence variants in a single study, many of which are rare or even novel. The limitation of classical single-marker association analysis for rare variants has been a challenge in such studies. A new generation of statistical methods for case-control association studies has been developed to meet this challenge. A common approach to association analysis of rare variants is the burden-style collapsing methods to combine rare variant data within individuals across or within genes. Here, we propose a new hybrid likelihood model that combines a burden test with a test of the positional distribution of variants. In extensive simulations and on empirical data from the Dallas Heart Study, the new model demonstrates consistently good power, in particular when applied to a gene set (e.g., multiple candidate genes with shared biological function or pathway), when rare variants cluster in key functional regions of a gene, and when protective variants are present. The model is potentially useful for detecting causal genes in whole-exome case-control studies with 1000-2000 subjects, when a suitable candidate gene set is known in advance. When applied to data from an ongoing sequencing study of bipolar disorder (191 cases, 107 controls), the model identifies six causal gene sets with nominal p-values < 0.05, of which one “neurological system process” (GO:005087) is significant after multiple testing correction.

2181W

Novel methods for combined linkage and association analyses of complex quantitative traits in very large families sheds light on genetics of musical aptitude. *Y. Huang[1], J. Oikkonen[2,3], L. Ukkola-Vuoti[2], P. Raijas[4], K. Karma[4], P. Onkamo[3], A. Thomas[5], I. Jarvela[2], V.J. Vieland[1,6].* 1) Battelle Ctr Math Med, Res Inst at Nationwide Children's, Columbus, OH; 2) Dept. of Medical Genetics, U. of Helsinki, Helsinki, Finland; 3) Dept. of Biological and Environmental Science, U. of Helsinki, Finland; 4) Sibelius Academy, DocMus Dept., Helsinki, Finland; 5) Dept. of Biomedical Informatics, U. of Utah, Salt Lake City, UT; 6) Dept. of Pediatrics and Statistics, Ohio State U., Columbus, OH.

A previous genome-wide microsatellite linkage scan for musical aptitude in 15 Finnish pedigrees showed evidence for a major locus at 4q22 assuming an additive model under the variance components framework of SOLAR, and targeted haplotype analyses indicated association with AVPR1A using HBAT, which required breaking up the pedigrees. In the current study the original 15 plus 85 new families were genotyped on the IlluminaHumanOMNIExpress 12 1.0v SNP chip. Using the software package Kelvin, which supports combined linkage (posterior probability of linkage, PPL) and linkage disequilibrium (posterior probability of linkage disequilibrium, PPLD) analysis, we were able to keep the very large families intact, conduct full multipoint linkage analyses while utilizing Kelvin's unique approach to integrating over the nuisance parameters of the trait model, and to run genome-wide association analyses on the full pedigrees conditioned on the linkage information. For the linkage portion, exact likelihood calculations were impossible due to inherent limitations of the Elston-Stewart algorithm on the number of markers and of the Lander-Green algorithm on pedigree size. We solved this problem by adapting MCMC sampling of meiosis indicators, as implemented in McLink, to work with Kelvin's direct (non-stochastic) numerical integration routines for handling the trait model, forming a unique “hybrid” approach using MCMC for the marker data and exact likelihood calculations for the trait data. Using two separate quantitative trait musical aptitude phenotyes (Seashore Pitch (SP); Karma Music Test (KMT)) and a combined phenotype (COMB), we obtained evidence of both common and unique loci: with SP peaks at 4p13 (PPL=85%), 16p12.3 (19%), 22q11.21 (36%); KMT peaks at 4q12 (13%), 4q21.21 (24%), and 16p13.11(23%); and COMB peaks at 4q12 (33%), 4q35.1 (12%), 18q22.2 (13%), and 22p16.1(20%). PPLD analyses are currently underway, together with fine-mapping in seven very large (> 30 individuals) individual pedigrees.

2182T

Linkage and Genome-wide Association Studies reveal distinct facets of the genetic architecture of Alopecia Areata. *L. Petukhova[1,2], A.M. Christiano[1,3].* 1) Department of Dermatology, Columbia University, New York, NY; 2) Department of Epidemiology, Columbia University, New York, NY; 3) Department of Genetics & Development, Columbia University, New York, NY.

Marker-based methods for genome-wide gene mapping are broadly categorized into two techniques: linkage and association (e.g. GWAS). Both methods utilize genome-wide genotype data to identify genomic regions with distributions of marker alleles that are consistent with the presence of a proximal disease gene. For linkage analysis, families are typed and analyzed to identify genomic regions that cosegregate with disease. Cosegregation will be observed when the disease allele is rare in the population and has strong effects on the phenotype. In contrast, GWAS compare allele frequencies for genetic markers across groups of unrelated cases and controls and is best suited for detecting disease alleles that are common in the population, and which consequently are not likely to have strong phenotypic effects. Thus, because different sets of assumptions underlie each of these two gene-mapping methods, it is expected that each will identify different sets of disease alleles. Therefore, the approaches are complementary to each other. Importantly, emerging empirical evidence, which is well-supported by theoretical arguments, suggests that the genetic architecture of common diseases contains niches for both rare and common alleles. We have previously published the first GWAS for alopecia areata, a prevalent autoimmune disease that targets the hair follicle causing disfiguring hair loss. Here we report for the first time linkage analysis in a cohort of families ascertained in the same population as the GWAS cases and typed with a commercial SNP genotyping array. Linkage analysis combined with pedigree-based association testing (PBAT) identified a set of loci which were not identified in our GWAS. Some of these newly identified loci map to biological processes implicated by the GWAS (e.g. VNN1), while others implicate processes known to be important for hair follicle biology (e.g. FDFT1, COL4A3, COL19A1), but were not identified in pathway analysis of significant GWAS loci. Our data are consistent with emerging evidence in other complex diseases which support the notion that complete enumeration of the genetic architecture may best be achieved by utilizing both linkage and association methods to identify the full spectrum of disease alleles.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2183F

Rare and common variants in extracellular matrix gene, FBN2, are associated with Mendelian and multifactorial macular degenerative disease. *R.R. Priya[1], X. Zhan[2], R.N. Fariss[3], C.F. Chakarova[4], K.E. Branham[5], M.M. Campos[4], J.S. Friedman[1], B. Brooks[1], H.K. Rajasimha[1], M.A. Morrison[6], S.G. Jacobson[7], M.L. Klein[8], E.Y. Chew[9], D. Stambolian[10], M.M. DeAngelis[6], S.S. Bhattacharya[4], J.R. Heckenlively[5], G.R. Abecasis[2], A. Swaroop[1].* 1) Neurobiology-Neurodegeneration & Repair Laboratory (NNRL), National Eye Institute, Bethseda, MD; 2) Center for Statistical Genetics, Department of Biostatistics, University of Michigan, Ann Arbor, MI, USA; 3) Biological Imaging Core Unit, National Eye Institute, Bethesda, MD, USA; 4) Department of Molecular Genetics, Institute of Ophthalmology, London, EC1 9EL, UK; 5) Department of Ophthalmology, University of Michigan, Ann Arbor, MI, USA; 6) Department of Ophthalmology and Visual Sciences, University of Utah, Salt Lake City, UT, USA; 7) Scheie Eye Institute, Department of Ophthalmology, University of Pennsylvania, Philadelphia, PA, USA; 8) Macular Degeneration Center, Casey Eye Institute, Oregon Health and Science University, Portland, OR, USA; 9) National Eye Institute, National Institutes of Health, Bethesda, MD, USA; 10) Departments of Ophthalmology and Genetics, University of Pennsylvania, Philadelphia, PA, USA.

Macular dystrophies constitute clinically and genetically heterogeneous group of disorders, characterized by progressive loss of visual acuity due to macular dysfunction. Age related macular degeneration (AMD) represents the most common late-onset multifactorial macular disease, associated with at least 19 susceptibility loci and several environmental factors. Stargardt disease, Best macular dystrophy, Sorsby's fundus dystrophy and Doyne honeycomb retinal dystrophy are among the early-onset Mendelian macular diseases that share clinical features with AMD. Using a combination of whole genome linkage analysis and exome sequencing, we have identified a novel heterozygous segregating mutation, p.Glu1144Lys in *FBN2* in a two-generation family with early form of dominant macular dystrophy. *FBN2* encodes a cysteine-rich glycoprotein, which is a principal component of extracellular matrix (ECM). This variant was absent from the 1000 Genomes Project, ESP5400 and dbSNP135 databases. Sequencing of patients with maculopathies and AMD identified additional rare variants. Immunohistological studies in monkey and human fetal eye localized the FBN2 protein in Bruch's membrane, choroid and sclera. Bruch's membrane is multilayered structure located between the choroid and retinal pigment epithelium, and extracellular deposits within or adjacent to Bruch's membrane are often clinically associated with AMD, making *FBN2* an interesting candidate. Genes encoding ECM components have been implicated in both early and late macular degenerative diseases, and recent genomewide association studies have identified variants at or near *TIMP3*, *COL8A1*, and *COL10A1*, conferring susceptibility to AMD patients. This prompted us to explore whether common variants in *FBN2* are associated with AMD. Our analysis revealed the association of a non-synonymous (Val965Ile) SNP, rs154001, with AMD in the analysis of 1988 cases and 2182 population matched controls (p value=1.03×10^{-4}; odds ratio= 0.85). Taken together our findings indicate the importance of *FBN2* in macular degenerative diseases and illustrate a unique scenario where both rare and common variants in a single gene contribute to the clinical phenotype.

2184W

Genetic Evidence for an Ethnic Diversity in the Susceptibility to Ménière's Disease. *J. Ohmen[1], C. White[4], J. Wang[1], L. Fisher[2], H. Zhang[5], M. Derebery[3], R. Friedman[1,3].* 1) Molec & Cellular Biol, House Ear Inst, Los Angeles, CA; 2) Department of Clinical Studies, House Ear Institute, Los Angeles CA, 90057; 3) House Clinic, Suite 111, 2100 W. Third St., Los Angeles, CA, 90057; 4) Bioinformatics and Systems Biology Graduate Program, Univ. of California at San Diego, 9500 Gilman Dr., La Jolla, CA, 92093; 5) Cleveland Clinic, Lerner College of Medicine, Case Western Reserve Univ. Cleveland, OH, 44195.

Ménière's disease (MD) is a debilitating disorder of the inner ear characterized by cochlear and vestibular dysfunction. The etiology of this disease is still unknown and epidemiological data for Ménière's disease has not been widely collected, with a small number of groups collecting the bulk of the data in a small number of cohorts. In this paper we characterize a large cohort of definite Ménière's disease patients for gender, age of onset, and use molecular genetic methodologies to characterize their ancestry. We confirm previous observations that there is a sex bias for the disease but note that both sexes have nearly identical age of onset values. Association analysis of genome-wide molecular genetic data for these samples failed to identify any variants associated with early onset of disease. Interrogation of these same data sets with principal component analysis allowed detailed observations about the ethnic ancestry of our patients. Surprisingly, comparison of the ethnicity of MD patients presenting to our tertiary care clinic with the ethnic composition of all patients visiting the clinic revealed a significant ethnic bias, with patients of Caucasian ancestry representing the most susceptible population and individuals with Asian, Hispanic and African ancestry being significantly less susceptible. To the best of our knowledge, this report is the first ethnic characterization of a large Ménière's disease cohort from a large metropolitan region using molecular genetic data. Our data suggest that there is a significant bias in ethnic disease susceptibility.

2185T

Genomics of Response to Healing following a Meniscectomy. *D.D. Vance[1], L. Wang[2,3], E. Rampersaud[2,3], T. Guettouche[2,3], W.H. Cade[2,3], R.L. Belton[2,3], B.P. Lesniak[1], J.M. Vance[2,3], M.A. Pericak-Vance[2,3], L.D. Kaplan[1].* 1) Division of Sports Medicine, Department of Orthopaedics, University of Miami, Miami, FL, USA; 2) John P. Hussman Institute for Human Genomics, University of Miami, Miami, FL, USA; 3) Dr. John T. Macdonald Foundation Department of Human Genetics, University of Miami, Miami, FL, USA.

Meniscal tear is a common orthopedic diagnosis that is often treated surgically by an arthroscopic partial meniscectomy, followed by physical therapy. Previous studies examining post-meniscectomy patients have shown that although the procedure is efficacious for some individuals, others retain chronic knee symptoms and functional limitations. While important, age, BMI, location of injury, and quality of physical therapy do not explain all the individual variations. One factor yet to be extensively examined is the contribution of genomic differences between individuals. Molecular analysis of age and sex-related gene expression of tissue collected from meniscal tears has shown significantly elevated expression levels of certain arthritis-related markers. The goal of this project is to better characterize the gene expression profiles in synovial fluid following meniscectomy between individuals from various ethnicities with varying levels of response. By comparing these results to the individual's unaffected knee analytic variability can be reduced. The Genomics of Response to Healing following a Meniscectomy project enrolls individuals through the U of Miami Sports Medicine clinic who have suffered a meniscus tear and are undergoing a partial meniscectomy. Baseline data gathered prior to surgery includes history of present illness (HPI) and applicable knee measurements and knee strength tests. At the time of surgery, intraoperatively, synovial fluid is collected from the affected knee for RNA extraction and gene expression analysis. Following surgery, the study participants undergo a standardized physical therapy regimen along with 2 and 6 month follow-up visits where quantitative assessments of knee function are administered by collecting knee measurements and the results of knee strengthening exercises using the Biodex from both the repaired and control knees of the same individual. Questionnaires are administered to collect information on demographics, medical history, lifestyle modifications as well as self-assessment of post-opt knee function and pain. Results of both the clinical knee assessment and questionnaires are used to compare and contrast gene expression profiles which are currently being assayed. Study design and implementation protocol along with baseline and expression data will be presented.

2186F

GWAS of host-pathogen interactions implicates methionine salvage and microtubules in regulation of inflammatory cell death. *D.C. Ko[1, 2, 3], K.P. Shukla[4], S.I. Miller[4].* 1) Department of Molecular Genetics & Microbiology, Duke University, Durham, NC; 2) Department of Medicine, Duke University, Durham, NC; 3) Center for Human Genome Variation, Duke University, Durham, NC; 4) Department of Microbiology, University of Washington, Seattle, WA.

Recent efforts by many laboratories have made progress in cataloguing human genetic variation and in identifying those variants associated with disease. When genome-wide association studies (GWAS) are combined with approaches that test how variants alter physiology, new biological insights can emerge. Here we used such an approach to reveal unexpected connections between inflammation, metabolism, and the cytoskeleton. We measured phenotypic variation in over 400 HapMap lymphoblastoid cell lines following infection with Salmonellae. Multiple read-outs of infection and host response were measured and genome-wide association of these traits revealed cis-eQTLs associated with pro-inflammatory Salmonella-induced cell death. One of the SNPs was associated with expression of APIP, an enzyme in the methionine salvage pathway. The direction of the association, RNA interference, and overexpression studies indicated that APIP is a previously uncharacterized inhibitor of inflammatory cell death. Furthermore, the precursor to the methionine salvage pathway, methylthioadenosine, was also demonstrated to regulate inflammatory cell death. A second SNP was associated with expression of an isoform of β-tubulin. Experimental studies demonstrated this protein inhibits inflammatory cell death as well, and may provide clues as to the mechanism of the microtubule-depolymerising drug colchicine in treatment of gout and other inflammatory diseases. Our results demonstrate novel insights into basic cell biology with clinical implications that are emerging from the application of GWAS to cellular phenotypes.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2187W

An integrative approach to mapping complex, glycemic traits in skeletal muscle tissue. *S. Keildson[1], J. Fadista[2], C. Ladenvall[2], A.K. Hedman[1], H-F. Zheng[3], L. Groop[2], P. Franks[2], O. Hansson[2], C.M. Lindgren[1], the Muther Consortium.* 1) Wellcome Trust Ctr, Roosevelt Drive, Oxford, OX37BN, United Kingdom; 2) Lund University Diabetes Centre CRC, Malmö University Hospital, Jan Waldenströms gata 35, Malmö, Sweden; 3) Department of Medicine, Human Genetics, Epidemiology and Biostatistics, Jewish General Hospital, McGill University, Montreal, Quebec, Canada.

Integrating genome-wide assays of gene expression and genome-wide genotype data may advance our understanding of the biological mechanisms related to diseases and traits, by pinpointing genetic variation in gene expression associated with disease. The MuTHER (Multiple Tissue Human Expression Resource) project lends itself to such analyses, as it includes SNP-genotyping, mRNA and miRNA expression, methylation and metabolite profiling data from a range of tissues collected from well-phenotyped female twins. Since skeletal muscle is an important peripheral tissue in insulin resistance, we examined variation in gene expression, from 9641 genes, associated with various glycemic traits (fasting glucose, fasting insulin, homeostasis model assessment-estimated insulin resistance (HOMA-IR) and HOMA-IR adjusted for BMI (HOMA-IRadjBMI)), as well as metabolic traits (BMI) in 39 skeletal muscle tissue samples, available as a part of MuTHER. We looked for associations between gene transcripts and phenotypes using a linear mixed-effects model adjusted for age, batch and twin structure. At a false discovery rate (FDR) of 5%, we found that 117 transcripts in skeletal muscle were associated with fasting glucose levels (pvalue<8.3E-04), 21 with fasting insulin levels (p-value<1.2E-04), 10 with HOMA-IR (p-value<6.2E-05) and 790 transcripts were associated with HOMA-IR(adjBMI) (pvalue<6.2E-04). Furthermore, 80 transcripts were associated with BMI (p-value<5.3E-04). Using the same mixed-effects model and 1000 Genomes imputed data, we then looked for associations between gene expression levels and genomic variants residing within a 1Mb window of the transcript. In total, 190 (~2%) transcripts were shown to have significant cis-regulatory effects (FDR < 5%, p-value<5.09E-06). When overlapping the results from these two analyses, we found that 1, 3 and 20 significant cis-eQTLs were also significantly associated with HOMA-IR, glucose and HOMA-IR(adjBMI) respectively, highlighting the potential importance of genetic control mechanisms on complex trait variance. To increase sample size and power, we will meta-analyse our data with similar data from the Malmo Men (n=38) and Malmo intervention (n=39) studies. Our study will facilitate mapping the genetic architecture of complex traits related to skeletal muscle.

2188T

cis-eQTL analysis of dysregulated interferon-pathway genes identifies OAS1 as a novel candidate for susceptibility to Sjögren's syndrome. *H. Li[1,2], J.A. Ice[1], J.A. Kelly[1], I. Adrianto[1], S.B. Glenn[1], K.S. Hefner[3], E.S. Vista[4], D.U. Stone[2], R. Gopalakrishnan[2], G.D. Houston[2], D.M. Lewis[2], M.D. Rohrer[5], P. Hughes[5], J.B. Harley[6,7], C.G. Montgomery[1], J. Chodosh[8], J.A. Lessard[9], J-M. Anaya[10], B.M. Segal[11], N.L. Rhodus[5], L. Radfar[2], M.B. Frank[1], R.H. Scofield[1], C.J. Lessard[1,2], K. Moser Sivils[1,2].* 1) Arthritis & Clinical Immunology, OMRF, Oklahoma City, OK; 2) University of Oklahoma Health Sciences Center, Oklahoma City, OK; 3) Hefner Eye Care and Optical Center, Oklahoma City, OK; 4) University of Santo Tomas Hospital, Philippines; 5) University of Minnesota, Minneapolis, MN; 6) Cincinnati Children's Hospital Medical Center, Cincinnati, OH; 7) Department of Veterans Affairs Medical Center, Cincinnati OH; 8) Harvard Medical School, Boston, MA; 9) Valley Bone and Joint Clinic, Grand Forks, ND; 10) Universidad del Rosario, Bogotá, Colombia; 11) Hennepin County Medical Center, Minneapolis, MN.

Sjögren's syndrome (SS) is a progressive autoimmune exocrinopathy present in 0.7-1% of Europeans characterized by symptoms of dry eyes and mouth. Dysregulation of interferon-related pathways is well documented in SS and related disorders, like systemic lupus erythematosus. We performed cis-expression quantitative trait loci (c-eQTL) analysis in 132 cases to identify genetic variants regulating proximal IFN-regulated genes through integration of gene expression data with genotypes from a SS genome-wide association study (GWAS). To further characterize the expression of interferon-inducible (IFI) genes in SS, we first performed global gene expression profiling (GEP) in peripheral blood on 48803 probes using the Illumina HumanWG-6 v3.0 Expression BeadChip in 201 SS cases and 76 healthy controls. Analyses were performed in the R Bioconductor suite. After quality control, 20342 probes (in 15607 genes) were quantile normalized and Welch's t-tests, q-values, and fold changes (FC) were calculated. Differentially expressed (DE) genes between cases and controls were selected by: q<0.05 and FC >1.25 or <0.87. Pathway analysis for DE genes was carried out in Genomatix. Thirty-seven IFI genes were among the top 123 DE genes (q<1x10E-10), with interferon alpha-inducible protein 27 (IFI27) demonstrating the greatest fold change (FC=77.25; q=1.88x10E-14). Among pathways in which DE genes were identified, pathways involved in type I interferon-mediated signaling (31/70 genes; p=5.56x10E-14) and viral transcription and genome expression (49/140 genes; p=3.25x10E-16) were highly significant. We selected 19 (q<0.001) DE genes involved in type I interferon-mediated signaling pathways for use as phenotypic traits in eQTL analyses. Single nucleotide polymorphisms (SNPs) from GWAS data spanning 20kb up- and downstream of the target genes were tested for genetic association with expression levels by linear regression. C-eQTL analysis identified significant association of 23 correlated (r2>0.8) SNPs within and flanking the IFI gene 2'-5'-oligoadenylate synthetase 1 (OAS1) with expression levels. Evidence for association of the OAS1 SNP rs10774671, a splice-site polymorphism associated with multiple sclerosis and type 1 diabetes, increased from p=6.99x10E-3 in our SS GWAS (424 cases/2120 controls) to p=5.1x10E-8 in c-eQTL analysis. These results provide a promising novel candidate locus in and around OAS1 for future functional studies to better understand SS pathophysiology.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2189F

Integrative Genomic Analysis Identifies Multiple Loci Influencing the Human Immune Response to Influenza Vaccination. *L.M. Franco[1,2], K.L. Bucasas[1], J.M. Wells[3], D. Nino[3], X. Wang[1,4], G.E. Zapata[1,4], N. Arden[5], J.M. Quarles[5], M.S. Bray[4], R.B. Couch[2,3], J.W. Belmont[1,4], C.A. Shaw[1].* 1) Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX; 2) Department of Medicine, Baylor College of Medicine, Houston, TX; 3) Department of Molecular Virology and Microbiology, Baylor College of Medicine, Houston, TX; 4) Children's Nutrition Research Center, Baylor College of Medicine, Houston, TX; 5) Department of Microbial and Molecular Pathogenesis, Texas A&M University, College Station, TX.

Identification of the host genetic factors that contribute to variation in vaccine responsiveness may uncover important mechanisms affecting vaccine efficacy. To identify these, we designed an integrative genomic analysis (IGA) study combining genetic, transcriptional, and antibody response data in humans immunized with a seasonal influenza vaccine. We enrolled a discovery cohort (n = 119) and an independent validation cohort (n = 128) of ethnically homogeneous healthy adults. Genome-wide genotyping was performed. Global gene expression was measured in peripheral blood before and on days 1, 3, and 14 after vaccination. Serum antibody titers were measured before and on days 14 and 28. We show that this study design is powerful in identifying loci that exhibit a transcriptional response to vaccination, genetic regulation of expression (cis eQTL), and correlation between expression and the antibody response. Using mixed-model regression, we identified 945 SNP-transcript pairs, corresponding to 135 genes, with significant genotype-expression association (genotype effect $p < 5 \times 10^{-8}$) and a transcriptional response to the vaccine (day effect $p < 0.05$) in the discovery cohort. Of these, 114 genes were validated in the second cohort (genotype effect $p < 0.05$ and day effect $p < 0.05$), of which 97 would pass equally stringent thresholds in both cohorts (genotype effect $p < 5 \times 10^{-8}$, day effect $p < 0.05$). Twenty-three of these genes also showed significant correlation with the antibody response in both cohorts. Remarkably, a majority of these genes encodes proteins involved in intracellular transport and/or antigen processing. We hypothesized that at some loci the magnitude of the genetic effect could be different before and after the environmental perturbation. This type of effect, which would not be observed in a cross-sectional eQTL or genome-wide association study, could be explicitly examined with our longitudinal data. We found that an increase in the expression variance explained (R_g^2) after vaccination was a general feature of the loci identified. QTL analysis of the antibody response conditioning on expression supports a direct causal role in vaccine immunogenicity for the genes identified. We conclude that variation at these loci significantly influences the human immune response to the vaccine. More broadly, we demonstrate that a longitudinal IGA study design is more efficient in complex-trait gene identification than current cross-sectional methods.

2190W

Barrett's esophagus and esophageal adenocarcinoma show substantial polygenetic variance, with genetic overlap between the diseases. *W.E. Ek[1], D. Whiteman[2], D.M. Levine[3], T.L. Vaughan[4], S. MacGregor[1], in behalf of BEAGESS study investigators.* 1) Statistical Genetics, Queensland Institute of Medical Research, Herston, Queenslans, Australia; 2) Cancer Control Group, Queensland Institute of Medical Research, Herston, Queensland, Australia; 3) Department of Biostatistics, University of Washington, Seattle, Washington, USA; 4) Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA.

Most complex diseases are regulated by many causative variants that independently show only a small effect on the trait. These could easily be missed in a standard genome-wide association study (GWAS). However, even though the individual effect of each SNP is low, collectively they might account for a substantial proportion of variation. This implies that, as a complement to GWAS, we can gain information about the genetic architecture by combining effects across SNPs. Our aim was to estimate the variance explained (heritability) for esophageal adenocarcinoma (EA), a rapidly fatal cancer, and its main precursor Barrett's esophagus (BE) when including all genotyped SNPs. We used a new method developed to estimate variance explained for case-control studies of unrelated individuals. We also examined if EA and BE share a similar genetic background by using a large number of SNPs in one trait to predict the risk of developing the other trait. Systematic evaluation of the overlap between the traits is based on “profile” scores, for which we calculated the estimated relative risk for every SNP based on a `discovery' set (BE) and then tested this “profile' in a target set of interest (EA). Our results, based on 1503 EA cases, 2383 BE cases and 3202 controls, show a significant variance explained for both EA (0.28, se= 0.05) and BE (0.39, se=0.05) when including all genotyped SNPs. We also found a substantial polygenetic overlap between EA and BE (p=1e-12). Given this overlap, we expect combining EA and BE as a larger case set versus all controls will improve power to map some loci. Examination of other risk factors showed that the genetic overlap was larger for patients with low BMI and non smokers. These results suggest that the genetic architecture for both diseases consist of many common variants, each of small effect, and the diseases seem to have a shared genetic background.

2191T

Rare germline copy number variants in Hodgkin lymphoma families. *M. Rotunno, M.L. McMaster, L.R. Goldin.* NIH/NCI/DCEG, Rockville, MD.

Hodgkin lymphoma (HL) shows strong familial aggregation but no major susceptibility genes for HL have been identified to date. Studies based on detection of copy number variation (CNV) are promising for identifying disease susceptibility genetic variants. The goal of this study was to identify high-penetrance CNVs in HL-prone families. Using array comparative genomic hybridization (aCGH) from © Oxford Gene Technology, we measured genome-wide germline CNVs in 45 HL cases or obligate carriers from 12 HL-prone families with three or more affected members (5 cases in 3 families, 4 cases in 4 families, and 3 cases in 5 families). The CytoSure Interpret software was used to analyze the CNV data. Each HL case presented with 67 to 150 CNVs (median 101), of average size 1.1Mb (1Kb to 45.3Mb). According to a dominant segregation model, between 10 and 32 (median 13) CNVs were shared by all cases or obligate carriers within each family. Since high-penetrance variants are likely uncommon in the general population, we then filtered out CNVs with frequency larger than 5% based on information from the Toronto Database of Genomic Variants (DGV). We thus identified 3 to 16 (median 8) rare shared CNVs per family, for a total of 56 unique CNVs (29 gains and 27 losses). Thirty-nine of these CNVs included genes, some of which are known to be involved in cancer. Forty-five of the identified CNVs were potentially novel, as they were not annotated in DGV. Thirty-five (33 novel) CNVs occurred in multiple families. In particular, a 24Kb CNV loss (DGV frequency < 3%) on chromosome 3p11 involving EPHA3, a tyrosine-protein kinase receptor gene, was shared by two HL families. CNVs in this gene have been associated with multiple types of hematological malignancies in previous somatic studies. In addition, CNV gains in 10q22.1, 6q16.3, and 3q26.1 were shared by 7, 9, and 11 families, respectively. In conclusion, our results, if validated, indicate that rare germline CNVs may play a role in familial HL risk.

2192F

The proportion of heritability of type 2 diabetes estimated through GWAS that is attributable to functional genetic variation. *J.M. Torres[1], E.R. Gamazon[2], S.K. Das[5], H.M. Highland[4], G.I. Bell[3], C.L. Hanis[4], N.J. Cox[2].* 1) Committee on Molecular Metabolism and Nutrition, University of Chicago, Chicago, IL; 2) Section of Genetic Medicine, University of Chicago, Chicago, IL; 3) Department of Medicine, University of Chicago, Chicago, IL; 4) Human Genetics Center, School of Public Health, University of Texas Health Science Center at Houston, Houston, TX; 5) School of Medicine, Wake Forest University, Winston-Salem, NC.

The growing prevalence of non-insulin dependent diabetes mellitus or type 2 diabetes (T2D) is a major health concern in the United States and disproportionately affects Mexican Americans. For example, Starr County, TX, where 97% of residents identify themselves as Hispanic, reports the highest diabetes-specific morbidity and mortality of any county in Texas despite being only the 53rd largest of Texas' 254 counties. Despite the identification of more than 60 highly significant and reproducibly associated variants for T2D, these variants do not, even collectively, account for much of the narrow sense heritability to T2D. Here, we apply genome-wide complex trait analysis (GCTA) software developed by Visscher et al. to a dataset consisting of 1,273 adult residents from Starr County (837 diabetic cases, 436 normoglycemic controls) genotyped on the Affymetrix Genome-Wide SNP Array 6.0 in order to estimate the phenotypic variance explained by the SNPs interrogated in the GWAS. After performing quality control measures and correcting for sex, BMI, and age at diagnosis, we applied the restricted maximum likelihood analysis (REML) implemented in GCTA and report that ~37% of the phenotypic variance in our dataset is explained by the additive genetic contribution of the interrogated SNPs. This quantity is approximately four times the amount of heritability for T2D often attributed to SNPs implicated through GWAS. We then extracted a subset of SNPs from our dataset that were mapped as expression quantitative trait loci (eQTLs) in either human skeletal muscle or adipose tissue and determined that the proportion of the total phenotypic variation attributable to these SNPs is significant. Furthermore, we applied analyses to validate that these trends are robust to population stratification in this admixed population. These results support the conclusion that a substantial fraction of the heritability of T2D estimated from the variation interrogated in GWAS can be attributed to a subset of genetic variation that impacts the function of metabolic tissues relevant to T2D.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2193W

Proportion of heritability and implications of pleiotropy attributable to GWAS hits for ECG phenotypes in the Erasmus Rucphen Family Study. *A. Isaacs[1,2], C. Silva[1,3], J.A. Kors[4], B.A. Oostra[1,2], C.M. van Duijn[1,2].* 1) Dept. of Epidemiology, Erasmus MC Rotterdam, Rotterdam, the Netherlands; 2) Centre for Medical Systems Biology, Leiden, the Netherlands; 3) Dept. of Biomedical Sciences, Universidad del Rosario, Bogotá, Colombia; 4) Dept. of Medical Informatics Erasmus MC Rotterdam, Rotterdam, the Netherlands.

Electrocardiogram (ECG) parameters are important measures of cardiac electrophysiology and provide crucial information on cardiac function. Genome-wide association studies (GWAS) achieved remarkable success over the past few years in locating genetic loci influencing a broad variety of ECG phenotypes, such as QRS complex duration, PR interval, and QT interval. Despite these advances, however, large portions of the variability of these traits remain unexplained by either genetic or environmental determinants, giving rise to the so-called "missing heritability" problem. The Erasmus Rucphen Family Study (ERF), consisting of more than 3000 individuals descendant from a limited number of founders approximately 250 years ago and connected in a single large pedigree, provided an excellent opportunity to directly estimate the effects of genes on heritability (h^2). Participants were genotyped using Illumina 318/370K single nucleotide polymorphism (SNP) arrays and were extensively phenotyped at baseline, which included 12-lead ECG measurements. Variance component methodology, as implemented in the SOLAR software package, was used to estimate the h^2 of several ECG phenotypes (QRS, QT and PR) and to determine the proportion of h^2 attributable to common genetic variants discovered in GWAS. Differences between models were tested using likelihood ratio tests, with degrees of freedom equivalent to the number of SNPs included in the saturated models. In models adjusted for age, gender, body-mass index, height and heart rate, highly significant ($P < 1.0x10^{-7}$) h^2 (± standard error) were observed for QRS (0.34 ± 0.06), PR (0.38 ± 0.06) and QT (0.34 ± 0.07). Inclusion of SNPs determined from GWAS for each trait decreased the estimated h^2 and increased the proportion of variance attributable to covariates (i.e. SNPs). GWAS SNPs decreased h^2 by 0.05 (15%, $P = 0.001$) for QRS, by 0.01 (1.6%, $P = 1.5x10^{-5}$) for PR, and by 0.09 (27%, $P = 2.2x10^{-15}$) for QT. The inclusion of all ECG SNPs (e.g. including QRS and QT SNPs in the PR model) further increased the explained proportion of h^2 by 0.2% for QRS ($P = 0.05$) and 8% for PR ($P = 5.7x10^{-4}$); for QT this difference (5.7%), although large, was non-significant. These results provide good estimates of the proportion of heritability attributable to common variants discovered thus far via GWAS. The inclusion of additional SNPs found for other ECG traits also increased the explained h^2, suggesting pleiotropic effects yet to be seen in GWAS.

2194T

The role of genes and life-course-related processes in onset of common diseases in late life. *A. Kulminski, I. Culminskaya, K. Arbeev, S. Ukraintseva, L. Arbeeva, A. Yashin.* Duke University, Durham, NC.

Complexity of aging-related diseases with post reproductive manifestation tempers progress in genome-wide association studies (GWAS) and contributes to the problem of missing genetic variance. Different strategies to address this problem are extensively discussed (Nat Rev Genet. 2010; 11, 446-450). One issue was that this problem is overblown. This presentation provides evidences supporting this view. We use the Framingham Heart Study original (FHS) and Offspring (FHSO) cohorts to elucidate whether life-course-related processes can substantially impact the role of lipid-related genes, the apolipoprotein E (APOE) e2/3/4 polymorphism and the APOB rs1042034 (C/T) SNP, in onset of cardiovascular disease (CVD). It is shown that the APOE e4 allele and the rs1042034 minor allele homozygote (CC) can play detrimental, neutral, and protective role in etiology of CVD in different ages and environments in sex-specific fashion. The APOE e4 allele is primarily associated with CVD in women. The e4 allele confers risk of CVD in younger-old women ($RR_{\leq 75yrs}$=1.73, p=1.4×10^{-3}) but it can be cardio-protective in older women (RR_{76+yrs}=0.68, p=0.021) in the same FHS cohort. Disregarding the role of aging-related heterogeneity results in insignificant effect (RR=0.84, p=0.143). The role of the e4 allele can change through generations which is a proxy for environment; the e4 allele in the FHSO only confers risks of CVD in older women ($RR_{>64yrs}$=1.57, p=6.6×10^{-3}). Just increasing the sample size by pooling data from the FHS and the FHSO cohorts merely scales the life-course-related heterogeneity in parallel (RR= 1.03; p=0.701). The APOB CC homozygote is associated with CVD in men. It confers highly significant risk of CVD in the FHS (RR=2.18, p=4.5×10^{-5}). In the FHSO, the CC homozygote can be cardio-protective although this association is not significant (RR=0.78, p=0.352). Pooling data from both cohorts makes the estimates to be significantly worse than in the FHS alone (RR=1.30, p=0.088). The observed associations are stable longitudinally ensuring that they are not the result of stochastic realizations at a specific assessment and robust to longitudinal attrition of the samples at CVD risks. The results suggest that life-course-related processes can play a key role in genetic predisposition to healthspan. More detail systemic analyses beyond those offered in standard GWAS will substantially advance the progress in the field.

2195F

Ancestry directed sequence analysis for rare variants in admixed populations. *E. Ziv[1,2,3], D. Hu[1,3].* 1) Medicine, UCSF, San Francisco, CA; 2) Helen Diller Family Comprehensive Cancer Center, UCSF, San Francisco, CA; 3) Institute for Human Genetics, UCSF, San Francisco, CA.

Genome wide association studies (GWAS) have only identified a fraction of the heritability of most complex traits. Much of the remaining heritability may be due to rare variants that are not well tagged by arrays commonly used in GWAS. Here we evaluate approaches to identifying rare variants in populations of mixed ancestry. Our approach is motivated by datasets from populations of mixed ancestry that have already undergone GWAS and where a majority of heritability remains unexplained. First, we demonstrate using data from sequencing studies that rare variants in an admixed population are likely to be private to one ancestral population. Then we demonstrate that such private alleles will create deviations in locus specific ancestry in cases. Finally, we consider strategies for leveraging the locus specific ancestry data in sequencing studies. We note that selection of individuals based on their locus specific ancestry can substantially increase the power/efficiency of sequencing analysis. We also examine the case where overall individual ancestry is associated with the trait of interest. We demonstrate that a selection strategy that identifies patients with discordance between local and global ancestry can be more powerful than one which focuses on locus specific ancestry only.

2196W

The Convergence of Functional Genomics, Heritability Estimation and Polygenic Modeling: Emerging Spectrum of Allelic Variation in Bipolar Disorder. *E.R. Gamazon[1], H.K. Im[1], C. Liu[2], D.L. Nicolae[1], N.J. Cox[1].* 1) University of Chicago, Chicago, IL, USA; 2) University of Illinois, Chicago, IL, USA.

It is widely held that a substantial genetic component underlies Bipolar Disorder and other neuropsychiatric disease traits. Recent efforts have been aimed at understanding the genetic basis of disease susceptibility, with genome-wide association studies (GWAS) unveiling some promising associations. Nevertheless, the genetic etiology of Bipolar Disorder remains elusive with a substantial proportion of the heritability - which has been estimated by ourselves and others to be 80% (congruent with previous reports based on twin and family studies) - unaccounted for by the specific genetic variants identified by large-scale GWAS. Furthermore, functional understanding of associated loci generally lags discovery. Studies we report here provide considerable support to the claim that substantially more remains to be gained from GWAS on the genetic mechanisms underlying Bipolar Disorder susceptibility, and that a large proportion of the variation in disease risk may be uncovered through integrative functional genomic approaches. We set out to combine recent analytic advances in heritability estimation and polygenic modeling and leverage recent technological advances in the generation of -omics data to evaluate the nature and scale of the contribution of functional classes of genetic variation to a relatively intractable disorder. We identified 2,375 cis-acting eQTLs in cerebellum that account for 36% (s.e. = 0.024) of the heritability, which represents 65% of the total heritability attributable to SNPs interrogated through GWAS. Our findings show that a much greater resolution may be attained than has been reported thus far on the number of common loci that capture a substantial proportion of the heritability to disease risk and that, importantly, the functional nature of contributory loci may be clarified en masse.

2197T

Contributions of cis and trans eQTLs to the heritability of schizophrenia. *C. Jumper[1], E.R. Gamazon[2], A. Konkashbaev[2], C. Liu[3], D.L. Nicolae[1,2,4,5], N.J. Cox[1,2,4].* 1) Department of Human Genetics, University of Chicago, Chicago, IL; 2) Section of Genetic Medicine, University of Chicago, Chicago, IL; 3) Department of Psychiatry, University of Chicago, Chicago, IL; 4) Department of Medicine, University of Chicago, Chicago, IL; 5) Department of Statistics, University of Chicago, Chicago, IL.

We estimated the heritability of variants interrogated in a genome-wide association (GWA) study of schizophrenia (NIMH nonGAIN Schizophrenia) and then determined the proportion of that heritability attributable to cis and trans-regulatory SNPs (eQTLs) in human brain using mixed-effects models. While the heritability estimated for all SNPs is 0.86, the heritability for 311973 cis eQTLs is 0.61, or 71% of the overall heritability. The heritability for 41404 trans eQTLs is 0.20, or 23% of the overall heritability. The total observed heritability of schizophrenia is consistent with other study estimates of approximately 0.8. This study clearly traces the genetic heritability due to identification of cis and trans components, demonstrating how future work would be able to estimate heritability partitioned by gene class. This should enable further insight into the mechanism of schizophrenia through identification of important genetic susceptibility regions and pathways.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2198F

The contribution of sequencing based GWAS in explaining the missing heritability of quantitative traits. *E. Porcu[1,2,3], C. Sidore[1,2,3], M. Steri[1], F. Busonero[1,3], A. Mulas[1], M. Dei[1], S. Lai[1], MG. Piras[1], S. Naitza[1], D. Schlessinger[4], G. Abecasis[3], S. Sanna[1], F. Cucca[1,2].* 1) Istituto di Ricerca Genetica e Biomedica (IRGB), CNR, Monserrato, Italy; 2) Dipartimento di Scienze Biomediche, Università di Sassari, SS, Italy; 3) Center for Statistical Genetics, University of Michigan, Ann Arbor, MI, USA; 4) Laboratory of Genetics, NIA, Baltimore, MD, United States.

Genome-wide association studies (GWAS) have yielded valuable insights into the genetic architecture of complex quantitative traits, but have often left much of the trait heritability unexplained. To assess the hypothesis that a portion of unexplained heritability can be due to variants not included or poorly tagged in common genotyping arrays and in the HapMap dataset, we analyzed 6,145 individuals from the SardiNIA study, for a set of ~8 million variants, directly genotyped or imputed with a reference panel derived from whole-genome low-pass sequencing of 1,146 Sardinian individuals, focusing on four quantitative traits: hsCPR, IL-6, MCP-1 and ESR. When considering only loci previously identified by us in the same cohort, for the first two traits we observed no improvement in the heritability, the top variants resulting from this analysis being in strong linkage disequilibrium (LD) with those already described. By contrast, for MCP-1 and ESR, we observed different leading variants not in LD with those found previously, as well as additional independent signals, which led to an increase in the heritability explained from 9.4% to 11.6%, and from 2.7% to 3.9%, respectively. The imputation of variants detected by sequencing also contributed to the identification of two novel loci for both MCP-1 and hsCRP. The inclusion of these markers further increased the heritability explained to 12.5% for MCP1, and from 5.0% to 6.7% for hsCRP. Our results thus indicate that the forthcoming sequencing based GWAS will contribute to better account for the heritability at previously described loci, and will further help to explain the missing heritability with the identification of novel associations.

2199W

Polygenic analysis shows that low frequency SNPs confer early onset myocardial infarction risk. *E. Stahl[1,2], R. Do[2,3], B. Vilhjamsson[2,4], N. Stitziel[3], S. Sunyaev[2,5], A. Price[2,4], R.M. Plenge[1,2,5], S. Kathiresan[2,3], S. Raychaudhuri[1,2,5].* 1) Dept. Medicine, Div. Rheumatology, Immunology and Allergy, Brigham and Women's Hospital, Boston, Massachusetts, USA; 2) Program in Medical and Population Genetics, Broad Institute, Cambridge, Massachusetts, USA; 3) Center for Human Genetic Research and Cardiovascular Research Center, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts, USA; 4) Dept. Biostatistics, Harvard School of Public Health, Boston Massachusetts, USA; 5) Dept. Medicine, Div. Genetics, Brigham and Women's Hospital, Boston, Massachusetts, USA.

Myocardial infarction (MI) and coronary artery disease (CAD) is a heritable trait that is caused by many loci. Recent genome-wide association studies have identified common variants associated with MI/CAD in multiple loci (Schunkert et al 2011 Nature Genetics); however, these loci explain only a small proportion of the heritability of MI/CAD, and many more false negative loci have not achieved the stringent statistical evidence required for discovery. Studies have identified a polygenic component of heritability, consisting of thousands of associated SNPs in GWAS data, for several complex diseases including MI (Stahl et al 2012 Nature Genetics). Another component of heritability could be rare variants that are not well captured by GWAS data. If there are enough variants with sufficient allele frequency (e.g. 0.1-5%) and effect size, these could contribute greatly to phenotypic variance in the population. Resequencing studies have identified low frequency SNPs that confer risk for MI, such as the PCSK9 R46L nonsynonymous variant; however, it is currently unknown whether these variants make up a substantial proportion of the missing heritability for complex disease.

We have imputed MIGEN and WTCCC CAD GWAS data into both common and low frequency SNPs using reference panels from whole genome and exome sequencing data (1000G CEU Dec 2010 release, and phased exome sequence+GWAS data) and performed heritability and polygenic risk score analysis for MI and CAD. We show that addition of low-frequency variants increases linear mixed model heritability estimates (MAF>5%, common SNP h2=0.38, P=3e-15; MAF>0.5%, h2=0.44, P=2.5e-15; 0.5%<MAF<5%, h2= 0.085, P=0.04), and quantify the overall contributions of SNPs from HapMap, 1000G and the exome. Furthermore, we find greater polygenic risk score prediction R2 for genic (R2=0.018, P=7e-16) versus intergenic (R2=0.0039, P=2e-4) SNPs, but not of nonsynonymous (R2=0.008, P=7e-8) vs other genic (R2=0.018, P=8e-16) SNPs. These results demonstrate that low-frequency variants make a detectable contribution to MI/CAD heritability, and help to better understand the genomic structure of their contribution.

2200T

Ubiquitous polygenicity of human complex traits: genome-wide analysis of 49 traits in Koreans. *J. Yang[1], T. Lee[2], J. Kim[3], S. Cho[4], P. Visscher[1,5], H. Kim[2,3,4].* 1) University of Queensland Diamantina Institute, University of Queensland, Brisbane, Queensland, Australia; 2) Department of Agricultural Biotechnology, Seoul National University, Seoul, Korea; 3) Interdisciplinary Program in Bioinformatics, Seoul National University, Seoul, Korea; 4) C&K Genomics, Seoul, Korea; 5) The Queensland Brain Institute, The University of Queensland, Brisbane, Queensland 4072, Australia.

Recent studies in population of European ancestry have shown that 30-50% of heritability for human complex traits such as height (Yang et al. 2010) and body mass index (Yang et al. 2011), and common diseases such as schizophrenia (Lee et al. 2012) and rheumatoid arthritis (Stahl et al. 2012) can be captured by common SNPs, and that genetic variation can be attributed to chromosomes, in proportion to their length. Using genome-wide estimation and partitioning approaches, we analyzed 49 human quantitative traits, many of which are relevant to human diseases, in 7,170 unrelated Korean individuals genotyped on 326,591 SNPs. For 43 of the 49 traits, we estimated a significant (P < 0.05) proportion of variance explained by all SNPs (h2G). On average across 47 of the 49 traits for which the estimate of h2G is non-zero, 13.4% (range of 3.4% to 31.6%) of phenotypic variance can be explained by all the SNPs being analysed, or approximately one-third (range of 7.8% to 76.8%) of narrow sense heritability. In contrast, on average across 25 of the 49 traits, the top associated SNPs at genome-wide significance level (P < 5e-8) explain 1.5% (range of 0.5% to 3.8%) of phenotypic variance. The majority (~92%) of explained variation estimated from all SNPs is captured by the SNPs with p-values < 0.031 in single SNP association analyses. Longer genomic segments tend to explain more phenotypic variation, with a correlation of 0.78 between the estimate of variance explained by individual chromosomes and their physical length. This correlation was stronger (0.81) for intergenic regions. Despite the fact that there are a few SNPs with large effects for most traits, these results suggest that polygenicity is ubiquitous for most human complex traits, and that a substantial proportion of heritability is captured by common SNPs.

2201F

Prediction of human height with large panels of SNPs - insights into genetic architecture. *Y.C. Klimentidis[1], A.I. Vazquez[1], G. de los Campos[2].* 1) Energetics, University of Alabama at Birmingham, Birmingham, AL; 2) Biostatistics, University of Alabama at Birmingham, Birmingham, AL.

Prediction of complex traits from genetic information is an area of major clinical and scientific interest. Height is a model trait since it is highly heritable and easily measured. Substantial strides in understanding the genetic basis of height have recently been made through genome-wide association studies (GWAS), and whole-genome prediction (WGP) which fits thousands of SNPs jointly. Here, we attempt to gain insight into the genetic architecture of human height by examining how WGP accuracy is affected by the choice of single-nucleotide polymorphism (SNPs). Specifically, we compare the prediction accuracy of models using: 1) SNPs selected based on the `top hits' of the GIANT consortium meta-analysis for height at different p-value thresholds, and 2) SNPs in genomic regions that surround the most significant `top hits'. We use the Framingham Heart Study and GENEVA datasets, imputed up to 10 million SNPs with 1000 Genomes reference data. The predictive accuracy of each model was evaluated in cross-validation. We find that prediction accuracy increases up to a certain point with the inclusion of more `top hits' from the GIANT study, that including SNPs from the regions surrounding `top hits' contributes minimally to prediction accuracy, and that prediction accuracy increases with the size of the training dataset. Finally, we find that prediction accuracy is greatest for individuals at the phenotypic extremes of height. Our results suggest that improvement of genomic prediction models will require the use of information from a large number of selected SNPs, and that these models may be most useful at the phenotypic extremes.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2202W

Heritability estimation of height from common genetic variants in a large sample of African Americans. *F. Chen[1], G.K. Chen[1], R.C. Millikan[2], E.M. John[3,4], C.B. Ambrosone[5], L. Berstein[6], W. Zheng[7], J.J. Hu[8], R.G. Ziegler[9], S.L. Deming[7], E.V. Bandera[10], W.J. Blot[7, 11], S.S. Strom[12], S.I. Berndt[9], R.A. Kittles[13], B.A. Rybicki[14], W. Issacs[15], S.A. Ingles[1], J.L. Stanford[16], W.R. Diver[17], J.S. Witte[18], L.B. Signorello[7,11], S.J. Chanock[9], L. Le Marchand[19], L.N. Kolonel[19], B.E. Henderson[1], C.A. Haiman[1], D.O. Stram[1].* 1) Preventive Medicine, University of Southern California, Los Angeles, CA; 2) Epidemiology, Gillings School of Global Public Health, and Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, NC; 3) Northern California Cancer Center, Fremont, CA; 4) School of Medicine, Stanford University, and Stanford Cancer Center, Stanford, CA; 5) Cancer Prevention and Control, Roswell Park Cancer Institute, Buffalo, NY; 6) Cancer Etiology, Population Science, Beckman Research Institute, City of Hope, CA; 7) Epidemiology, Vanderbilt Epidemiology Center and Vanderbilt-Ingram Cancer Center, Vanderbilt University, Nashville, TN; 8) Sylvester Comprehensive Cancer Center, Department of Epidemiology and Public Health, University of Miami, Miami, FL; 9) Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bathesda, MD; 10) The Cancer Institute of New Jersey, New Brunswick, NJ; 11) International Epidemiology Institute, Rockville, MD; 12) Epidemiology, The University of Texas M.D. Anderson Cancer Center, Huston, TX; 13) Medicine, University of Illinois at Chicago, Chicago, IL; 14) Biostatistics and Research Epidemiology, Henry Ford Hospital, Detroit, MI; 15) James Buchanan Brady Urological Institute, Johns Hopkins Hospital and Medical Institutions, Baltimore, MD; 16) Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA; 17) Epidemiology Research, American Cancer Society, Atlanta, GA; 18) Institute of Human Genetics, Dept of Epidemiology and Biostatistics, University of California, San Francisco, CA; 19) Epidemiology, Cancer Research Center, University of Hawaii, Honolulu, HI.

Height has an extremely polygenic pattern of inheritance. Genome-wide association studies (GWAS) have revealed hundreds of common variants that are associated with human height at genome-wide levels of significance. Each of these common variants has a very modest effect, and only a small fraction of phenotypic variation can be explained by the aggregate of these common variants. In this large study of African-American men and women, we genotyped and analyzed 975,519 autosomal SNPs across the entire genome using a variance components approach, and found that 46.4% of phenotypic variation can be explained by these SNPs in a sample of 9,779 evidently unrelated individuals. We noted that in two samples of close relatives defined by probability of identical-by-descent (IBD) alleles sharing (Pr (IBD=1)>=0.3 and Pr (IBD=1)>=0.4), the proportion of phenotypic variation explained by the same set of SNPs increased to 75.5% (se: 14.8%) and 70.3% (26.9%), respectively. We conclude that the additive component of genetic variation for height may have been overestimated in earlier studies (~80%) and argue that this proportion also includes variation from epistatic effects. Using simulation, we showed that by using common SNPs that are only weakly correlated with causal SNPs, the model could explain a large proportion of heritability. We therefore argue that the heritability estimate from the variance components approach is not necessarily the variation explained by a given set of SNPs, but also possibly reflects distant relatedness between nominally unrelated participants. Finally, we explored the performance of the variance components approach and concluded that the approach fails when a large number of independent variables are included in the model as the structure of the two components becomes similar. Thus some degree of population stratification seems to be required in order for the method to perform well for very large numbers of SNPs; however when modest stratification is present there is a risk of miss-attribution of effects of unmeasured (and untagged) variants to measured variants.

2203T

Optimizing risk stratification to discover novel neuroblastoma genes. *A. Goldenberg[1,2], B. Wang[1], S. Diskin[4], J.M. Maris[4], M. Irwin[3].* 1) Genetics and Genome Biology, SickKids Research Institute, Toronto, ON, Canada; 2) Department of Computer Science, University of Toronto, Toronto, ON, Canada; 3) Cell Biology, SickKids Research Institute, Toronto, ON, Canada; 4) Pediatrics, Children's Hospital of Philadelphia, Philadelphia, PA.

Heterogeneity in the clinical definition of complex traits is one of the major obstacles to identifying genetic risk factors. Neuroblastoma (NB), a sympathetic nervous system tumor is the most frequent cause of pediatric cancer-related deaths. The classification for NB into low (LR), intermediate (IR) and high risk (HR) groups is based on age, tumor stage, histology, MYCN status, ploidy, and 11q aberrations. However, each subtype is still very heterogeneous. Importantly, for the HR group the classification provides little guidance for estimating survival and response to therapy. Importantly, we and others hypothesize that there is significant NB germline variation based on 1) early age of onset; 2) fairly quiet somatic landscape of NB. Thus, we have developed a non-parametric, unsupervised method to identify homogeneous clusters of patients (NB subtypes) and the associated genetic variants. We first transform the data into patient similarity (e.g using correlation). We then use a network fusion method to iteratively combine normalized Laplacians of the patient phenotype and genotype networks. We introduce a new feature (SNP) ranking score that is able to identify local contribution for each feature. A feature is identified as important if it is associated with one or more patient clusters. Our simulations show that our method is significantly more powerful than other approaches. Considering IR and HR neuroblastoma patient groups, without assigning these labels to the method, we identified 4 novel clusters - two for HR and two for IR patients. For example, one HR cluster was associated with amplified MYCN status, earlier NB onset and several variants in the T-cell activating gene CD28, while a second HR cluster was associated with MYCN non-amplification and a previously detected variant in BARD1 (Capasso et al, Nat Gen, 2009). The identified candidate SNPs are currently being analyzed for functional significance in vitro and in vivo. Our methodology has several advantages over existing Genome Wide Association methods: 1) it does not require knowledge of disease subtypes; 2) works with a large number of discrete and continuous phenotypes; 3) highly scalable; 4) robust to very large numbers of irrelevant features (SNPs); 5) can be easily applied to any genomic data in a p>>n setting, (p=number of features, n=number of samples), to identify the intrinsic clustering in the data and the small set of features that drive the coherence within those clusters.

2204F

Correlations between associated SNPs cause upward bias in whole-genome based heritability estimation. *C. Li[1], C. Yang[2], J. Ferguson[2], J. Cho[3,4], H. Zhao[2].* 1) Program in Computational Biology and Bioinformatics, Yale University, New Haven, CT; 2) Department of Biostatistics, Yale School of Public Health, New Haven, CT; 3) Department of Genetics, Yale University, New Haven, CT; 4) IBD Center, Division of Gastroenterology, Department of Medicine, Yale University, New Haven, CT.

Although results from initial genome wide association studies (GWAS) suggested that much of the heritability for human complex traits is “missing”, a series of papers published recently argued that common SNPs could explain a significant proportion of heritability. The underlying models behind these developments for `heritability' estimate are linear mixed models (LMM) using whole-genome common SNPs. However, these models make implicit assumptions that may not be satisfied by the underlying genetic model that affects the traits of interest. In this presentation, through both theoretical and empirical studies, we demonstrate that the estimates obtained in the published papers could have upward biases due to correlations among SNPs in real data. In addition, we show that smaller sample sizes lead to larger biases when correlations exist. For instance, when strong correlation exists between the SNPs associated with the phenotype, the heritability estimate can be increased from 0.25 to 0.86 and 0.94 when using 2000 and 1000 samples respectively. The bias can be reduced when ultra-stringent marker pruning is used based on linkage disequilibrium (LD). For example, for a Crohn's disease data set, with LD pruning at r^2 cut-off of 0.5, the heritability estimates based on all 2000 samples and 1000 subsamples are 0.28 and 0.32 respectively. However, as we do more stringent LD pruning, these two estimates get closer, eventually converging at 0.11 when we set r^2 cut-off at 0.01. Therefore, the heritability estimates based on GWAS data using the existing LMM methods should be interpreted with caution.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2205W

Enrichment of rare non-synonymous variants in 196 lipid-associated genes in subjects with extreme levels of HDL-C: Evidence for rare allele model in etiology of dyslipidemia. *M.M. Motazacker[1,2], J. Peter[1], M. Treskes[3], J.A. Kuivenhoven[4], G.K. Hovingh[2].* 1) Department of Experimental Vascular Medicine, Academic Medical Center, Amsterdam, Netherlands; 2) Department of Vascular Medicine, Academic Medical Center, Amsterdam, The Netherlands; 3) Department of Biochemistry, Tergooi Ziekenhuizen Hilversum, The Netherlands; 4) Department of Pathology and Medical Biology, Molecular Genetics, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands.

Plasma HDL-C levels are inversely correlated with cardiovascular disease but the molecular basis and clinical consequences of extreme HDL-C phenotypes are poorly understood. Meta-analysis of genome-wide association studies has revealed that common variants can only explain a minor fraction of heritability of these phenotypes. This study explores the spectrum of rare non-synonymous variants in lipid-associated genes in individuals with extreme HDL-C phenotypes. Eighty individuals with HDL-C levels below the 1st percentile and above the 99th percentile for age and gender (n=40 in each group) were selected from a clinical chemistry database of a community hospital. Using a custom DNA capture library (SureSelect, Agilent), coding regions and exon-intron boundaries of 196 genes implicated in lipid metabolism (lipid geneset) and 78 genes related to coagulation (control geneset) were captured and sequenced on an Illumina HiSeq2000 platform (average coverage >50x). We identified 124 and 136 rare (frequency < 0.5% in European-Americans) non-synonymous variants in lipid geneset present solely in the low-HDL-C or high-HDL-C groups, respectively. In coagulation geneset, 41 variants were identified. The lipid geneset was enriched for rare non-synonymous variants (OR 1.72; p=0.001) when compared to coagulation geneset after correcting for the number of nucleotides sequenced. This enrichment remained significant after excluding 12 well-established HDL-C genes (OR 1.69; p<0.005). Similar comparison of the two genesets in a Dutch general population cohort (n=500) did not reveal a significant difference. Furthermore, lipid geneset in extreme HDL-C cohort was enriched for rare non-synonymous variants when compared to the Dutch general population cohort (OR 1.36; p<0.05). A similar comparison for control geneset did not show a significant difference. Nine percent of the lipid gene variants in extreme HDL-c cohort were coding-indel, nonsense or splice-site variants while 66% concerned missense variants of conserved nucleotides. Eighty percent of the individuals in extreme HDL-C cohort carried at least two rare non-synonymous variants. Our data suggests that in the majority of cases, multiple rare non-synonymous variants in genes with established or proposed roles in lipid metabolism could be responsible for extreme HDL-C phenotypes. Cataloguing the spectrum of such variants may help in better understanding of complex inheritance patterns of the lipid phenotypes.

2206T

Missing heritability and the Use of Genomewide Expression Data as an Indicator of Genetic Variation in Genome Scans. *P. Schliekelman, G. Sun, S. Zhang.* Department of Statistics, University of Georgia, Athens, GA.

It has become clear that HapMap-based Genomewide Association Studies (GWAS) are not adequate to capture all of the genetic variation for complex traits. Thus, there is currently great interest in developing alternative strategies for exploring this variation. In this talk, we advocate for the use of gene expression information as an indicator of genetic variation in genome scans. Using simulations and data analysis, we highlight properties of gene expression variation that may give it advantages as an indicator of genetic variation relative to traditional molecular markers. Consider a genetic polymorphism that is causative for a disease and also for variation in expression in some set of transcripts. First, we show that the test of association between gene expression and disease is equivalent to the test of association between a marker and the disease, with variation in gene expression taking the place of variation due to recombination between marker and locus. As the quality of the expression information increases, the power to detect a disease locus converges to the power of a GWAS to detect a marker in complete disequilibrium with the disease locus. However gene expression has two advantages. First, it is much less sensitive to the number of alleles at the locus than is a GWAS. Second, recent evidence suggests that the influence of multiple causative loci may converge on single transcripts or groups of transcripts. If this is the case, then the expression level is an indicator of a multilocus genotype and then the power to detect the association can be increased substantially higher than that for the GWAS. Transcript groups can be linked to specific sequence polymorphisms using eQTL mapping. Simulations and an application to data show that an expression-based approach can have substantially higher power than a standard GWAS for detecting some disease polymorphisms.

2207F

Estimating the proportion of variation in susceptibility to PTSD captured by common SNPs. *C. Yang[1], J. Gelernter[2], C. Li[3], H. Kranzler[4], L. Farrer[5], H. Zhao[6].* 1) Department of Epidemiology and Public health, Yale University, New Haven, CT, 06, New Haven, CT; 2) Department of Psychiatry, Division of Human Genetics, Yale University School of Medicine, New Haven, Connecticut, 06516, USA; 3) Program in Computational Biology, Yale University, New Haven, CT, 06520. USA; 4) Department of Psychiatry 3900 Chestnut Street Philadelphia, PA 19104, USA; 5) Departments of Medicine, Neurology, Ophthalmology, and Genetics and Genomics, Boston University School of Medicine and Departments of Epidemiology and Biostatistics, Boston University School of Public Health, Boston, MA, USA; 6) Department of Epidemiology and Public health and Department of Genetics, Yale University, New Haven, CT, 06520, USA.

Post-traumatic stress disorder (PTSD) is a complex disorder caused both genetic and environmental factors. In order to understand the genetic architecture of PTSD, the subjects were recruited in the course of genetic studies of alcohol, cocaine, and opioid dependence from 5 US sites: Yale, Univ. CT, UPenn, the Medical University of South Carolina, and McLean Hospital of Harvard Medical School. Originally, the number of European-American (EA) subjects having GWAS data was 1974. Subjects who are related, who were determined to be population group outliers, had abnormal heterozygosity rate, and who were missing PTSD phenotype information were excluded, before estimating the proportion of variation in susceptibility to PTSD captured by common SNPs. After quality control, we have 300 cases, 1278 controls and 695,134 SNPs (SNPs with MAF<0.05, HWE<0.0001 are excluded). We have performed both genome-wide single-SNP analysis and genome-wide interactions analysis. We do not find genome-wide significant association signals for both single-SNP analysis and interaction analysis. However, we estimate that about 18% of variation in liability scale (heritability) is captured by these common SNPs collectively. We also show that the variance explained by each chromosome is linearly related to its length, except chromosome 10. Specifically, the marginal signals of chromosome 10 can are very weak (with smallest p-value 10e-05), but it can contribute to about 10% of heritability. Our result indicates that the genetic architecture of PTSD is polygenic (many genes contribute to the trait but with small effects). It also implies that more individual SNP associations can be identified as the sample size increases.

2208W

Simulation-based study investigating "Missing heritability" — a euphemism for the abject failure of modern genomics to elucidate the etiology of complex traits. *J.H. Lee[1, 2, 3], R. Cheng[1, 2], K.M. Weiss[9], T. Hiekkalinna[8, 10], J.D. Terwilliger[4, 5, 6, 7, 8, 10].* 1) Sergievsky Center, Columbia University, New York, NY; 2) Taub Institute, Columbia University, New York, NY; 3) Epidemiology, Columbia University, New York, NY; 4) Genetics and Development, Columbia University, New York, NY; 5) Psychiatry, Columbia University, New York, NY; 6) Columbia Genome Center, Columbia University, New York, NY; 7) Division of Medical Genetics, New York State Psychiatric Institute, New York, NY; 8) Chronic Disease Prevention, Public Health Genomics Unit, National Institute for Health and Welfare, Helsinki, Finland; 9) Anthropology, Pennsylvania State University, State College, PA; 10) Institute for Molecular Medicine - Finland, Helsinki, Finland.

Human geneticists are forced to deal with "missing heritability" when the findings from their experiments explain a small portion of the genetic component of the etiology of complex traits. Is there anything really "missing"? To address the question systematically, we developed two sophisticated software packages: ForSIM and simQTL. ForSIM simulates a phenogenetic relation over evolutionary time with risk alleles arising by mutation and then subjected to natural selection on phenotypes. simQTL simulates a phenogenetic architecture conditional on a set of epidemiological parameters. Both packages also simulate environmental risk factors as well (both discrete and continuous, familial and random). ForSIM simulation under a wide variety of assumptions consistently lead to genes with a large number of risk variants influencing the phenotype. The signal captured by tagSNPs in those genes explain a small portion of the actual heritability when the effects of all variants in the gene are combined. simQTL simulation allows us to examine the effects of common variants in hundreds of polygenes where those genes have varying effect size and minor allele frequency, and they are randomly scattered throughout the genome. These analyses show weak power even when the functional variants themselves are typed, with methods like GCTA explaining a reasonable portion of the actual polygenic variance. However, when markers in LD with the functional variants are simulated instead, much more of the heritability, as expected, appears to be "missing". We took this a step further, and integrated real sequence and GWAS data in the simulations to incorporate real LD structures, and assigned polygenic function to actual variants. When we repeated the process, the results were even worse because cryptic LD as it exists in the real world further blurs the interpretation. Our simulation experiments clearly show that current models failed to model complexity in a reasonable way. Interestingly GWAS argues strongly against CVCD as a model for common disease.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2209T

A quantitative population genetic framework to constrain the range of genetic architectures underlying complex human traits. *V. Agarwala[1,2,3,5], J. Flannick[1,5], S. Sunyaev[1,2,4], D. Altshuler[1,4,5].* 1) Program in Medical & Population Genetics, Broad Institute, Cambridge, MA; 2) Program in Biophysics, Harvard University, Cambridge, MA; 3) Harvard-MIT Division of Health Sciences and Technology, MIT, Cambridge, MA, USA; 4) Department of Genetics, Harvard Medical School, Boston, MA, USA; 5) Center for Human Genetic Research, Massachusetts General Hospital, Boston, MA.

The genetic architecture of complex human traits -the number, frequencies, and effect size distribution of causal alleles- determines both the success of disease gene mapping as well as the likely future of personalized medicine. Despite a vast array of genetic studies, debate continues about the relative contribution of rare and common variants to disease heritability. Given the growing body of data for many complex traits, a critical question is: to what extent can empirical data constrain the space of possible trait architectures?

We present a principled, population genetic framework to evaluate disease models in light of three classes of empirical data: a) observed human genetic variation, b) disease prevalence and heritability, and c) prior genetic mapping results. We forward simulate population-scale, full-sequence variation in hundreds of thousands of individuals. We select a human demographic history which includes recent explosive population growth, and also model natural selection. We show that the resulting simulated data, at both synonymous and non-synonymous sites, is consistent with empirically observed frequency spectra and linkage patterns in over 1,300 European samples.

We find that two disease model parameters - the mutational target size and the strength of coupling between purifying selection and phenotypic effects - are sufficient to produce a wide range of genetic architectures, including the highly debated “common polygenic” and “rare Mendelian” models of complex disease. For type 2 diabetes, we find that both these models were consistent with initial discovery genome-wide association studies (GWAS). Empirical data from larger scale GWAS (over 50K samples) and polygene score analysis, however, can now exclude both models. We rule out models positing either strong or no coupling to selection and demonstrate that the `common-vs-rare' dichotomy is artificial: models with genetic variance largely restricted to one frequency class are in fact inconsistent with T2D empirical data. More generally, as next-generation sequencing studies provide a deluge of empirical data, our framework offers a robust method by which to predict power in these studies, test precise genetic hypotheses, and ultimately further constrain the space of architectures underlying complex traits.

2210F

Multi-ethnic polygenic prediction in CARe and other cohorts contrasts the genetic architectures of height and body mass index. *E.K. Speliotes[1], J. O'Connell[2], R. Do[3,4], B. Vilhjalmasson[3,5], S. Pollack[5], Y. Gong[6], N. Patterson[3], M. Akylbekova[7], A. Cupples[8], M. Fornage[9], J. Hirschhorn[3,10], W.H.L. Kao[11], L. Lange[12], G. Lettre[13], M. Li[14], J. Mychaleckyj[15], S. Musani[16], G. Papanicolaou[17], J.I. Rotter[18], D. Siscovick[19], X. Zhu[20], J.G. Wilson[17], . the GIANT Consortium[3], J. Johnson[6], P.I.W. de Bakker[21], S. Raychaudhuri[3,22], R.M. Plenge[3,22], A. Price[3,5], E.A. Stahl[3,22].* 1) University of Michigan, Ann Arbor, MI; 2) University of Maryland, Baltimore, MD; 3) Broad Institute, Cambridge, MA; 4) Massachusetts General Hospital, Boston, MA; 5) Harvard School of Public Health, Boston, MA; 6) University of Florida College of Pharmacy, Gainesville, FL; 7) Jackson Heart Study, Jackson State University, Jackson, MS; 8) Boston University School of Public Health, Boston, MA; 9) University of Texas Health Sciences Center, Houston, TX; 10) Childrens Hospital, Boston, MA; 11) Johns Hopkins School of Public Health, Baltimore, MD; 12) University of North Carolina, Chapel Hill, NC; 13) Montreal Heart Institute, Montreal Quebec, Canada; 14) University of Pennsylvania School of Medicine, Philadelphia, PA; 15) University of Virginia, Charlottesville, VA; 16) University of Mississippi Medical Center, Jackson, MS; 17) National Heart, Lung, and Blood Institute, Bethesda, MD; 18) Cedars-Sinai Medical Center, Los Angeles, CA; 19) University of Washington School of Public Health, Seattle, WA; 20) Case Western Reserve University School of Medicine, Cleveland, OH; 21) University Medical Center Utrecht, The Netherlands; 22) Brigham & Women's Hospital, Harvard Medical School, Boston, MA.

We utilized polygenic prediction analyses to gain insight into the genetic architectures of two human quantitative traits, height and body mass index. Classical heritability methods applied to population GWAS data indicate that common SNPs explain substantial proportions of heritability across complex traits- about 50% for height and 30% for body mass index (BMI). Recent European GWAS meta-analyses for height (N~130,000; Lango Allen Nature 2010) and BMI (N~120,000; Speliotes Nature Genetics 2010) were used as training data. Test data included the PEAR antihypertensive response GWAS (N=450 Europeans [Eur], results shown below), a new GIANT cohort (N=~5,000 Eur) and the CARe cohort (N~8,000 African Americans [AA]). SNP sets were based on a range of P_training thresholds with and without LD-pruning (r2<0.1). Polygenic scores (sums of SNP dosages, weighted by effect sizes assuming a proportion p<=1 of SNPs associated [see Vilhjalmsson Abstract 120121058]) were tested for association to gender-, age- and ancestry-adjusted phenotypes. Polygenic scores based on large-scale European GWAS meta-analysis were highly significant predictors of height and BMI (max R2_height,Eur=0.20, P=3e-10; max R2_BMI,Eur=0.086, P=9e-11), and yielded prediction R2 far greater than genome-wide significant SNPs. Significant SNPs (P_training<0.01, 3,491 SNPs) achieved the best height prediction, whereas including many non-significant SNPs (P_training<0.2, 186K SNPs) improved BMI prediction. LD-pruning yielded better height prediction in Europeans, whereas all SNPs yielded better prediction of BMI in Europeans and of both traits in African Americans. Attenuation of prediction R2 in African Americans (by a factor of 0.18; max R2_height,AA=0.037, P=2e-52, P_training<0.01, 10,592 SNPs; max R2_BMI,AA=0.015, P=7e-25, P_training<1, 363,370 SNPs) is likely due to a combination of population-specific allele frequencies, LD patterns, and causal variants. The results are consistent with a smaller proportion of SNPs genome-wide contributing to variation in height compared with BMI, and with BMI causal variants on average being more poorly tagged and population-specific (and therefore lower-frequency) than height causal variants. Our recent method for inferring common-SNP heritability as well as the allele frequency/effect size distribution from polygenic prediction results (Stahl et al., Nature Genetics, 2012; analysis in progress), promises to directly elucidate these possible explanations.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2211W

Can "asthma genes" predict asthma? *J.A. Curtin, D. Belgrave, J. Hankinson, A. Custovic, A. Simpson.* Translational Medicine, The University of manchester, Manchester, United Kingdom.

Background: Large scale genome wide association studies (GWAS) of asthma have yielded many significant associations between single nucleotide polymorphisms (SNPs) and asthma. We investigated whether these genetic factors predict the development of asthma in a population-based birth cohort (Manchester Asthma and Allergy Study). **Methods:** We reviewed all published reports of asthma GWAS and identified 94 SNPs for which information on the risk allele and the odds ratios was available. We genotyped DNA from 880 children using Illumina 610 quad chips and Sequenom. Asthma at age 11 years was defined if all of the following criteria were met: (1) physician-diagnosed asthma; (2) asthma symptoms (wheeze, cough, or both) within the previous 12 months; and (3) current use of anti-asthma medication (n=105); children with none of the above criteria were assigned as controls (n=513). Information on maternal asthma was ascertained during pregnancy. Current wheeze at ages 3 and 11 years was defined as wheezing in the previous 12 months. Atopic sensitisation at age 3 years was defined as a positive skin test to at least one of the common inhalant/food allergens. We generated receiver-operating-characteristic (ROC) curves to determine the ability of genetic factors, atopy, clinical and family history to predict asthma at age 11. **Results:** Using all 94 SNPs, we were able to explain 1% of the variation in asthma with an Area Under the Curve (AUC) of 0.568. When we assessed if the individual SNPs were associated with asthma ($p<=0.05$), we found that 19 SNPs met these criteria. This reduced set of genetic variants explained 4.8% of the variation in phenotype (AUC: 0.642). Maternal asthma explained 0.7% of the variation (AUC: 0.539). In contrast current wheeze at age 3, sensitisation at age 3 and current wheeze at age 11 years each explained 13.6%, 11.2% & 72.0% of the variation respectively, (AUC: 0.697, 0.679 & 0.959 respectively). **Conclusion:** We found that 94 SNPs associated with asthma in published GWAS have limited ability to predict the disease in an unselected birth cohort representative of the general population. Although a subset of these SNPs (for which we confirmed significant association with asthma in our study) improved the predictive ability, these "asthma genes" explained only a small proportion of the variation of asthma (<5%). In contrast, current wheeze and sensitisation at age 3 years were much better predictors of subsequent asthma.

2212T

What is the total SNP-associated heritability for alcohol dependence? *N.G. Martin[1], G. Zhu[1], P.A. Lind[1], A.C. Heath[2], P.A.F. Madden[2], M.L. Pergadia[2], G.W. Montgomery[1], J.B. Whitfield[1].* 1) Genetic Epidemiology, Queensland Institute of Medical Research, Brisbane, Australia; 2) Department of Psychiatry, Washington University School of Medicine, St Louis, MO, USA.

Background. Much has been written about the so-called "missing heritability" for complex traits. Nowhere is this more pertinent than for alcohol and nicotine dependence (AD, ND) for which there are estimates of heritability of up to 65% from twin studies, yet few causal variants have been replicated from GWAS studies, despite large sample sizes, suggesting that individual effect sizes of SNPs must be very small. Recently new statistical genetic techniques have been developed which allow estimation of the total variance associated with all SNPs on a GWAS chip, but this has yet to be applied to AD and ND. Methods. The current analysis is based on AD and ND symptom count data from over 8000 participants in our population-based twin-family studies who have used either alcohol or cigarettes at some stage of their lives. They were individually genotyped with Illumina 370K or 660K chips and 7.034M genotypes were imputed from HapMap 3 and 1000-Genomes data. The GCTA program of Yang, Visscher et al is used first to detect the degree of relatedness between apparently unrelated subjects, based on a set of about 300,000 SNPs pruned for LD. Phenotypic similarity is then regressed on IBS sharing for all possible relative pairs to estimate the total amount of variance due to SNPs on the chip. Results. Based on GCTA analysis for other complex traits we expect to find SNP associated variance accounting for about half the heritability estimated from conventional genetic epidemiology designs. However, these estimates are highly sensitive to population stratification so great care will be taken to remove all traces of population stratification during the analysis. Conclusions. The gap between the SNP-associated variance estimated by GCTA and twin and family estimates of heritability is most likely due to several factors. First, the tag SNPs on the chip are not in perfect LD with the causal SNPs; for other traits, simulation has shown that correcting for imperfect LD raises the SNP "heritability" by about 10%. Another major factor is that commercial chips only interrogate common SNPs so large effects of rare SNPs are simply not captured. Reasonable estimates from simulations suggest that this could account for another 20% of variance. Finally, we recognize that there are large sections of the genome containing highly repetitive DNA which are very poorly tagged by current chips, and where substantial proportions of genetic variance may be hidden.

2213F

A knockout mouse model for canine juvenile epilepsy. *E.M. Nevalainen[1,2], E. Seppala[1,2], H. Lohi[1,2].* 1) Veterinary Biosciences, University of Helsinki, Helsinki, Finland; 2) Folkhalsan Research Unit, Helsinki, Finland.

Epilepsy afflicts more than 1% of the human population. Epilepsies are the most common neurological diseases in two to 10 years old children. Epilepsy is also the most common neurological disorder in dogs. Our group has recently described benign focal juvenile epilepsy in the Lagotto Romagnolo breed (Jokinen et al, J Vet Intern Med 2007). These dogs manifest an epileptic tremor syndrome with similarities to human idiopathic childhood epilepsies. The onset of the epilepsy in lagottos is at the age of five to nine weeks and affected puppies present focal seizures with interictal EEG activities. The disease resolves without treatment by the age of four months. We have identified the genetic cause in the LGI2 gene (Seppälä et al. PLoS Genet 2011). LGI2 is a paralog of the human epilepsy gene, LGI1. We have shown that LGI2, like LGI1, is neuronally secreted and interacts with the members of the ADAM family of neuronal receptors. LGI-ADAM complexes are suggested to function in synaptic remodeling and participate in the development of neuromorphology. Based on the expression pattern, Lgi2 was proposed to have a role in the maintenance of the electric stability of the neural networks in the early phases of the developing brains. We have generated a LGI2-deficient mouse model (using KOMP) to further elucidate the biological roles of LGI2 complexes in the normal and epileptic developing brains. Lgi2-deficient mice develop normally to adulthood, are fertile, and the gross tissue anatomy and morphology is similar to that of wild-type littermates. Different tissues have been collected and the expression of several genes and proteins will be compared in knockout and wild-type mice at varying ages using routine quantitative and immunohistochemical protocols. The homozygous mice have been monitored to follow possible epileptic activities and seizures. We will study the electrophysiological properties of the neuronal cells of the knockout and wild-type mice. Canine epilepsy models are unique in representing naturally-occurring spontaneous models of human epilepsy. The data will shed light on the role of LGI2 in multicellular organisms, and elucidate its role in the development of epilepsy in juvenile dogs and humans.

2214W

ER stress response genes and their natural variation, discovered using the *Drosophila* model system. *C. Y. Chow, M. F. Wolfner, A. G. Clark.* Dept Molec Biol & Gen, Cornell Univ, Ithaca, NY.

The endoplasmic reticulum (ER) is responsible for synthesis and maturation of many proteins essential for cellular function. ER dysfunction can have devastating consequences when misfolded proteins accumulate in the ER, resulting in "ER stress" associated with pathogenesis of diseases from diabetes to neurodegeneration. To understand the contribution of ER stress resistance to disease, it is important to determine the extent of natural variation in ER stress resistance. Drosophila provides an excellent new model for studying variation and for rapid investigation of novel ER stress genes. We characterized variation in ER stress resistance in the Drosophila Genetic Reference Panel (DGRP), a collection of 192 inbred wild-derived lines whose whole-genome sequences are available. We compared their relative survival on food supplemented with the ER stress-inducing drug tunicamycin. Molecular markers of ER stress, such as Xbp1 splicing, indicate that tunicamycin induces ER stress in the fly. Mortality rates, as measured by the Cox proportional hazard ratio, varied by more than 100 fold, indicating extensive genetic variation in ER stress resistance among the DGRP lines. To understand the genetic variation underlying differential survival to ER stress, we used Agilent microarrays to compare gene expression between lines that were either highly sensitive vs. resistant to tunicamycin-induced ER stress. We found that genes involved in cellular responses to ER stress, including genes previously implicated in human disease, varied in response across the genotypes. Some of these canonical ER stress genes showed up to 10-fold difference in ER stress-induced expression. We also discovered additional genes, not previously implicated in ER stress response, that correlated with ER stress and its variation in our lines. Finally, we observed that transcriptome differences observed prior to ER stress induction may be predictive of resistance or sensitivity to ER stress. Our results indicate that natural variation in Drosophila provides a rich new resource for identifying novel ER stress genes and their functions, further elucidating this cellular pathway critical to many human diseases.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2215T

Physical and genetic interactions between BBS genes and CEP290. *Y. Zhang[1], K. Bugge[1,2], C.C. Searby[1,2], R.F. Mullins[1], S. Seo[1], V.C. Sheffield[1,2].* 1) Dept Pediatrics, Univ Iowa, Iowa City, IA; 2) Howard Hughes Medical Institute, USA.

Bardet-Biedl syndrome (BBS) is a heterogenous autosomal recessive inherited disorders with clinical features that include retinal degeneration, obesity and developmental anomalies. At least 17 BBS genes have been reported. Seven BBS proteins form a molecular complex know as the BBSome, and three additional BBS proteins form a second complex known as the BBS chaperone complex based on the homology of these proteins to chaperonins. The chaperone complex is required for BBSome assembly. However, the interaction or function of other BBS proteins remains to be investigated. Recent studies have shown that mutation of a novel centrosomal protein, Cep290, also called NPHP6 and BBS14, results in BBS and other ciliopathies, but the exact mechanism remains unknown. In this study, we examined the BBS-related function of Cep290 and investigated interactions between Cep290 and other BBS proteins. We demonstrate that Cep290 interacts with BBS4 and BBS9, which are components of the BBSome. In addition, we found that the BBSome is required for correct localization of Cep290 in RPE Cells. Furthermore, Cep290 is mislocalized in retina of BBS1M390R/M390R mice, which are homozygous for a methionine to arginine substitution at codon 390 of Bbs1, another component of the BBSome. Moreover, the body weight and leptin level of BBS4 +/- Cep290 +/- (double heterozygous) mice are higher than those of single heterozygous mice. BBS4 +/- Cep290 +/- mice also show leptin resistance compared to single heterozygotes. These data strongly indicate that there is both physical and genetic interactions between Cep290 and the BBSome.

2216F

***Vacuolated lens (vl)*-associated Neural Tube Defects (NTDs) are regulated on a multigenic basis by *Cdx1* and Retinoic Acid (RA) signaling pathway.** *B. Li[1,2], P. Matteson[1], M. Ababon[1], A.Q. Nato[3], T. Matise[3], V. Nanda[1,4], J. Millonig[1,2,3].* 1) Center for Advanced Biotechnology and Medicine, Piscataway, NJ; 2) Department of Neuroscience and Cell Biology, UMDNJ-RWJMS, Piscataway, NJ; 3) Department of Genetics, Rutgers University, Piscataway, NJ; 4) Department of Biochemistry, UMDNJ-RWJMS, Piscataway, NJ.

To understand the molecular genetic basis of neurulation, we are studying the *vacuolated lens* (*vl*) mutation that arose spontaneously on the C3H/HeSnJ background. The mutation affects apposition/fusion of the neural folds, which leads to NTDs, embryonic lethality and abnormal pigmentation. The *vl* phenotypes are caused by a mutation in an orphan G-protein coupled receptor (*Gpr161*), that is expressed in the neural folds. Crossing *vl*^{C3H} mice to the MOLF/EiJ strain rescues the *vl* defects. QTL analysis mapped three modifiers (*Modvl*: **Mod**ifier of ***vl***). One QTL being studied is *Modvl5* (LOD= 5.0; Chr18). A *Modvl5*^{MOLF} congenic strain was generated to study whether this locus is sufficient to rescue the *vl* phenotypes. Analyses on congenic embryos determined that *Modvl5*$^{C/M}$ rescues the embryonic lethality (P= 0.004) and partially rescues the NTDs. Using our bioinformatics pipeline, we determined that the transcription factor,*Cdx1*, is the only gene within the *Modvl5* 95% Confidence Interval (CI) co-expressed with *Gpr161* in the lateral neural folds. Re-sequencing *Cdx1* identified a poly-glutamine polymorphism (5Q-C3H, 7Q-MOLF) predicted to affect secondary and tertiary protein structure. Additional mating to Balb/cJ (5Q allele) does not rescue the *vl* defects, while mating to PWK/PhJ (7Q allele) rescues the phenotypes (p=0.0001). Electrophoretic Mobility Shift Assay (EMSA) demonstrated that 7Q binds better than 5Q allele (p<0.001) to the promoters of Retinoic Acid signaling genes. QRT-PCR indicated that the expression of these genes (*Aldh1a2* and *Cyp26a1*) is down-regulated by *vl* and returns to nearly wild-type levels in *Modvl5*$^{C/M}$. To directly measure RA signaling intensity, E8.5 embryos were cocultured with the RA-responsive reporter cell line F9-RARE-lacZ. Preliminary data suggests that RA signaling is down in *vl* and rescued by *Modvl5*$^{C/M}$(p=0.01). Consistent with this possibility, preliminary analysis indicates that the NTDs phenotype is partially rescued by injecting all-trans RA into E9.5 *vl*^{C3H} embryos. These results establish *vl*as one of the first multigenic mouse mutant models of NTDs and demonstrate that *Cdx1*^{MOLF} bypasses the mutation by up-regulating the downstream RA signaling pathway.

2217W

HLA class-I supertypes are associated with specific M. tuberculosis strain infections. *M. Salie[1], L. vd Merwe[2], M. Moller[1], M. Martin[3, 4], X. Gao[3, 4], R. Warren[1], M. Carrington[3, 4], E. Hoal[1].* 1) Biomedical Science, Stellenbosch University, Tygerberg, Western Cape, South Africa; 2) Biostatistics Unit, Medical Research Council, Tygerberg, Western Cape, South Africa; 3) Ragon Institute of MGH, MIT and Harvard, Boston, Massachusetts, USA; 4) Laboratory of Experimental Immunology, Frederick, Maryland, USA.

Tuberculosis (TB) is the second highest cause of mortality due to an infectious disease. Development of active clinical TB is not solely due to the pathogen, Mycobacterium tuberculosis, but environmental and host genetic factors also contribute. The potential role of interactions between the host and bacterial genotype in influencing disease outcome has been proposed and a few studies have indicated the adaptation of M. tuberculosis strains to specific populations. The Human Leukocyte Antigen (HLA) class-I molecules are involved in various immunological processes, including antigen presentation to T cells. The HLA class-I genes; HLA-A, -B and -C were genotyped in 300 TB cases by sequencing using the ABI-3130XL DNA analyzer. M. tuberculosis strain identification was done by means of spoligotyping. Haplotypes were inferred using the EM algorithm. Logistic regression models were used to determine the likelihood of cases having a specific strain, adjusting for age and sex. Significant associations between host genotype and bacterial strain were identified for genotypes, haplotypes and the additive allelic model. For HLA-A, the odds ratio (OR) of having LAM, was 3.89 and 6.33 for the A01/A02 and A03/undefined genotypes, respectively, compared to A01/A01. For HLA-B, the significant ORs of having Beijing were very large and imprecise for B07/B08 (OR=19.6), B01/B44 (OR=10.4) and B08/B62 (OR=25.4) compared to B07/B07. For HLA-C, the ORs of having Beijing were 3.61, 4.39 and 4.46 for C1/C2, C2/C2 and C2/undefined, respectively, compared to C1/C1. Three haplotypes were also found to be significantly associated with infection with a Beijing strain (P=0.0002). Several associations were also identified for HLA alleles and M. tuberculosis strains. We show that HLA alleles could be involved in M. tuberculosis strain susceptibility. Both pathogen and host genetics need to be taken into consideration when studying TB disease development.

2218T

Peripheral blood cell counts are heritable in a large, unselected rhesus macaque pedigree. *A. Vinson[1,2], A.D. Mitchell[1,2], D. Toffey[1].* 1) Oregon National Primate Research Center, Oregon Health & Science University, Beaverton, OR, 97006; 2) Dept. of Molecular and Medical Genetics, Oregon Health & Science University, Portland, OR, 97239.

Circulating blood cell counts are associated with risk for cardiovascular and other complex human disease, and a better understanding of the genes that influence their spontaneous variation will ultimately inform personalized therapies. The goals of this study were to assess heritability of blood cell counts in rhesus macaques, a widely-used non-human primate model for complex human disease. Large, managed populations of rhesus macaques are a powerful resource for genetic analysis of quantitative traits like blood cell counts, offering substantial genetic similarity to humans, available pedigree information, large numbers of informative relative classes, and a controlled environment that will enhance genetic signal over noise. Based on these advantages, we have characterized an unselected pedigree of Indian-origin rhesus macaques for quantitative genetic analysis and gene mapping of complex traits. This pedigree was developed to optimize several criteria: animals must have parentage confirmed by molecular genotyping, belong to a minimum 3-generation lineage, be available for sampling, not have common ancestors outside the pedigree, and form a single, extended pedigree structure. The resulting pedigree contains 1,289 macaques spanning 6 generations, with 800 females and 489 males. Using a maximum likelihood-based variance decomposition approach with clinical data collected on 995 animals, we found significant heritability for white blood cell, red blood cell, and platelet counts, as well as measures of mean cell volume, mean cell hemoglobin, mean cell hemoglobin concentration, hematocrit, and hemoglobin (h^2 range 0.144—0.453, P-value range 1.14×10^{-3} —2.61×10^{-21}). These results indicate that genetic contributions to these important biomedical phenotypes can be detected and measured in this pedigree, and support the need for genome-wide analysis of these and related risk factors for human disease in this population.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2219F

A copy number variant at the KIT ligand locus confers risk for canine squamous cell carcinoma of the digit. *E.A. Ostrander[1], B. Decker[1], E. Carlins[1], B. VonHoldt[2], G. Carpintero-Ramirez[1], H.G. Parker[1], R.K. Wayne[2], D.M. Karyadi[1].* 1) Cancer Genetics Branch, NHGRI/NIH, Bethesda, MD; 2) Dept of Ecology and Evolutionary Biology, UCLA, Los Angeles CA.

The domestic dog is a robust model for studying the genetics of disease susceptibility as many breeds harbor an elevated risk for certain cancers, despite the simple population structure common to all breeds. One example of breed-specific disease propensity is squamous cell carcinoma of the digit (SCCoD), a locally aggressive cancer that frequently causes lytic bone lesions and multiple toe recurrence. SCCoD is uncommon in most dog breeds, but is prevalent in Standard Poodles. Intriguingly, Poodles with dark coat color are at high risk for SCCoD, whereas Poodles of light coat colors are entirely unaffected, suggesting that interactions between multiple pathways are necessary for oncogenesis. We performed a genome-wide association study on Standard Poodle SCCoD cases compared to unrelated black Poodle controls. Allelic association was calculated with the single-locus chi-squared significance test. The six most strongly associated SNPs (P_{raw}= $5.62x10^{-5}$-$1.20x10^{-7}$) were contiguous markers on canine chromosome 15 and all were statistically significant at the chromosome-wide level ($P_{empirical}$= $0.0299\text{-}7.00x10^{-5}$). Comparison of Poodle cases to other at-risk breeds initially refined the locus to 144.9Kb, a region that harbors 128 variants segregating with risk. The initial risk-associated haplotype is present in 92.9% cases, with 55% of cases are homozygous for the risk allele. While 52.8% of controls are heterozygotes, none are homozygous for the risk haplotype. The region lies upstream from the KIT Ligand (KITLG) gene. Fine mapping and sequencing reduced the locus to a minimal region of 29.7Kb. A dosage dependent effect of a CNV within the region is strongly associated with the disease phenotype in Standard Poodles. Allele-specific expression assays are ongoing to determine the relationship between the CNV and KITLG expression. Additional investigation highlighted a compensatory nonsense mutation in light colored standard poodles that explains their lack of disease, even if they carry the risk allele. The presence of disease is not due to coat color pigments, i.e. light colored Briards lack the compensatory mutation and get the disease if they carry the risk allele, similar to black Briards. This work highlights the utility of the canine system for disentangling the genetics of multigene cancer disorders. Our ongoing studies should highlight other genes contributing in the disease and, by extension, what role these genes play in human cancers.

2220W

Large-scale Transcriptional and Epigenetic Changes Bridge Maternal Nutritional Imbalance and Metabolic Syndrome. *X. Yang[1,2], L. Peterson[3], T. Fare[1], J. Zhu[1,4], R. Kleinhanz[1], C. Suver[1,5], A.M. Cumiskey[3], R.L Rosa[3], A. Vosatka[3], M. Skinner[6], E.E Schadt[1,4], P.Y. Lum[1,7].* 1) Merck Research Laboratory, Merck & Co, Inc. Seattle, WA; 2) Integrative Biology and Physiology, UCLA, Los Angeles, CA; 3) Department of Pharmacology, Merck & Co, Inc. Rahway, NJ; 4) Department of Genetics and Genomic Sciences, Institute of Genomics and Multiscale Biology, Mount Sinai School of Medicine, New York, NY; 5) Sage Binetworks, Seattle, WA; 6) Center for Reproductive Biology, School of Biological Sciences, Washington State University, Pullman, WA; 7) Ayasdi Inc, Palo Alto, CA.

Objective: The increasing prevalence of Metabolic Syndrome (MetSyn), a combination of cardiovascular disease risk factors with poorly understood etiology, has resulted in increasing morbidity and mortality worldwide. Among the potential causes for the predisposition to MetSyn is early nutrition during gestation and lactation. The aim of our study is to understand the molecular mechanisms underlying the phenotypic effects of early nutritional perturbation. Research Design and Methods: We characterized the alterations in the transcriptome and epigenome of mouse offspring exposed in utero to a high fat high sucrose diet, a protein-restricted diet, or a standard Chow diet using genome-wide transcriptional and DNA methylation profiling of two key metabolic tissues - hypothalamus and liver - at 15 weeks of age. Results: We found that poorly balanced maternal diets (the high fat high sucrose diet and the low protein diet) induced alterations in the expression of thousands of genes involved in biological processes relevant to MetSyn in the hypothalamus and liver tissues. We also identified significant changes in DNA methylation status of hundreds of genetic loci along with a genome-wide hypomethylation shift in the liver tissue of the mouse offspring. Conclusions: These findings provide molecular insights into the influence of early nutrition on metabolic disease susceptibility and may help direct therapeutic interventions to prevent and treat MetSyn.

2221T

Gene Expression Profiling in Mild Traumatic Brain Injury, Post-traumatic Stress Disorder, Sleep Disorders and Depression among Redeployed Military Personnel. *H. Lee[1], D. Wang[1], R. Pierce[2], T. Baxter[2], R. Dionne[1], H. Kim[1], V. Mysliwiec[2], T. Barr[1,3].* 1) National Institute of Nursing Research (NINR), National Institutes of Health (NIH), Bethesda, MD; 2) Madigan Healthcare System, Tocoma, WA; 3) West Virginia University, Morgantown, WV.

Objectives: Post-traumatic Stress Disorder (PTSD) and mild traumatic brain injury (mTBI) remain leading causes of disability among post-deployed military service members and both are frequently comorbid with sleep disorders or depression. Our study aims to identify biomarkers for PTSD, mTBI, sleep disorders and depression through peripheral blood transcriptome profiling in a redeployed military population for developing more sensitive diagnostic criteria. Methods: A total of 106 military service members (>18 years old) within 18 months post-deployment were enrolled (99% male; 68% European American) and were assessed for the 4 clinical factors—PTSD, mTBI, sleep disorders and depression. Total RNA was extracted from blood samples followed by amplification and hybridization to Affymetrix GeneChip Human Gene U133 Plus 2.0 Arrays. In Partek Genomics Suite Beta 6.6, normalized data (log base 2 for each probeset) were subjected to quality control procedures. For gene filtering, probesets with less than 5% of coefficient of variation across the entire sample were filtered out. Differential gene expression was then compared between groups. Considering potential influence of ethnicity and gender on gene expression, analysis was initially limited to 57 European American male subjects.Results: We first performed a separate bivariate analysis for each of the 4 clinical factors, and then compared differentially expressed genes found in the bivariate analyses with ones from multivariate analysis controlling for each other clinical factors. For PTSD, different genes were found from bivariate (*CEACAM8*, *DEFA4*, *CAMP*, and *MMP8*) and from multivariate (*SDHD*,*SLPI*, *RHCE/RHD*, *BEX1*, and *CLCN4*) analyses. However, there were same genes expressed in both bivariate and multivariate analyses, for blast neurotrauma (*MYSM1*, *SLAMF7*, *CBLL1*,*SMAD7*, and *TTC22*), depression (*SLC11A1* and *PI4K2B*), or sleep disturbances (*FAM3B*). Validation study using qPCR follows to confirm these genes. Conclusions: Our interim findings support the use of peripheral blood transcriptome profiling as a potential biomarker which can help diagnose military personnel with these illnesses. Further investigation with a large cohort as well as more objective measurements is necessary to validate the diagnostic accuracy of these biomarkers. Acknowledgements: The study is supported by the Center for Neuroscience and Regenerative Medicine (CNRM) at the Uniformed Services University of the Health Sciences (USUHS).

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2222F

Genetics of obesity across development in a model system. *C.A. Schmitt[1], S. Service[1], R.M. Cantor[2], A.J. Jasinska[1], M.J. Jorgensen[3], J.R. Kaplan[3], N.B. Freimer[1].* 1) Center for Neurobehavioral Genetics, University of California, Los Angeles, Los Angeles, CA; 2) Department of Human Genetics, University of California, Los Angeles, Los Angeles, CA; 3) Department of Pathology, Section on Comparative Medicine, Wake Forest School of Medicine, Winston-Salem, NC.

Obesity is a problem of epidemic proportions worldwide, and has severe negative impacts on public health. Obesity arises from a complex interaction of genetic predisposition and environment that can accumulate throughout life, but increasing evidence points to the importance of biological processes occurring early in development in the manifestations of adult disease. In humans, the elucidation of the genetic underpinnings of obesity and concomitant ontogenetic patterns is complicated by the ubiquity of diets extremely high in fats and simple carbohydrates, and the difficulty in assessing what study subjects are actually eating and in what quantities. This research investigates a baseline genetic predisposition to obesity during development in a genetically well-characterized model system under a controlled diet and environment: the African green monkey or vervet (*Chlorocebus aethiops sabaeus*) in the Vervet Research Colony (VRC) at Wake Forest School of Medicine.

We used basic measures taken thrice yearly on body size and composition (such as BMI and waist circumference) in a population of 657 monkeys measured from 2008 through 2012 to search for heritable patterns in obesity. We found that 118 (29% of adults) presented with signs of abdominal obesity at least once, with obesity defined as having an adult waist circumference above 40.5 cm; and 43 (10% of adults) were chronically obese (obese during at least three successive measurements). Loess regressions of mean waist circumference reveal stark differences in growth trajectories between the chronically obese and non-obese colony members: whereas non-obese colony members' waist circumference growth levels off just after attaining full adult size, obese colony members appear to maintain a pubertal growth slope up to a decade after puberty has passed. Analyses of estimated additive genetic heritability using SOLAR show that BMI, waist circumference, and chronic obesity are all highly heritable (BMI: $h^2 = 0.44$, $p < 0.0001$; waist circumference: $h^2 = 0.37$, $p < 0.02$), at levels much higher than those seen in previous studies of NHP populations, with significant covariates represented by age and sex. This preliminary study suggests that a more comprehensive examination of growth trajectories may elucidate previously unknown patterns in the development of obesity. Such novel patterns may be used to assess early obesity risks and promote the discovery of novel biomedical interventions.

2223W

Modeling 3D facial shape in relation to sex and genomic ancestry estimated from DNA. *M.D. Shriver[1], P. Claes[2], D.K. Liberton[1], K.M. Rosana[1], E.E. Quillen[1], L.N. Pearson[1], B. McEvoy[3], M. Bauchet[1], H. Tang[4], G. Barsh[4,5], D.M. Absher[5], D.A. Puts[1], J. Rocha[6,7], S. Beleza[4,8], R.W. Pereira[9], J.K. Wagner[10], J. Boster[11].* 1) Dept Anthropolgy, Penn State Univ, University Park, PA; 2) Katholic University of Leuven, KU Leuven, Belgium; 3) Smurfit Institute of Genetics, Dublin, Ireland; 4) Department of Genetics, Stanford University, Palo Alto, CA; 5) HudsonAlpha Institute for Biotechnology, Huntsville, AL; 6) CIBIO: Centro de Investigação em Biodiversidade e Recursos Genéticos, Universidade do Porto, Portugal; 7) Departamento de Biologia, Faculdade de Ciências, Universidade do Porto, Portugal; 8) IPATIMUP: Instituto de Patologia e Imunologia Molecular da Universidade do Porto, Portugal; 9) Programa de Pós-Graduação em Ciências Genômicas e Biotecnologia, Universidade Católica de Brasília, Brasilia, Brasil; 10) Center for the Integration of Genetic Healthcare Technologies, University of Pennsylvania, Philadelphia, PA; 11) Department of Anthropology, University of Connecticut, Storrs, CT.

Human facial diversity is substantial, complex, and largely scientifically unexplained. We use spatially dense quasi-landmarks and partial least squares regression (PLSR) to measure face shapes in population samples with mixed West African and European ancestry from three locations (United States, N=154; Brazil, N=191; and Cape Verde, N=247). We modeled facial variation as shape regressed on ancestry direction (SRAD) and on sex direction (SRSD). We then examined the extent to which distance along these directions (dSRAD and dSRSD) predicted genomic ancestry, sex, and observer judgments of ancestry and sex. We used alternate marker panels and noise injection (NI) experiments to investigate the robustness of facial modeling and find that accurate shape regressions are possible with imprecise estimates of sex and genomic ancestry. The independent variables estimating both sex and ancestry can be reduced in precision by as much as 50% (i.e., r2=0.5 between the original variable and the NI variable) and still show correlation coefficients of about 0.95 between the original and NI dSRAD and dSRSDs. The PLSR approach is remarkably efficient in accessing the information on both genomic ancestry and sex to establish the directions in face space and measurements of the relative positions of individual faces with respect to these directions. Judgments of ancestry and sex shift sharply at certain morphological thresholds, reflecting categorical perception. The strong correlations that we observe between judged sex and judged femininity and dSRSD, and between judged proportional ancestry and judged categorical ancestry dSRAD formally support the idea that we are generating valid and meaningful estimates in the context of human perceptions. Facial models combining the two independent variables - genomic ancestry and sex - allow systematic transformations of individual faces along these directions, establish a framework for mapping genes affecting facial features, and provide the first tool for estimating the appearance of a person's face from genetic markers.

2224T

Endogenous DUX4 induces myotube-specific apoptosis in FSHD Muscular Dystrophy and is regulated by Wnt/β-catenin signaling. *G.J. Block, D. Narayanan, A.M. Amell, L.M. Petek, D.G. Miller.* Pediatrics, University of Washington, Seattle, WA.

The gene encoding Double Homeobox Protein 4 (DUX4) is located within each unit of the D4Z4 macrosatellite repeat and is transcriptionally silenced in most cells. In patients with facioscapulohumeral muscular dystrophy (FSHD), D4Z4 loses markers of heterochromatin and DUX4 is transcribed and translated. Previous studies showed that DUX4 protein is rare in muscle cultures from FSHD patients and the cultures do not display defects in differentiation potential or susceptibility to stress. Although over-expression of exogenous DUX4 induces apoptosis in multiple cell and tissue types, there is no evidence that DUX4 induces cell death at physiologic levels, or explanation as to why skeletal muscle is preferentially affected in the disease. Here we describe a novel myotube differentiation protocol to increase the efficiency of myotube formation and detection of DUX4 protein. We show that muscle cultures from FSHD patients only express DUX4 in differentiated myotubes, where the DUX4(+) nuclei then initiate apoptosis. Knockdown of DUX4 by siRNA abolished myotube apoptosis. Finally we demonstrate the utility of our model by modulating DUX4 expression through activation and suppression the Wnt/Beta-catenin signaling pathway. Our data provide the first evidence that DUX4 induces apoptosis of myotubes at physiological levels, and validates a platform that can be used to understand the regulation of DUX4-induced apoptosis in FSHD.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2225F

Development of a Novel *Col3a1* Transgenic Mouse Model for Vascular Ehlers-Danlos Syndrome. *F. Malfait[1], Y. Kang[1], S. Symoens[1], P. Coucke[1], M. Renard[1], B. Guillemyn[1], J. De Backer[1], F. Branco-Madeira[2], W. Toussaint[2], L. Vanhoutte[2], S. Janssens[2], B. Descamps[3], C. Vanhove[3], P. Segers[4], B. Lambrecht[2], A. De Paepe[1]*. 1) Center for Medical Genetics, Ghent University Hospital, Ghent, Belgium; 2) Department of Molecular Biomedical Research, VIB, Ghent, Belgium; 3) Department IBITech-MEDISIP-INFINITY-GROUP-ID Consortium, Ghent University, Ghent, Belgium; 4) IBiTech-bioMMeda, Ghent University, Ghent, Belgium.

Type III collagen (*COL3A1*) mutations cause vascular Ehlers-Danlos syndrome (vEDS), a life-threatening disorder, characterized by thin, translucent skin, abnormal wound healing, easy bruising and increased risk for arterial, intestinal and uterine rupture. The majority of *COL3A1* mutations are glycine-substitutions in the triple-helical domain, resulting in misfolded collagen trimers. In less than 5% of the patients *COL3A1* haploinsufficiency mutations have been reported.

We generated a transgenic mouse model in which mutant mouse *Col3a1* gene with a p.Gly183Ser mutation was introduced into mouse genome by BAC technology. This mutation has been documented in several patients with typical vEDS. This transgenic approach allows near-normal expression of *Col3a1* as the construct contains the endogenous promoter region, regulatory elements and untranslated regions. Currently, five different *Col3a1* transgenic mice lines are available, each harboring different copies of the mutant *Col3a1* transgene. Initial results on the transgenic line harboring the highest transgene copy number show that these mice have decreased pre-weaning survival in comparison with WT-mice (p-value < 0.05). Male transgenic mice are smaller at 12 and 16 weeks of age (12 and 16.2%, respectively) when compared to strain-, age-, and sex-matched WT controls. Moreover, the transgenic males recapitulate the human phenotype and display a fragile, thin skin. At two months they spontaneously develop open wounds starting in the shoulder area and penetrating skin and subdermal tissue. These skin lesions evolve fast and animals need to be euthanized rapidly. Preliminary vascular corrosion-casting experiments on the euthanized mice are complicated by rupture of arterial vessels, suggesting increased vascular fragility. Additionally, post-mortem fluid-induced rupture experiments reveal friable arteries which rupture at a lower pressure, compared to WT mice.

Although further experiments are necessary to fully delineate the phenotype, this mouse model seems to mimic the characteristics of human vEDS.

2226W

An integrative genetic and metabolomic analysis suggests altered phosphatidylcholine metabolism in asthma. *J.S. Ried[1], H. Baurecht[2], F. Stückler[3], J. Krumsiek[3], J. Adamski[4], C. Gieger[1], J. Heinrich[5], M. Kabesch[6], A. Peters[7, 8], E. Rodriguez[2], K. Suhre[3, 9, 10], R. Wang-Sattler[8], K. Strauch[1, 11], H.-E. Wichmann[5, 12, 13], F.J. Theis[3, 14], T. Illig[4, 15], S. Weidinger[2]*. 1) Institute of Genetic Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany; 2) Department of Dermatology, Allergology, and Venerology, University Hospital Schleswig-Holstein, Campus Kiel, Kiel, Germany; 3) Institute of Bioinformatics and Systems Biology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany; 4) Institute of Experimental Genetic, Helmholtz Zentrum München - German Research Center for Environmental Health, Germany; 5) Institute of Epidemiology I, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany; 6) Department of Pediatric Pneumology, Allergy and Neonatology, Hannover Medical School, Hannover, Germany; 7) Institute of Epidemiology II, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany; 8) Research Unit of Molecular Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany; 9) Faculty of Biology, Ludwig-Maximilians-Universität, Munich, Germany; 10) Department of Physiology and Biophysics, Weill Cornell Medical College, Doha, Qatar; 11) Institute of Medical Informatics, Biometry and Epidemiology, Chair of Genetic Epidemiology, Ludwig-Maximilians-Universität, Munich, Germany; 12) Institute of Medical Informatics, Biometry and Epidemiology, Chair of Epidemiology, Ludwig-Maximilians-Universität, Munich, Germany; 13) Klinikum Grosshadern, Munich, Germany; 14) Department of Mathematics, Technische Universität München, Munich, Germany; 15) Hannover Unified Biobank, Hannover Medical School, Hannover, Germany.

Background: Genome-wide association studies (GWAS) have identified many risk loci for asthma, but effect sizes are mostly small and in most cases the biological mechanisms are unclear. Targeted metabolomics that is based on the quantitative measurement of a panel of small molecules provide information about a whole range of pathways of intermediary metabolism in a biological sample. Previous studies show that combining genetic and metabolomic information can aid in characterizing genetic association signals with high resolution. Our objective is to investigate the interrelation of current asthma, asthma risk alleles and blood serum metabolites.

Methods: We investigated 151 serum metabolites, quantified by targeted mass spectrometry, in fasting serum of asthmatics and non-asthmatic individuals from the population-based KORA F4 study (n=2,925 participants). For the characterization of differences in metabolomic profile between asthmatics and no asthmatic individuals we performed logistic regression analyses and network reconstruction with Gaussian graphical modeling (GGM). In addition, we analyzed effects of single nucleotide polymorphisms (SNPs) at 40 asthma risk loci on these metabolites.

Results: Two phosphatidylcholines (PC.aa.C42:1 and PC.ae.C42:5) were significantly associated with current asthma (p-value= 8.16×10^{-5}/7.35×10^{-5}, OR= 1.38/1.39). Eight metabolites of different metabolite classes were significantly associated with asthma risk alleles from six loci. Acyl-alkyl-phosphatidylcholines (PC.ae.C42:5 and PC.ae.C44:5) associated with risk alleles in*PSMD3* and *MED24* in turn coincide with asthma risk. The combined effect directions of the SNPs on these two phosphatidylcholines and of phosphatidylcholines on asthma were consistent with known effects of SNPs on asthma risk.

Conclusions: Our study demonstrates the potential of metabolomics to infer asthma-related biomarkers by the identification of deregulated phosphatidylcholines that associate with asthma and asthma risk alleles. This knowledge may be used in the future for better understanding of asthma prediction, diagnosis and treatment.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2227T

Genomic data integration reveals molecular modularity of Parkinson disease. *A. Dumitriu[1, 2], Y. Xia[2], R.H. Myers[1, 2].* 1) Department of Neurology, Boston University School of Medicine, Boston, MA 02118, USA; 2) Bioinformatics Program, Boston University, Boston, MA 02215, USA.

Parkinson disease (PD) is a complex neurodegenerative disorder, with mostly unknown etiology. Several human and model organism genetic studies have been performed for the analysis of PD. While individual studies can provide important knowledge about the biological processes involved in PD, a more complete picture of the disorder can be obtained by combining multiple existing genetic data sources. Complementary to disease-specific data, available functional information can also be considered for the investigation of PD genetics.

We constructed a comprehensive catalogue of PD-relevant genes by integrating data from three different types of genetic studies: whole-genome expression, genome-wide association, and PD animal models. We then showed that PD-relevant gene sets are significantly more interconnected than expected by chance in the network of functional linkages among human genes, and that gene functional information improves the prediction accuracy of PD-relevant genes. Finally, we used the combined genetic data and the functional information to rank a set of canonical pathways for their involvement in PD. Two additional microarray expression studies performed for 1) PD and 2) Hepatitis C Virus (HCV) were used as our positive and negative replication sets, respectively. Only the pathways found to be significant for the combined genetic data set, as well as for the PD microarray, but not for the HCV microarray were considered important. The pathway analysis showed evidence of association with PD for pathways involved in neurotransmission, vesicular transport, signaling, and cell death.

Our results demonstrated the relevance of functional information in the study of PD. We also found that PD candidate genes from expression and association studies are largely different, which suggests that commonly used genetic methods assess distinctive aspects of the disease. Finally, we showed that integrative analyses of PD can benefit from a set of comprehensive and unbiased disease gold-standard genes, which is not currently easily available.

2228F

Pathway analysis of the GENEVA Alcohol Consumption GWAS Meta-Analysis. *O. Harari, A. Goate, GENEVA Alcohol Consumption Working Group.* Psychiatry, Washington University School of Medicine, St. Louis, MO.

The genetic architectures of complex psychiatric traits are likely to be the product of multiple genetic risk factors. Pathway analysis offers a complementary perspective to interpreting GWAS, incorporating repositories of expert knowledge, represented in biological pathways to allow the identification of variants that do not necessarily reach stringent genome-wide significance levels. Our approach to analysis of the GENEVA Alcohol Consumption GWAS meta-analysis, which encompasses 6 studies, included two stages: i) we evaluated each of the studies by itself; and ii) we interrogated the meta-analysis aggregated evidence. We restricted the analysis to subjects with European-American ancestry (EA), including 7,433 females and 7,275 males with harmonized alcohol consumption data (ln-tranformed drinks/week) and genome-wide genotype data (10,329 subjects were genotyped with the Affymetrix 6.0 platform; and 4,379 with different Illumina chips). After verifying that each of the studies included more independently significant SNPs than expected by chance, we performed a pathway analysis employing the ALIGATOR method. We observed that only the analysis of SNPs with p-values < 1.0E-3 identified an excess of Gene Ontology (GO) terms and pathways from the Kyoto Encyclopedia of Genes and Genomes (KEGG). There was not a pathway or term significant (p-value<0.05) for 3 or more studies. When we evaluated the meta-analysis signals, an excess of significantly overrepresented pathways was detected for women but not for men. Variants in alcohol dehydrogenase (ADH) genes modulate drinking levels. Rs1229984 in ADH1B, with a MAF < 5% for EAs, achieves a genome-wide significant level. This SNP is not ascertained in the employed chips and is poorly imputed. Consequently, only SNPs with higher frequency and in high D' could detect this signal in 5 of the GENEVA studies with moderate p-values (between 5.0E-3 and 0.05). We explored if 8 GO terms and 6 KEGG pathways including ADH1B were enriched with SNPs showing this significance level, and found 3 studies with positive results. However the pathways had moderate significance that did not surpass a Bonferroni multiple-correction adjustment. We observed similarly moderate p-values for the GENEVA meta-analysis. Alcohol consumption can be variously measured and it is likely that phenotypic heterogeneity across samples influenced our finding. Larger samples with systematically ascertained alcohol consumption data are necessary.

2229W

Pathway-Based Meta Analysis of Ulcerative Colitis Genome-Wide Association Studies. *C. Kao[1], Z. Wei[2], J. Li[1], W. Wang[2], J. Glessner[1], C. Cardinale[1], J. Bradfield[1], E. Frackelton[1], C. Kim[1], F. Mentch[1], H. Qui[1], S. Grant[1,3], R. Baldassano[1,3], H. Hakonarson[1,3], International IBD Genetics Consortium.* 1) Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, PA; 2) Department of Computer Science, New Jersey Institute of Technology, Newark, NJ; 3) Department of Pediatrics, University of Pennsylvania School of Medicine, Philadelphia, PA.

Inflammatory Bowel Disease (IBD) is a complex inflammatory condition, where the etiology is comprised of various genetic and environmental factors. The major forms of IBD include ulcerative colitis (UC) and Crohn's disease (CD), which are of distinct clinical phenotypes. Genetics studies to date have revealed at least 99 susceptibility loci for IBD, which, however, explain only a small percentage of the overall variance in disease risk, especially for UC. To overcome the power limitation of genome-wide association (GWA) studies, we employed pathway-based approaches to identify gene sets that cooperatively make contribution to the genetic etiology of UC. We compiled 6 UC European cohorts, totaling 5584 UC cases and 11587 controls. Our pathway-based meta-analysis of the 6 cohorts confirmed multiple immune-related pathways that have been implicated in the genetic etiology of IBD, such as “Antigen processing and presentation”, “Type I diabetes mellitus” and “Asthma”, at the genome-wide significance level. We noticed the genes in the major histocompatibility complex (MHC) region made a remarkable contribution to the significance of these pathways. The large GWA signals from MHC-linked genes could potentially mask involvement of other important pathways. To delineate how genes outside of the MHC region contribute to the etiology of UC, we excluded the SNPs in the MHC region and re-surveyed top pathways. As expected, many other immune-related or immunoregulatory pathways remain highly significant, and their association signals replicated in an independent pediatric cohort of European ancestry (CHOP) and in an adult cohort genotyped on the Immunochip. Pathways without obvious immune involvement with significant signals replicating in all three cohorts include those that regulate phosphorylation, such as “Lipid phosphorylation” (p values of 0.00053, 0.034, 0.0099 in discovery, CHOP, and Immunochop cohorts, respectively). Interesting pathways with suggestive signals (i.e. significant in discovery and one replication cohort) include the “Classical complement cascade”, “Parkinson's disease”, “Methionine metabolism”, and “Pantothenate and CoA biosynthesis” pathways. The possible contribution of these pathways to the etiology of UC and insight into the genetic mechanisms of UC are discussed.

2230T

Shared genetic variants between traits and diseases reveal novel disease risk factors. *L. Li[1], D. Ruau[1], C. Patel[1], R. Chen[1,2], A. Butte[1].* 1) Pediatric Department, Stanford University, Stanford, CA; 2) Personalis Inc.

Clinical manifestations of risk factors, or traits, are critical for prognostics and preventive care of complex disease. Currently, these risk factors are identified for disease through large scale epidemiological studies, which are time consuming, costly, and lose power to discover associations for less prevalent diseases. In the last decade, GWAS and candidate gene studies have identified genetic variants for thousands of diseases and traits. We utilized those variants and apply a method to systemically identify common genetic associations between traits and diseases. We hypothesize that traits sharing common genetic variants with diseases could serve as non-invasive means to prognose disease susceptibility. We validated our novel findings using independent patient cohort databases. We used a manually curated database (Varimed) of 8962 GWAS and candidate gene studies, containing 87553 SNPs from 2376 diseases and traits. We obtained 97 diseases and 102 traits by restricting to variants with p<1e-6 and diseases or traits associated with more than 3 genes. Disease-trait similarity was first estimated by the cosine distance between genome-wide binary vectors of associated genes. We then hierarchical clustered using 1000 bootstrap resampling. To assess the clinical relevance of our new predictions, we used summary statistics extracted from Stanford Translational Research Integrated Database Environment (STRIDE) as well as the National Health and Nutrition Examination Survey (NHANES). There were 1986 genes associated in 97 diseases (median=10) and 1718 genes associated in 102 traits (median=9). We identified 42 significant disease-trait pairs. 83% of the pairs were disease and traits from different publications. 28 pairs were known risk factors or predictive markers and 14 pairs were novel findings. We obtained 5 cohorts from STRIDE or NHANES and all were validated: elevated serum pro-brain natriuretic peptide levels and IgA nephropathy (OR:2.4;[1.5-3.7];p<0.001), elevated serum IgE levels and psoriasis (OR:1.7;[1.2-2.6]; p=0.005), elevated erythrocyte sedimentation rate and Alzheimer's disease (OR:1.9;[1.5-2.5]; p<0.0001), higher prostate-specific antigen levels in basal cell carcinoma (p<0.05), and higher bone mineral density in sudden cardiac death (p=0.04). We developed a systemic way to identify novel associations between traits and diseases, suggesting some of these traits can serve as a novel non-invasive means for disease prognosis.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2231F

Comparison of pathways implicated in anti-citrillunated peptide antibody positive and negative rheumatoid arthritis patients. *P. Martin[1], S. Viatte[1], A. Brass[2], S. Eyre[1], Rheumatoid Arthritis Consortium for Immunochip (RACI).* 1) Arthritis UK Epidemiolgy Unit, University of Manchester, Manchester, United Kingdom; 2) School of Computer Science, Kilburn Building, Oxford Road, Manchester, United Kingdom.

Introduction: Rheumatoid Arthritis (RA) is a chronic autoimmune disease affecting up to 1% of the adult population. GWAS have been successful in identifying SNPs associated with disease. However, previous studies have been under-powered to detect differences between anti-citrillunated peptide antibody (ACPA) positive and negative patients. Using a custom Illumina array, the RA Immunochip consortium allows comprehensive analysis of disease associations in both ACPA positive and negative patients. Pathway analyses test whether a molecular pathway is associated with disease by testing for enrichment of associated genes. This involves mapping SNPs to genes which can often be problematic or subjective. Here we present a robust workflow based method to assign genes and a comparison of the pathways significantly enriched in ACPA positive and negative patients.

Method: Excluding HLA, 17 confirmed RA SNPs attaining GWAS significance ($p<5x10^{-8}$) in ACPA positive patients were selected from the Immunochip analysis along with the 17 most significant SNPs from known RA associations from the ACPA negative analysis. Genes were assigned to SNPs using a Taverna workflow which defines a region using SNPs in linkage disequilibrium (LD) with the associated SNP. This prevents the bias introduced by assigning SNPs to `biologically plausible' genes and provides a robust, reproducible method for assigning SNP associations to genes. These genes were then tested for enrichment in PANTHER pathways using the Expression Analysis tool.

Results: Overall 4 pathways were significantly ($p<0.05$) enriched in the ACPA positive gene list and 5 in the ACPA negative gene list, with 3 pathways significantly associated in both. The toll receptor signalling pathway was associated in the ACPA positive ($p=0.0072$) group, the interleukin signalling pathway in ACPA negative ($p=0.0014$) patients and the JAK/STAT signalling pathway in both ($ACPA^+$ $p=0.038$; $ACPA^-$ $p=2.05x10^{-5}$). **Discussion:** Although preliminary, this study provides a promising pathway analysis of ACPA positive and negative RA patients. These findings not only suggest shared processes between the two disease sub-groups, but also identify areas where possible differences exist helping to define patient groups and potentially leading to better targeting of therapies.

2232W

Patterns of genetic expression in mental retardation associated or not with microcephaly. *T.F. Almeida[1], D.V. Bernardo[2], C.R.D.C. Quaio[1], G.L. Yamamoto[1], E.D.F. Carvalho[1], C.A. Kim[1], D.R. Bertola[1].* 1) Genetic Unity, Children's Institute, HCFMUSP, São Paulo, São Paulo, Brazil; 2) Laboratory of Human Evolutionary Studies, University of São Paulo, São Paulo, Brazil.

Mental retardation (MR) commonly runs with several genetic disorders and may or may not be associated with microcephaly. There is a vast description of single genes in which the disruption of their expression may be involved in these two groups of phenotypes (MR with microcephaly and MR without microcephaly), but fewer studies analyzing a great amount of genes at the same time. The pattern of a gene expression in a specific tissue directly relates with its importance in developing a phenotype. We primarily expected to find a greater expression in the nervous system of genes related to microcephaly than those not related to it. Therefore, we compared the pattern of the expression of genes between these two groups in different areas of the nervous system to determine if some of these regions might have a more important role in the development of microcephaly. Fifty genes for each group were selected in accordance to the order of appearance in OMIM. The level of expression of these genes was obtained by GNF expression atlas 2 data from U133A and GNF1H chips (UCSC site) for different tissues of the nervous system (fetal brain, adult brain, 16 different regions of the CNS and five regions of the peripheral nervous system). The level of expression of each gene was classified as: high expression (if expressed more in the area than the mean for the whole body), average expression and low expression for all of these areas. Frequency tables based on the expression of each gene for each area of the nervous system were compared between the two groups using Chi-squared test and Bonferroni correction for multiple analyses. The expression of genes related to microcephaly was statistically different when compared with genes not related to microcephaly in the fetal brain, the adult brain and four different regions of the CNS: the cerebellum, the cingulate cortex, the subthalamic nucleus and the globus pallidus. There were no major differences in gene expression for the other areas. The expression was comparatively lower in the microcephaly group, contrarily to our first expectation. This may be explained by the fact that 13 genes associated with microcephaly and none determining MR without microcephaly are involved in the cellular cycle and are thought to be equally expressed in the nervous system and the whole body. The regions where the expressivity of the genes was discordant between the two groups may also have an important role in determining microcephaly.

2233T

Molecular analysis of *GABRA1* and *EFHC1* genes in patients with juvenile myoclonic epilepsy and other idiopathic generalized epilepsies. *C.V. Soler[1], M.C. Gonsales[1], F.A. Oliveira[1], L.E. Betting[2], F. Cendes[2], I. Lopes-Cendes[1].* 1) Medical Genetics, University of Campinas, Campinas, Brazil; 2) Department of Neurology, Faculty of Medical Sciences, University of Campinas, Campinas, Brazil.

Purpose: Juvenile myoclonic epilepsy (JME) is the most common syndrome among the idiopathic generalized epilepsies (IGEs) with an onset between the ages of 9 and 27 years. It remains unclear whether different forms of IGEs could share a common genetic predisposition. Several genes have been associated with JME and other IGEs but the most relevant are the α1-subunit of the $GABA_A$ receptor gene (*GABRA1*), and the EFHC1 protein gene (*EFHC1*). The first is important to neuronal excitability regulation, while the second is involved in the central nervous system development and in cortical and subcortical architecture alterations. Therefore, the purpose of this study is to search for mutations in *GABRA1* and *EFHC1* in patients with JME other IGEs.

Method: We screened for *GABRA1* and *EFHC1* mutations in 52 patients with JME and 33 with other types of IGEs. In addition, computation algorithms were used for predicting the impact of amino acid exchanges in protein function.

Results: In *GABRA1*, we identified three silent mutations (c.96A> G, c.156 T> C and c.1323G> A) and an intronic (IVS1 +9 A> T), all present in SNPs databases and none were detected as potentially deleterious. In *EFHC1*, we identified two new variants (896A> G and 1766G> A), and six variants already described in SNPs databases (c.475C> T; 545g> A, 685T> C, 1343T> C; 1821A> G and 1855A> C). The predicted pathogenic potential of these variants was not conclusive amongst all analyses.

Conclusion: Since deleterious mutations were not found in *GABRA1*, this gene does not seem to be related to IGE in our cohort. We found eight sequence variants in *EFHC1*, seven of those were predicted as deleterious by at least one algorithm. All of these were present only in patients with JME. However, the deleterious prediction in *EFHC1* function did not reach a consensus. In addition, we did not observe preferential location of the mutations along *EFHC1*. Study supported by FAPESP and CNPq.

2234F

Genome-wide expression analysis of idiopathic scoliosis endophenotypes. *K. Gorman[1], S. Bouhnik[1], Q. Yuan[1], C. Julien[1], A. Franco[1], K. Choquet[1], L. Suvarnan[1], M.-Y. Akoume[1], A. Moreau[1,2].* 1) Viscogliosi Laboratory in Molecular Genetics of Musculoskeletal Diseases, Sainte-Justine University Hospital Research Center, Department of Biochemistry, Faculty of Medicine, Université de Montréal, Canada; 2) Department of Stomatology, Faculty of Dentistry, Université de Montréal.

Idiopathic scoliosis (IS) is a complex pediatric spinal deformity with an unknown etiopathogenesis. Although a genetic basis is supported, phenotypic variability for morphological and developmental (i.e. propensity for curve progression) factors have complicated efforts to understand it. Genetic heterogeneity is presumed from the limited consensus for loci identified. For complex syndromes, the use of well characterized endophenotypes will stratify patients and give studies more power to identify genetic correlates. We previously reported elevated levels of osteopontin (OPN) and a differential Gi protein signalling defect among IS patients. Three distinct Groups are associated with a hypofunctionality of Gi proteins and OPN resistance or sensitivity on a cellular level. The goal of this study is to identify differential gene expression associated with these biochemically distinct Groups. Fifteen age-matched female patients (5 from each Group) with severe curves were selected for genomic expression analysis. Microarrays, on osteoblasts derived from bone samples obtained intraoperatively, were performed at Centre d'Innovation de Génome Québec à Montréal in triplicate using the Affymetrix GeneChip® Human gene 1.0 ST array. Raw data was normalized using Robust Means Analysis, genes were filtered using a Kruskal-Wallis test and a 3-fold threshold for differential expression. Data was corrected using the Benjamini & Hochberg algorithm and genes with a $p \leq 0.05$ FDR were considered. There are 185 genes significantly differentially expressed among the three Groups. Supervised clustering showed that Group 1 is distinct from Groups 2 and 3, and Groups 2 and 3 have similar expression patterns. Unsupervised clustering showed two distinct gene clusters. These are defined by opposing expression patterns between Group 1 and Groups 2 and 3. Importantly, this pattern of gene expression mirrors the cellular activity for OPN previously demonstrated in our lab. Candidate genes from the filtered gene set will relate to a OPN sensitivity and may also explain some of the differential co-morbidities and prognoses among the Groups. Candidate genes are being confirmed on an expanded patient cohort using quantitative PCR. We expect that key genes related to IS endophenotypes will allow for better patient diagnoses and more personalized therapies, and thus better patient prognoses.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2235W

Molecular profiling of osteogenic disregulation in non syndromic craniosynostosis: new insights from microarray and morphological studies. *M. Barba[1, 2], W. Lattanzi[1, 2], G. Tamburrini[3], M.C. Geloso[1], L. Massimi[3], M.A. Puglisi[4], M. Caldarelli[3], C. Di Rocco[3], F. Michetti[1, 2], C. Bernardini[1].* 1) Institute of Anatomy and Cell Biology, Università Cattolica del Sacro Cuore, Rome, Italy; 2) Latium Musculoskeletal Tissue Bank, Rome, Italy; 3) Pediatric Neurosurgery Unit, Università Cattolica del Sacro Cuore, Rome, Italy; 4) Institute of Internal Medicine and Gastroenterology, Università Cattolica S. Cuore, Rome, Italy.

Introduction: The genetic basis and the molecular pathogenesis of non syndromic craniosynostosis (NS-CRS) are still largely unknown. This study attempted to clarify the molecular mechanisms underlying the premature ossification of calvarial sutures, using microarray analysis to perform exon-level genome-wide expression on calvarial suture specimens of NS-CRS patients. Methods: Calvarial specimens from both fused and unfused sutures were collected during surgery from patients affected by different craniosynostosis. Matched samples of each patient was used for comparative microarray analysis using GeneChip human exon 1.0 arrays by Affymetrix, which enable two complimentary levels of analysis, gene expression and alternative splicing. The functional relevance of these data was analyzed comparatively in calvarial cells isolated from fused-versus-patent suture specimens using confocal microscopy. Results: Gene-level analysis allowed identifying 28 significantly modulated genes, mainly involved in cell adhesion, cell-matrix interaction, matrix mineralization, osteogenesis, and tissue development. Exon-level analysis produced a list of over 600 genes with significant ($p<0.01$) alternatively spliced variants in fused-versus-patent sutures, indicating the differential somatic expression of different gene isoforms at the dysmorphic site. Interestingly, this list included the CRS-associated genes FGFR1-3 and TWIST1 along with several genes involved in the structure and functions of the centrosome/primary cilium/mitotic spindle constituents. The confocal microscopy analysis showed that cells isolated from fused sutures display a predisposition to produce differently shaped and developed primary cilia during differentiating stages (G0). During cell replicative stages (M), these cells showed more frequent mitotic events with irregularly oriented spindles, compared to patient-matched controls. Conclusions: These data seemed to suggest that a constitutive somatic alteration in the centrosome-primary cilium-mitotic spindle apparatus could underlie the overworking osteogenic process existing at the site of calvarial suture fusion.

2236T

Network-based Genomic Mapping highlights the influence of Central-Nervous System Pathways on Obesity. *M. Nikpay[1], P. Lau[1], R. Dent[2], R. McPherson[1].* 1) Atherogenomics Laboratory, University of Ottawa Heart Institute, Ottawa, Ontario, Canada; 2) Weight Management Clinic, The Ottawa Hospital, Ottawa, Ontario, Canada.

Introduction: Given the modular nature of a complex trait, the underlying genetic factors share protein interactions and display coordinated gene expressions; as such, we used these criteria to narrow the list of GWAS loci for obesity and investigate the underlying biological processes in a cohort of 958 obese (BMI ≥ 32 kg/m2) and 869 lean (BMI ≤ 23 kg/m2) subjects (OBLE study) and two other samples selected from controls of OHGS (Ottawa Heart Genomic Study)_B2 (288 obese and 421 lean subjects) and OHGS_A2 (98 obese and 271 lean subjects) to determine the validity of this approach. **Method:** Subjects in OBLE and OHGS_B2 were genotyped using the Affymetrix 6.0 Array, and in OHGS_A2 genotyping was done using 500K Array; moreover, the densities of our SNP Maps were increased by genotype-imputation from the 1000 Genomes dataset. In each cohort, nominally associated loci ($P < 0.001$) were searched in STRINGdb and their protein interactions were extracted by specifying the highest confidence score (Score > 0.9). Next, genes that shared protein interactions were subjected to co-expression analysis using data from COXPRESdb database and those that formed a network of co-expressed (Pearson's $r < 0.5$) genes were selected. We further reduced the number of genes by performing the statistical test of epistasis and retaining those with significant interactions. **Results:** We determined the expression of the resulting genes from OBLE (33 Genes), OHGS_B2 (12 Genes) and OHGS_A2 (12 Genes) across 65 normal human tissues using data from COXPRESdb and found that the identified genes in each study display high and coordinated expression profiles in cluster of brain-related tissues, and more particularly in parietal and frontal brain lobes. To control for possible systematic bias in our approach, we carried out the same set of analyses in OHGS_A2 and OHGS_B2 for the phenotype, Coronary Artery Disease; however, a very dissimilar pattern was obtained. Finally, while the basic model with only sex and age explains for 0.09 of variability of obesity in OHGS_A2, 0.2 in OHGS_B2 and 0.02 in OBLE study; these values increased to 0.45, 0.41 and 0.24, respectively, when SNPs identified by our method were included. **Conclusion:** Our findings support the notion that obesity is in part a central nervous system-mediated trait and show that the network properties of complex trait genes can be utilized to elucidate the genetic nature and the biological processes underlying a complex phenotype.

2237F

A comprehensive knowledgebase of autism genetic evidence. *Y. Huang, L. Xu, J. Li, M. Zhao, X. Tang, L. Wei.* Center for Bioinformatics, School of Life Sciences, Peking University, Beijing, P.R.China, Beijing, China.

Autistic spectrum disorder is a complex neurodevelopmental disorder with three cores of behavioral impairment. Twin studies have highlighted its strong genetic etiology, and many previous efforts have been put on identifying the genetic components of autism. To evaluate how many genes contribute to autism and how strong it is, a collection and integration for genetic evidence of autism is called for. Here we compiled a comprehensive autism knowledgebase including 2193 genes, 2806 genetic markers, 158 linkage regions, and 4544 structural variations which are associated with autism by various types of studies. The metadata of each studies and genes including clinical features, experimental procedures, and statistical values were also extracted. By using a ranking system, we identified a total of 434 highly-confident genes, and these genes are located in pathways related to synapse functions. With integrating annotations extracted from more than 30 external resources, we constructed a public online database - autismKB (http://autismkb.cbi.pku.edu.cn) for autism research community.

2238W

Mathematical Modeling of the Insulin Signal Transduction Pathway for Prediction of Insulin Sensitivity from Expression Data. *C.K. Ho[1], J.C. Liao[1,2], G. Sriram[2,3], K.M. Dipple[1,4].* 1) Biomedical Engineering Interdepartmental Program, University of California Los Angeles, Los Angeles, CA; 2) Department of Chemical and Biomolecular Engineering, University of California Los Angeles, Los Angeles, CA; 3) Department of Chemical and Biomolecular Engineering, University of Maryland, College Park, MD; 4) Departments of Human Genetics and Pediatrics, University of California Los Angeles, Los Angeles, CA.

Insulin sensitivity is a key component in understanding obesity and type II diabetes mellitus (T2DM). We developed a mathematical model of the insulin signal transduction pathway, which is an extension of a previous model (Sedaghat et al., *Am J Physiol Endocrinol Metab*, 283, E1084, 2002), to quantitatively predict insulin sensitivity in biological systems. Our extended mathematical model incorporates key recently-elucidated mechanisms of action in the insulin signal transduction pathway and predicts the level of glucose transporter (GLUT4) translocation to the plasma membrane in response to a given insulin dose. The model output also provides information regarding glucose clearance, a measurement used routinely in patients with T2DM. We validated this model using published *in-vivo* data that reported differential expression levels of the key genes involved in the insulin signaling pathway. We further evaluated our extended model by applying the model to biological systems, including mice, humans, and rats. Using microarray data of glycerol kinase knockout (Gyk KO) and wild type mice, our simulation showed that the GLUT4 translocation and glucose uptake rate of the Gyk KO is significantly reduced by 3.3% compared to wild type mice. Additionally, we simulated the effect of acute and short term exercising based on insulin sensitivity levels for i) obese non-diabetic or type II diabetic human subjects and (ii) conditioned rat models. Our model predicted that glucose clearance of T2DM patients is improved by 39.1% after performing 7 days of exercise, which is consistent with published data. Our simulations with the rat models also showed the conditioned rats have improved glucose clearance by 11.8%, compared to the sedentary rats. This is also consistent with the literature, indicating that our model is a good predictor of glucose clearance. This extended model is a useful tool for predicting insulin sensitivity in mammalian tissues with altered gene expression in the insulin signal transduction pathway.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2239T

Examining genetic risk factors of obesity: Pathway analysis of GWAS data. *M.A Simonson[1], M.C Keller[1,2], M.B McQueen[1,2].* 1) Psychology and Neuroscience, Institute for Behavioral Genetics, Boulder, CO; 2) Integrative Physiology, Institute for Behavioral Genetics, Boulder, CO.

Classic genome-wide association studies are generally limited in their ability explain a large portion of genetic risk for most common diseases. We sought to use both classic GWAS methods, as well as multiple methods of pathway-based analysis applied to single-nucleotide polymorphism (SNP) data to examine the underlying biology of obesity risk. We used data from the NHLBI Multi-Ethnic Study of Atherosclerosis (MESA) SNP Health Association Resource (SHARe), composed of 2343 case-control subjects, as well as the Atherosclerosis Risk in Communities (ARIC) Study, composed of 12771 case-control subjects, and the Coronary Artery Risk Development in Young Adults (CARDIA) Study - Gene Environment Association Studies Initiative (GENEVA), composed of 1808 case-control subjects. A total of 16922 subjects genotyped at ≈1 million SNPs using the Affymetrix Affy 6.0 array. Our investigation performed 3 complimentary forms of pathway analysis examining curated gene sets from online pathway databases, and publications in PubMed. The results of this analysis suggest common SNPs affect biological processes that relate to obesity risk. Work was supported by the National Institute of Child Health and Human Development - National Institutes of Health: Institutional Training Grant (T32 HD 7289-27, John K. Hewitt, Director).

2240F

Hypothesis independent pathway analysis identifies biologic pathways implicated in POAG etiology. *B.L. Yaspan[1], J.L. Wiggs[2], L.R. Pasquale[2,3], S.J. Loomis[2], J.H. Kang[3], D.L. Budenz[4], D.S. Friedman[5], D. Gaasterland[6], T. Gaasterland[7], R.K. Lee[8], P.R. Lichter[9], Y. Liu[10, 17], C.A. McCarty[11], S.E. Moroi[9], L.M. Olson[1], J.E. Richards[9], T. Realini[12], J.S. Schuman[13], K. Singh[14], D. Vollrath[15], G. Wollstein[13], D.J. Zack[5], R.R. Allingham[10], M.A. Pericak-Vance[8], R.N. Weinreb[16], K. Zhang[16], M.A. Hauser[10,17], J.L. Haines[1].* 1) Center for Human Genetics Research, Vanderbilt University, Nashville, TN; 2) Department of Ophthalmology, Mass Eye & Ear Infirmary Boston, MA; 3) Department of Medicine, Channing Division of Network Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA; 4) Department of Ophthalmology, University of North Carolina, Chapel Hill, NC; 5) Wilmer Eye Institute, Johns Hopkins University Hospital, Baltimore, MD; 6) Eye Doctors of Washington, Chevy Chase, MD; 7) Scripps Genome Center, University of California at San Diego, San Diego, CA; 8) Bascom Palmer Eye Institute and Human Genomics, University of Miami Miller School of Medicine, Miami, FL; 9) Department of Ophthalmology and Visual Sciences, University of Michigan, Ann Arbor, MI; 10) Department of Ophthalmology Duke University Medical Center, Durham, NC; 11) Essentia Institute of Rural Health, Duluth, MN; 12) Department of Ophthalmology, WVU Eye Institute, Morgantown, WV; 13) Department of Ophthalmology, UPMC Eye Center, University of Pittsburgh, Pittsburgh, PA; 14) Department of Ophthalmology, Stanford University, Palo Alto, CA; 15) Department of Genetics, Stanford University, Palo Alto, CA; 16) Department of Ophthalmology and Hamilton, University of California, San Diego, SD, CA; 17) Department of Medicine, Duke University Medical Center, Durham, NC.

Using genome-wide SNP data, we assessed biologic pathways as annotated in the KEGG database for association with risk of primary open-angle glaucoma (POAG). Our pathway analysis includes the entirety of the SNPs in the GWAS to examine associations that do not meet the genome-wide significance threshold. The PARIS (Pathway Analysis by Randomization Incorporating Structure) algorithm used in this study creates random pathways to mimic the size and structure of the actual pathways. The algorithm is also designed to determine whether statistical associations are due to the biologic properties of the pathway or instead its size and structure. We conducted a POAG case-control analysis of 1002 cases and 1183 controls from the GLAUGEN study and 2,517 POAG cases and 2,428 controls from the NEIGHBOR study using the Illumina 660W quad platform. We first analyzed each study individually for single-SNP associations using logistic regression models to test for association. A subsequent meta-analysis of the two studies was performed. After multiple testing correction, we found 14 pathways associated with risk of POAG ($p<0.001$ for all pathways). These pathways are loosely grouped into three categories: metabolism, cellular adhesion and signaling, and autoimmune disorders. Twelve of these pathways contain a substantial amount of genic overlap in the form of the HLA gene family. After removal of the HLA genes from these pathways, 11 were no longer significant, while the HLA gene set was highly significant ($p<0.001$). Furthermore, 2 pathways, butanoate metabolism (hsa00650) and basal transcription factors (hsa03022) contain genes that do not overlap with any other pathway in the database and thus are entirely independent gene sets. We then tested the 14 pathways for association with risk of POAG in the subset of our cases with normal tension glaucoma, the clinical subdivision that our previous GWAS analysis indicated harbors most of the genetic information. Only the butanoate metabolism pathway held statistical significance after correction for multiple testing. A hypothesis-independent approach identified the HLA gene family and the butanoate metabolism pathway as being associated with risk of POAG. These data provide insight into the complex genetic etiology of POAG and provide hypotheses for future functional studies.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2241W

Selecting likely causal genes from genome-wide association studies of BMI and height by integrative network-based analysis. *T.H. Pers[1,2,3], S. Vedantam[1,2,4], T. Esko[5], F. Day[6], S. Berndt[7], S. Gustafsson[8], A.E. Locke[9], A.R. Wood[13], B. Kahali[9], D.C. Croteau-Chonka[10], C. Powell[9], P. Dworzynski[3], C.E. Thomas[3], A.G. Pedersen[3], S. Brunak[3], A. Justice[10], K.L. Mohlke[10], E. Ingelsson[8], R.J.F. Loos[6,11], E.K. Speliotes[9,12], T.M. Frayling[13], J.N. Hirschhorn[1,2,4], The Genetic Investigation of Anthropometric Traits (GIANT) Consortium.* 1) Division of Endocrinology, Children's Hospital Boston, Boston, MA; 2) Metabolism Initiative and Program in Medical and Population Genetics, Broad Institute, Cambridge, Massachusetts 02142, USA; 3) Center for Biological Sequence Analysis, Technical University of Denmark, Lyngby, Denmark; 4) Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA; 5) Estonian Genome Center of University of Tartu (EGCUT), Tartu, Estonia; 6) MRC Epidemiology Unit, Institute of metabolic Science, Cambridge, UK; 7) Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Department of Health and Human Services, Bethesda, Maryland 20892, USA; 8) Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, 171 77 Stockholm, Sweden; 9) Department of Biostatistics, Center for Statistical Genetics, University of Michigan, Ann Arbor, Michigan 48109, USA; 10) Department of Genetics, University of North Carolina, Chapel Hill, North Carolina 27599, USA; 11) Mount Sinai School of Medicine, New York, NY; 12) Department of Internal Medicine, Division of Gastroenterology, University of Michigan, Ann Arbor, Michigan, USA; 13) Genetics of Complex Traits, Peninsula Medical School, Exeter, UK.

Genome-wide association (GWA) studies have identified thousands of loci where common variants are associated with polygenic diseases or traits. Because of linkage disequilibrium between variants and the possibility of distant regulatory effects, these associated loci typically contain multiple genes. Many loci have no single obviously causal gene, so the challenge for moving from association to novel biological insight is to identify which gene at each locus most likely explains the association. Because follow-up functional studies are often intensive, a key first step is to use computational approaches to prioritize genes within loci as most likely to be biologically relevant. Previous computational approaches have shown some success but often focus on single types of data, limiting their discriminatory power. We have developed a data-driven network-based approach that integrates a diverse panel of data types (including tissue-specific gene expression data, protein-protein interaction data, phenotypic information from mouse knockout studies, and information text-mined from the scientific literature) to systematically identify the most likely causal gene(s) at a given locus. Our method compares each gene within a locus to a set of likely positive and negative genes (which can be derived in unbiased fashion from the GWA data itself, excluding the locus under consideration) and partitions the results for each gene into data type-specific likelihoods. We evaluate our method using the GIANT consortium's GWA meta-analyses for human height and body mass index (BMI). For GWA loci of both human height and BMI, our integrative approach performs consistently better in predicting likely causal genes than approaches based on one or two data types only. As unbiased benchmarks, we tested for enrichment of genes that have associated missense variants or that are closest to the best associated SNPs, both of which are partial indicators of causality. Genes prioritized by our method are enriched for genes with missense variants (height 2.3-fold, P= 0.002; BMI 3.7-fold, P=0.03), and genes nearest to the lead SNPs (height 2.4-fold, P=4.5*10-18; BMI 2.7-fold, P=6.5*10-5); based on these and other benchmarks, this approach outperforms methods using fewer data types. We have developed an unbiased computational approach that integrates a variety of data types and performs well in prioritizing potentially causal genes from GWA data.

2242T

A systems biology approach to ionotropic GABAergic networks in epilepsy, a model of complex neurogenetic disease. *M. Jaworski.* Mental Health Services, University of Ottawa Health Services, Ottawa, ON, Canada.

Ionotropic GABAergic neurotransmission requires numerous co-ordinated steps involving multimolecular complexes distributed in spatially and developmentally diverse brain networks. Complex processes whereby GABAergic neurons synthesize and transport GABA into and out of synaptic vesicles, and into the synaptic cleft involve over 300 interacting gene products. Mechanisms which bring presynaptic GABAergic neurons into apposition with their correct post-synaptic neuronal targets are also finely tuned. Regulatory mechanisms modulating the cell surface expression, and post-translational modification of ionotropic GABA receptor subunits are both complex and diverse. Through mammalian neurosciences, particularly proteomics, we now possess a detailed, though incomplete, inventory of the proteins mediating these networks. This current state of knowledge is applied to epilepsy, or perhaps more accurately, the epilepsies, as a model of a common, genetically complex human disease which has the advantage of objective (endo)phenotyping (EEG and increasingly, neuroimaging). The functional effects of epilepsy susceptibiity mutations (including copy number variation) are increasingly understood. For some forms of epilepsy, (Juvenile Myoclonic Epilepsy and Childhood Absence Epilepsy), integrating knowledge from genomics, proteomics, and connectomics is becoming an increasingly achievable goal. Mutated gene products known to be expressed in the GABAergic presynapse (KCNA1 and KCNC3, as well as others), as well as postsynaptic receptor subunits (GABRA1, GABRB3 and GABRG2, amongst others) and their associated scaffolding/trafficking interactants (ARHGEF9, GPHN, SRC, FYN and others) which are also expressed in the GABAergic postsynapse, as well as mutated gene products involved in GABAergic cell proliferation (BRD2) in anatomically defined brain networks, increasingly point to the ionotropic GABAergic system as one of the important contributors to genetic susceptibility to epilepsy. This systems biology (neuroscience) approach, drawing on proteomics and cell circuitry, may suggest one possible mechanism for the extensive locus and alleleic heterogeneity, as well as the pleiotropy of susceptibility genes implicated in neurogenetic and neuropsychiatric diseases.

2243F

***P2RX7* SNPs and external apical root resoprtion during orthodontia.** *J. Hartsfield[1,2,4], G. Falcao-Alencar[1], J. Dempsey[1], E. Jacobson[1], G. Kluemper[1], D. Fardo[2], J. Macri[3], L. Morford[1,4].* 1) College Dentistry, Univ Kentucky, Lexington, KY; 2) College Medicine, Univ Kentucky, Lexington, KY; 3) College Public Health, Univ Kentucky, Lexington, KY; 4) Center for Oral Health Research, Univ Kentucky, Lexington, KY.

Objective: Previously, an *IL-1B* genetic marker responsible for decreased IL-1β in cell culture has been associated with external apical root resorption (EARR) during orthodontia. Functional*P2RX7* gene polymorphisms can also influence secretion of IL-1β. We hypothesized that there is an association between functional *P2RX7* polymorphisms and EARR in orthodontic patients. Materials and Methods: Genomic DNA was isolated from 459 Caucasian orthodontic patients (74 EARR, 385 control). Pre/post-treatment radiographs were analyzed by three investigators for maxillary incisor EARR. Variations at rs1718119, rs208294, and rs2230912 were assayed using Taqman®-based SNP genotyping. Hardy-Weinberg Equilibrium (HWE) testing was employed to assess genotyping quality, and Cochrane-Armitage trend analysis was used for association testing. Results: No departures from HWE were detected. Individuals homozygous for the major allele and heterozygous at rs208294 were more likely to have EARR (p=.02) than those homozygous for the minor allele. For rs1718119, people homozygous for the major allele were more likely to experience EARR (p=0.03) than those either heterozygous or homozygous for the minor allele. Individuals homozygous for the major rs2230912 allele were more likely to experience EARR (p=.03) than either the heterozygous or homozygous for the minor allele. Conclusion: Functional SNPs in the *P2RX7* gene are associated with EARR during orthodontia.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2244W

Pathway analysis of variation in DNA repair genes indicates involvement of repair of oxidative DNA lesions in human longevity. *M. Soerensen[1, 2], B. Debrabant[1, 3], J. Mengel-From[1, 2], T. Stevnsner[4], V.A. Bohr[4, 5], K. Christensen[1, 2], Q. Tan[1, 2], L. Christiansen[1, 2].* 1) The Danish Aging Research Center, Epidemiology, Institute of Public Health, University of Southern Denmark, Odense C, Denmark; 2) Department of Clinical Genetics and Department of Clinical Biochemistry and Pharmacology, Odense University Hospital, Odense C, Denmark; 3) Biostatistics, Institute of Public Health, University of Southern Denmark, Odense C, Denmark; 4) The Danish Aging Research Center, Institute of Molecular Biology and Genetics, University of Aarhus, Aarhus, Denmark; 5) Laboratory of Molecular Gerontology, National Institute on Aging, National Institute of Health, Baltimore, Maryland, USA.

Approximately 25% of the variation in human life span is thought to be caused by genetic variation, an effect believed to be minimal before age 60 and most profound from age 85 and onwards. An immense number of association studies reporting association of single-variants to human longevity have been published. These single-variants are located in genes taking part in different biological processes including lipoprotein metabolism, oxidative stress and DNA repair.

In the present study we aim to assess the involvement of DNA damage response and DNA repair processes in human longevity by performing case-control pathway analysis of 592 tagging single nucleotide polymorphisms (SNPs) of 77 genes in 1089 oldest-old (92-93) and 736 middle-aged (46-55) Danes. The 77 genes encode proteins participating in nine sub-pathways of key importance for genetic stability and, hence, the aging process: DNA damage response, telomere maintenance, mitochondrial DNA processes, base excision repair, nucleotide excision repair, mismatch repair, non-homologous end-joining, homologous recombinational repair and the RecQ helicases.

For the pathway analysis we applied a method presented by Bucan and co-workers (Wang et al. 2007), which extends gene set enrichment analysis to SNP data. We analyzed the nine different sub-pathways basing the pathway analysis on a best model procedure, i.e. we considered all possible modes of inheritance (an assumption free genotypic model, as well as additive, recessive and dominant genotypic models) for each SNP and combined the highest test statistics for the SNPs in the gene statistics. The pathway statistics showed that of the nine sub-pathways homologous recombinational repair and base excision repair displayed association ($p<0.05$) with case-control status, however, only for base excision repair the p-value (p-value = 0.016) was based on associations of SNPs in more than one gene. Therefore, our data indicate that longevity might be influenced by variation in the genes encoding proteins conducting repair of DNA damage induced by for instance reactive oxygen species.

All participants gave their informed consent prior participation. The project has been approved by the Danish National Committee on Biomedical Research Ethics. Wang K, Li M, and Bucan M, American journal of human genetics, 81(6):1278-1283, 2007.

2245T

Constructing Endophenotypes of complex disease using Nonnegative Matrix Factorization and Adjusted Rand index. *H.M. Wang[1,2], C.L Hsiao[2], A.R Hsieh[2], Y.C Lin[2], S.J Fann[2].* 1) Division of Biostatistics, Institute of Public Health, National Yang-Ming University, Taiwan; 2) Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan.

Complex diseases are typically caused by combinations of molecular disturbances that vary widely among different patients. Endophenotypes, a combination of genetic factors associated with a disease, offer a simplified approach to dissect complex trait by reducing genetic heterogeneity. Because molecular dissimilarities often exist between patients with indistinguishable disease symptoms, these unique molecular features may reflect pathogenic heterogeneity. To detect molecular dissimilarities among patients and reduce the complexity of high-dimension data, we have explored an endophenotype-identification analytical procedure that combines non-negative matrix factorization (NMF) and adjusted rand index (ARI), a measure of the similarity of two clusters of a data set. To evaluate this procedure, we compared it with a commonly used method, principal component analysis with k-means clustering (PCA-K). A simulation study with gene expression dataset and genotype information was conducted to examine the performance of our procedure and PCA-K. The results showed that NMF mostly outperformed PCA-K. Additionally, we applied our endophenotype-identification analytical procedure to a publicly available dataset containing data derived from patients with late-onset Alzheimer's disease (LOAD). NMF distilled information associated with 1,116 transcripts into three metagenes and three molecular subtypes (MS) for patients in the LOAD dataset: MS1 (n1=80), MS2 (n2=73), and MS3 (n3=23). ARI was then used to determine the most representative transcripts for each metagene; 123, 89, and 71 metagene-specific transcripts were identified for MS1, MS2, and MS3, respectively. These metagene-specific transcripts were identified as the endophenotypes. Our results showed that 14, 38, 0, and 28 candidate susceptibility genes listed in AlzGene database were found by all patients, MS1, MS2, and MS3, respectively. Moreover, we found that MS2 might be a normal-like subtype. Our proposed procedure provides an alternative approach to investigate the pathogenic mechanism of disease and better understand the relationship between phenotype and genotype.

2246F

Candidate genes of several complex diseases via an integrated analysis of GWAS and eQTL. *C. Fuller[1], X. He[2], Y. Song[1], H. Li[1].* 1) Integrative Program in Quantitative Biology, 1700 4th Street, University of California, San Francisco, CA 94158; 2) Lane Center of Computational Biology, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA 15213.

Genome-wide association studies (GWAS) for common disease can yield large numbers of loci, often in non-coding regions, with limited annotation and no obvious functional consequence. Many more loci may fail to reach genome-wide statistical significance but appear in excess relative to expectations under the null distribution. Together, such issues frustrate efforts to understand common disease etiology, since statistically significant, well-annotated loci usually explain only a small fraction of disease heritability. What is needed are straightforward methods to mine weak GWAS associations and place sets of loci in the context of functionally-relevant genes and gene networks. Our method attempts to do this by scoring GWAS results against thousands of genes from expression quantitative trait loci (eQTL) studies. Genes that share associations at multiple loci with GWAS are unexpected by chance and, thus, may play a causal role in the disease. Unlike earlier methods, this approach utilizes all SNPs that associate with a given gene, not just the proximal ones. The method uses a novel Bayesian framework to score the overlap between disease-loci associations from GWAS and gene-loci associations from an eQTL database. Multiple coincident associations boost the likelihood of a causal gene-disease relationship, whereas a penalty is assigned for each mismatched association. This approach is generalizable to any functionally-informative molecular trait for which trait-loci associations are known; it is not inherently limited to gene expression. We demonstrate our approach using GWAS for a number of common diseases: type 1 and type 2 diabetes, Crohn's disease, HIV progression, various cancers, and other conditions. Typically, our gene lists are enriched for highly-plausible genes, as suggested by either recent studies or the underpinnings of related diseases. Importantly, many supporting loci are found in -trans (i.e. distal to the gene itself) and have associations that fall below genome-wide statistical significance. Individually, such loci are uninformative, but collectively, they may yield key functional insights that cannot be extracted from the GWAS results alone.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2247W

Association study-based gene-set enrichment analysis identified biological pathways associated with fetal hemoglobin levels in sickle cell disease patients. *G. Galarneau[1,2], S. Coady[3], M. Puggal[3], N. Jeffries[3], D. Paltoo[3], A. Kutlar[4], G.J. Papanicolaou[3], G. Lettre[1,2].* 1) Montreal Heart Institute, Montreal, Quebec, Canada; 2) Université de Montréal, Montréal, Québec, Canada; 3) National Heart, Lung and Blood Institute, Bethesda, MD, USA; 4) Medical College of Georgia, Augusta, GA, USA.

INTRODUCTION: Fetal hemoglobin (HbF) levels, a highly heritable trait (h2~0.6-0.9), modify disease severity in the β-hemoglobinopathy sickle cell disease (SCD) and β-thalassemia. Single nucleotide polymorphisms (SNPs) at three loci (BCL11A, HBS1L-MYB, β-globin) are associated with HbF levels in anemic and non-anemic populations and clinical complications in SCD (pain crisis) and β-thalassemia (transfusion-dependency). These SNPs only explain 50% of the heritable variation in HbF levels.

METHODS & RESULTS: To identify new HbF regulators, we genotyped 1,213 SCD patients from the Cooperative Study of Sickle Cell Disease (CSSCD) on the IBC array, a gene-centric platform that includes ~50,000 SNPs and covers genetic variation at ~2,100 genes. We confirmed the association with BCL11A and the β-globin locus (the HBS1L-MYB intergenic region is not covered by the IBC array), but did not detect additional array-wide significant association signals. To identify additional HbF loci of smaller effect that might implicate shared biological processes, we performed a pathway analysis using gene-set enrichment methodologies (Wang et al., AJHG, 2009). We used the KEGG database to define biological pathways and phenotype permutations to assess statistical significance. We identified 12 highly redundant pathways that were nominally significant (P<0.05) and replicated seven of these in an independent set of 276 SCD patients. Our analysis suggests a role for acetyl-coA and short fatty acids metabolism as well as amino acid degradation in HbF levels variation.

CONCLUSION: We have identified and replicated biological pathways that are enriched for SNPs that influence HbF levels in SCD patients. Additional genetic validation and functional follow-up of these results will help understand the regulation of HbF production, a strong modifier of severity for two of the most common diseases worldwide.

2248T

Performances of single-SNP and pathway-based analyses of genome-wide data to detect genetic factors shared by eosinophil and basophil counts in asthma-ascertained families. *C. Loucoubar[1,2,3], M. Brossard[1,2,3], PE. Sugier[1,2,3], A. Vaysse[1,2,3], P. Jeannin[1,2,3], MH. Dizier[1,2,3], M. Lathrop[3], E. Bouzigon[1,2,3], F. Demenais[1,2,3], EGEA collaborative group.* 1) INSERM U946, Paris, France; 2) Université Paris Diderot, Paris, France; 3) Fondation Jean Dausset-CEPH, Paris, France.

Eosinophils (EOS) and basophils (BASO) are white blood cells (WBC) that coordinately mediate allergic inflammation. Correlation of their blood counts suggests the existence of shared genetic determinants. Genome-wide association studies of WBC counts in population-based studies have identified a number of loci and loci shared by EOS and BASO were reported in one study. Pathway-based methods have been proposed to facilitate the detection of trait-associated genes and these methods may be useful to identify pleiotropic genes. Our goal was to compare the ability of single-SNP and pathway-based analyses to detect genetic factors shared by EOS and BASO counts in 388 asthma-ascertained families from the French Epidemiological study on the Genetics and Environment of Asthma (EGEA) study. These families included 1,660 subjects who had been genotyped using an Illumina 610K chip. We used linear regression adjusted for family dependence to conduct single-marker association analysis of EOS and BASO counts with 2.5 millions Hapmap2-imputed SNPs. The pathway analysis,implemented in ALIGATOR, searched for overrepresentation of gene ontology(GO)classes containing association signals. The single-SNP analysis showed that the top hits were specific to EOS or BASO counts; the most significant loci, which represented new signals, were on 1p21.3 ($p=2x10^{-7}$)for EOS and on 8q21.11 ($p=4x10^{-6}$)for BASO. However, there were 7 loci(1q41, 2q36.1, 4q28.3, 6p21.3, 13q21.1, 15q14, 18p11.32) that included SNPs showing association signals ($2x10^{-4}\leq p\leq 8x10^{-4}$) with both EOS and BASO and having the same direction of effects. The pathway based-analysis examined association results from 711,681 SNPs that lied within 18,482 genes which were assigned to 7,070 GO categories. We found an overrepresentation of 7 and 10 GO categories for EOS and BASO respectively that achieved a category-specific p-value$\leq 6x10^{-4}$. These GO categories differed between EOS (immune response and lung development pathways) and BASO (calcium and sodium-potassium channel activities) and did not share any gene. This study shows that, in a situation where most trait-associated SNPs belong to genes that are spread across various biological pathways or are intergenic, pathway analysis does not have higher performance than single-SNP analysis to detect shared genetic determinants. Alternative gene-set methods to characterize pleiotropy will be further explored. Funding: ANR-11-BSV1-027-01 grant.

2249F

Association of common genetic variants with hypoxemia among Non-Hispanic White and African American COPD Cases highlights hematological disease genes. *M.N. McDonald[1], M.H. Cho[1,2], E. Wan[1,2], P. Castaldi[1], D.A. Lomas[3], H.O. Coxson[4], L.D. Edwards[5], W. MacNee[6], J. Vestbo[7], J.C. Yates[5], A. Agusti[8], P.M.A. Calverley[9], B. Celli[2], C. Crim[5], S. Rennard[10], E. Wouters[11], P. Bakke[12], X. Kong[14], R. Tal-Singer[14], B. Miller[14], A. Gulsvik[15], E.A. Regan[13], B.J. Makke[16], C. Lange[17], J.E. Hokanson[18], J.D. Crapo[16], T.H. Beaty[19], E.K. Silverman[1,2], C.P. Hersh[1,2] on behalf of the ECLIPSE and COPDGene Coauthors.* 1) Channing Lab, Brigham & Women's Hospital, Harvard Medical School, Boston, MA, USA; 2) Division of Pulmonary and Critical Care Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA; 3) Department of Medicine, University of Cambridge, Cambridge, UK; 4) UBC Department of Radiology, Vancouver General Hospital, Vancouver, Canada; 5) GlaxoSmithKline, Research Triangle Park, NC, USA; 6) Department of Respiratory and Environmental Medicine, University of Edinburgh, Edinburgh, Scotland; 7) Department of Respiratory Medicine, Manchester Academic Health Sciences Centre, University Hospital of South Manchester, Manchester, UK; 8) Thoracic Institute, Hospital Clinic, Barcelona, Spain; 9) Department of Pulmonary and Rehabilitation Medicine, University of Liverpool, Liverpool, UK; 10) Department of Medicine, Nebraska Medical Center, Omaha, Nebraska, USA; 11) Center for Chronic Diseases, University Hospital Maastricht, Maastricht, The Netherlands; 12) The Institute of Internal Medicine, University of Bergen, Bergen, Norway; 13) Epidemiology Department, National Jewish Health, Denver, CO, USA; 14) GlaxoSmithKline, King of Prussia, PA, USA; 15) Haukeland University Hospital and Institute of Medicine, University of Bergen, Bergen, Norway; 16) Division of Pulmonary Sciences and Critical Care Medicine, National Jewish Health, Denver, CO, USA; 17) Department of Biostatistics, Harvard School of Public Health, Boston, MA, USA; 18) Department of Epidemiology, Colorado School of Public Health, University of Colorado Anschutz Medical Campus, Aurora, CO, USA; 19) Johns Hopkins School of Public Health, Baltimore, MD, USA.

Background: Chronic Obstructive Pulmonary Disease (COPD) is the third leading cause of death in the U.S. Hypoxemia, defined as reduced oxygen levels in arterial blood, is a major complication of COPD and correlates with disease prognosis. Identifying genetic variants associated with hypoxemia may provide clues for deciphering underlying heterogeneity in prognosis among COPD patients. Methods: We performed genome-wide association studies on 2,819 Non-Hispanic White (NHW) and on 820 African American (AA) COPD cases from the COPDGene Study. All cases were genotyped on the Illumina HumanOmniExpress Beadchip for 730,525 markers. After quality control steps, each SNP was tested for its correlation with resting hypoxemia as measured by the linear trait $Sa0_2$, determined by transcutaneous oximetry at rest, under an additive model corrected for sex, current smoking status, lifetime smoking intensity, age and population substructure.The results from the NHW and AA studies were combined in a meta-analysis. Genes with significant SNPs ($P<5x10^{-5}$) were input to Ingenuity Systems Ingenuity Pathway Analysis (IPA) for overrepresentation of genes from known pathways. Results: One variant achieved genome-wide significance in the AA population (rs8025537 on chr 15, $P_{AA}=4.7x10^{-8}$). This variant was excluded from analysis in the NHW due to a low minor allele frequency (MAF=0.001%). This SNP lies 5' to the gene *RHCG*, which codes for Rh blood group CcEe antigens. The SNP rs6736403 on chr 2 was the most highly associated SNP in NHW ($P_{NHW}=2.7x10^{-7}$). IPA revealed that hematological disease genes (*RUNX1*, *COL18A1* and *UROD*) were overrepresented in the most significant meta-analysis results. The cellular development, embryonic development, nervous system development and function network was the most overrepresented network with an IPA score of 47. This network includes the *RUNX1*, *COL18A1* and *UROD* genes. Conclusion and Future Directions: We have identified one SNP in the*RHCG* gene that is associated with hypoxemia in AA COPD cases and several suggestive associations in NHW cases. Future directions include imputation of untyped variants using the 1000 genomes panel and development of new methods that allow for allelelic heterogeneity between NHW and AA populations. Any significant associations in NHW cases will be validated using the ECLIPSE Study. Funding: R01HL094635, R01NR013377, U01HL089856, U01HL089897. ECLIPSE was funded by GlaxoSmithKline (SCO104960; NCT00292552).

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2250W

ICAM3 expression level integrates the effects of six MS susceptibility alleles: dissecting a molecular cascade relevant to MS susceptibility. *T. Raj[1,2,3,4], M. Kuchroo[1,4], L.B. Chibnik[1,3,4], L. Ottoboni[1,3,4], X. Hu[2,4,5], B.T. Keenan[1,4], S. Raychaudhuri[2,4,5], B.E. Stranger[2,3,4], P.L. De Jager[1,3,4].* 1) Program in Translational NeuroPsychiatric Genomics, Institute for the Neurosciences Department of Neurology, Brigham and Women's Hospital, 77 Avenue Louis Pasteur, Boston, MA 02115, USA; 2) Division of Genetics, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA; 3) Harvard Medical School, Boston, MA 02115, USA; 4) Medical and Population Genetics Group, Broad Institute, Cambridge, MA 02142, USA; 5) Division of Rheumatology, Immunology, and Allergy, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA.

Objectives: There are now 49 confirmed multiple sclerosis (MS) susceptibility loci. We used several RNA expression datasets to: (1) systematically identify the proximal consequences of susceptibility alleles on the expression of nearby genes and (2) uncover functional networks that integrate these individual effects into dysfunctional immune responses. **Methods and results**: First, using a gene expression-based method and an atlas of CNS and immune cell RNA profiles, we identified CD11b+ dendritic cells and CD8+ T cells, in addition to CD4+T cells, as cell types that are preferentially targeted by MS risk alleles. This implicates many different immune but not CNS cell types in MS susceptibility. Second, using transcriptome profiles generated from (1) peripheral blood mononuclear cells of subjects with MS (n=225) and (2) 7 other immune cell types of healthy subjects, we report that 19 of the 49 MS risk-associated alleles alter RNA expression of genes in their vicinity using an additive expression quantitative trait locus (eQTL) analysis ("cis" effect). Further, using the Relative Trait Concordance Method, we show that altered RNA expression is the likely primary functional consequence of 9 of these alleles. Third, we assessed whether the susceptibility loci influenced each other's RNA expression levels ("trans" effect). For example, the MPHOSPH9 SNP influences ICAM3 RNA expression ($p=9.6x10^{-7}$). In fact, in addition to being associated in "cis" with a nearby MS risk allele, ICAM3 expression levels are associated with risk alleles of five other MS loci on different chromosomes: MPHOSPH9, ZNF767, NDFIP1, MPV17L2 and TMEM39A. Furthermore, we formally demonstrate that the effect of the NDFIP1 SNP on ICAM3 RNA expression is significantly mediated by increased NDFIP1 RNA expression. Similar, though not significant, relationships are seen with ZNF767 and TMEM39A RNA expression. **Conclusion:**Our systematic analyses provide compelling evidence for a primary role of altered RNA expression in the function of many MS susceptibility alleles. Some of these effects (1 cis- and 5 trans- eQTL associations), coalesce in increasing ICAM3 RNA expression, and we present evidence for a mechanistic effect for the NDFIP1 locus via increased NDFIP1 RNA expression. Given its role in early activation of T, NK and other immune cell-types, higher ICAM3 expression could lower the threshold for immune responses against self-antigens and therefore be a key node in susceptibility pathways.

2251T

Random enrichment of minor alleles of common SNPs affects complex traits and diseases in model organisms and humans. *S. Huang[1], X. Tan[1], J. Liang[1], C. Zeng[1], J. Zhang[2], J. Chen[2], L. Ma[1], A. Dogan[3], G. Brockmann[3], O. Goldmann[4], E. Medina[4], M. Xian[1], K. Yi[1], Y. Li[1], Q. Lu[1], Y. Huang[1], D. Wang[5], J. Yu[5], H. Guo[1], K. Xia[1], Z. Zhu[1], D. Yuan[1].* 1) State Key Laboratory of Medical Genetics, Central South University, Changsha, Hunan, China; 2) High Performance Computing Center, Modern Educational Technology Center, New Campus, Central South University, Changsha, Hunan 410083, China; 3) Department of Crop and Animal Sciences, Faculty of Agriculture and Horticulture, Humboldt-Universität zu Berlin, Invalidenstraβe 42, 10115 Berlin, Germany; 4) Infection Immunology Group, HZI - Helmholtz Centre for Infection Research, Inhoffenstraβe 7, 38124 Braunschweig, Germany; 5) CAS Key Laboratory of Genome Sciences and Information, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing 100029, China.

The null hypothesis in population genetics posits that most common SNPs are neutral but this has yet to be formally tested by experimental science. We employed two strategies to determine whether the minor alleles (MAs) of common SNPs are minor because of natural selection. First, we analyzed multiple panels of genetic reference populations or recombinant inbred lines (RILs) in model organisms (yeast, worm, fly, mouse, and rat), and identified the MAs of common SNPs in each panel and the fraction of MAs that each strain carries. We measured the brood size of 104 C. elegans RILs and did genotype-phenotype correlation analysis for the brood trait as well as for ~4700 published and unpublished traits for various RIL panels, most of which were archived at GeneNetwork. Although beneficial to immunity as is expected, more MAs correlated significantly with poor measurements in many adaptive traits, including reproductive fitness, life span, tumor susceptibility, anxiety, depression, startle response, and learning and memory. In addition, more MAs were significantly linked with sensitivity to alcohol, methamphetamine, cocaine, pain, and antipsychotic drugs, and levels in glucose, resistin, insulin, IL-17, iron, and dopamine. The majority of the traits examined did not correlate with MAs, including blood pressure and morphine response. Different MA-linked traits may or may not share the same set of MAs with related traits sharing more MAs than non-related traits. Second, we analyzed 21 published GWAS datasets of common diseases and identified the MAs of common SNPs in each control population and the fraction of MAs each control or case carries. In Europeans or European Americans, more MAs were significantly and repeatedly linked with type 2 diabetes, Parkinson's disease, psychiatric disorders, autoimmune diseases, alcohol and cocaine addictions, lung cancer, less life span, and lower education level achieved, but not hypertension and opiate and marijuana addictions. Thus, the effects of excess MAs in humans are remarkably similar to those in model organisms, suggesting that the link between excess MAs and diseases in humans is causal since it can be replicated in model organisms. These data indicate that most SNPs are functional and open a new avenue of inquiry into the genetic basis of complex traits/diseases. They confirm a self-evident intuition that any system can allow some random errors/noises in building blocks but only to a limit.

2252F

Functional characterization of DcR3 and NF-κB in EBV transformed cell lines from IBD patients of different allelic background and role in disease pathogenesis. *R. Pandey[1], C. J. Cardinale[1], S. Panossian[1], F. Wang[1], E. Frackelton[1], C. Kim[1], F. Mentch[1], R. Chiavacci[1], K. Kachelries[2], S. Grant[1], R. Baldassano[2], H. Hakonarson[1].* 1) Center For Applied Genomics, Children's Hospital Of Philadelphia, Philadelphia, PA 19104; 2) Division of Gastroenterology, Hepatology and Nutrition, Department of Pediatrics, University of Pennsylvania School of Medicine, PA 19104.

Aim: Decoy receptor3 (DcR3), member of the TNFR superfamily is a soluble receptor for FasL and plays significant role in immune suppression and tumor progression by neutralizing the FasL mediated apoptosis signal. Here we investigate the possible immuno-modulation mediated by DcR3 and NF-κB in EBV transformed control and patient derived cell lines with and without risk variants in the TNF Receptor Superfamily 6B gene, TNFRSF6B captured by the tagging SNP, rs2315008. **Methods:** Expression of DcR3 and its kinetics were examined by immunoblot analysis in whole cell lysates from EBV transformed control and patient derived cell lines of different genotype background for rs2315008 (AA, AT, TT). DcR3 induced rapid activation of nuclear factor κB (NF-κB) monitored by immnunoblot analysis of IkBα. DcR3 knockdown was performed using specific DcR3 siRNA. Cell proliferation and cell death was measured by MMT assay. Caspase8, caspase9, caspase3, and Bcl2 expression were analyzed by western blots. **Results:** EBV transformed cell lines derived from IBD patients harboring risk variants in TNFRSF6B (A allele) exhibit differential pattern of DcR3 expression and NF-κB activation kinetics in comparison with wild type. siRNA mediated knockdown post 24hrs of nucleofection results in decreased DcR3 expression, increased cell death and decreased cell proliferation, effects that were also genotype-dependent. **Conclusion:** EBV cell lines from IBD patients harboring risk variants in the TNFRSF6B gene exhibit differential pattern of DcR3 expression and NF-κB activation. Therefore, therapeutic intervention targeting the NF-κB activation pathways represents a promising opportunity for future therapy of IBD.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2253W

Integrative analysis of whole-gene co-expression network and exome sequence data to characterize pathways relevant to rheumatoid arthritis. *H. Nakaoka[1], S. Mitsunaga[2], H. Inoko[2], I. Inoue[1].* 1) Division of Human Genetics, National Institute of Genetics, Mishima, Shizuoka, Japan; 2) Division of Molecular Life Science, School of Medicine, Tokai University, Isehara, Kanagawa, Japan.

New technologies enable to quantify biological measurements on a genome-wide scale (omics data). Systems genomics approach for integrating multiple types of omics data will provide fundamental information to reveal biological pathways associated with disease or phenotype of interest. We propose a systems approach to the integration of whole-gene co-expression networks and exome sequence data to identify susceptibility genes and pathways relevant to rheumatoid arthritis. We constructed multi-tissue gene co-expression networks by using 200 microarrays from synovial tissues with and without rheumatoid arthritis. A clustering algorithm was performed to identify functional modules in these networks. In order to compare the functional modules between normal and diseased networks, a statistical framework was developed. We focused on “aberrant functional modules” that were groups of genes whose expression measurements were tightly correlated in the normal network but the correlations became disrupted in the diseased network. We found some overlaps of genes in the aberrant functional modules with three functional modules identified in our previous work; leukocyte activation and differentiation, pattern-recognition receptor signaling pathway, and chemokines and their receptors. We also found novel pathways such as signal transduction. We performed exome sequencing for 39 patients with rheumatoid arthritis by using Illumina HiSeq2000 and Agilent SureSelect Human All Exon 50Mb Kit. A module-based association study was examined, in which frequencies of rare variants within each aberrant module were compared between 39 affected and 144 unaffected subjects. These results suggest that our integrative approach is useful to identify pathways relevant to disease etiology and to prioritize genes for rare variant association studies at a large scale.

2254T

Integrating genetic linkage analysis and computational prioritization in identifying host genetic factors response to influenza A virus infection. *S. Bao[1], X. Zhou[2], X. Zhang[2], P. Tang[3], J. Zhou[4], Y. Li[5], K. To[4], Y. Song[1].* 1) Department of Biochemistry, The University of Hong Kong, Hong Kong, China; 2) Bioinformatics Division, Tsinghua National Laboratory of Information Science and Technology, Tsinghua University, Beijing; 3) Genomics Core Facility, the Institute for Human Genetics, UCSF; 4) Department of Microbiology, The University of Hong Kong, Hong Kong; 5) Chinese Academy of Sciences, QinDao.

The genetic background of the host has a major role on its resistance to pathogens infection. Several independent studies accumulatively added to the evidence that host genetic make-up may be a key factor in the epidemiology of avian influenza A (H5N1) virus infection in humans. Recently, studies based on mouse models, e.g. quantitative trait loci (QTL) analysis on inbred mouse strains, have shed lights on deciphering the molecular basis of host choice behavior. However, few in silico platforms have been developed to systematically identify genetic factors before they were further experimentally verified. Here, we tried to identify the host factors for influenza virus resistance by integrating genetic linkage analysis and gene prioritization algorithms. We collected all genes within mouse QTLs linked to host response to influenza virus by several independent studies as candidate genes. The genetic polymorphism of the two mouse strains (C57BL/6J, DBA/2J), which exhibited different kinetics after they were infected by influenza A virus, help us filter out candidates which were unlikely to be genetic factors. Totally, we identified 20 genes as candidates by integrating the results from three freely available prioritization tools (Endeavour, ToppGene and SUSPECT). After annotated by the Database for Annotation, Visualization and Integrated Discovery (DAVID), most of them functioned as protein binding genes and participated in the biological processes related to stimulus response, immunity and defense. Interestingly, among the genes expressing non-synonymous coding SNPs on DBA genome, Enpp2 and Tnfrsf11b were both mapped to Bhr2 MGI QTL, which is related to influenza infection consequence. The results indicate that genes identified by our model have great potential as genetic factors in causing host specific response to influenza infection and need to be further studied.

2255F

Molecular Markers of Severe Cerebral Malaria: Expression of Complement and Toll-Like Receptor Pathway Genes Associates with the Malaria Severity in Mali. *R. Sobota[1, 2], J. Manning[2, 3], A. Dara[2, 4], A. Niangaly[4], A. Djimde[4], O. Doumbo[4], A. Kone[4], M. Thera[4], G. Vernet[5], P. Leissner[6], C. Plowe[7].* 1) Human Genetics, Vanderbilt University, Nashville, TN; 2) Fogarty International Scholars Program, NIH, Bethesda, MD; 3) Brigham and Women's Hospital, Boston, MA; 4) Malaria Research and Training Center, Bamako, Mali; 5) Fondation Merieux, Lyon, France; 6) BioMerieux, Grenoble, France; 7) University of Maryland, Baltimore, MD.

Malaria continues to be a significant cause of mortality in the developing world, with over 3 million deaths and 300 to 500 million cases each year. The host response to infection by Plasmodium falciparum, the parasite most often responsible for severe malaria, ranges from asymptomatic parasitemia to death. There are numerous factors that dictate the clinical trajectory of a given infection, including the patient's age, nutritional status and level of P. falciparum transmitted in a given bite. Our study focused on the evaluation of host gene expression in search of molecular markers of severe malaria. In order to study the differential gene expression between those who develop severe versus uncomplicated malaria symptoms, we enrolled children of both sexes between 6 months and 5 years of age, residing in an endemic region of Bandiagara, Mali. A total of 23 uncomplicated and 24 severe cases were enrolled and followed from an acute episode of malaria through the following low transmission season. Affymetrix microarrays were used to obtain expression data on 10 patients (5 severe, 5 uncomplicated) for the acute disease presentation time point, 6 patients (3 sev, 3 unc) for a convalescence time point one week after presentation and following treatment, and 8 patients (4 sev, 4 unc) in the low transmission dry season following disease. KEGG and Ingenuity Pathway analyses showed that the Complement and Toll-Like Receptor pathways were differentially expressed between the severe and uncomplicated groups, with the severe patients showing higher expression of probes for the C1q, TLR2, TLR4, TLR8 and CR1 genes. All of the other significantly differentiated genes previously associated with malaria pathogenesis, GZMB, FOS, HSPA6, were also higher in the severe malaria group. The longitudinal component of the study revealed that in the severe group, the TLR2, TLR4, TLR8, CR1, GZMB, FOS, HSPA6 genes were all expressed at lower levels during dry season in comparison to their uncomplicated counterparts. Assuming that the post disease dry season data correspond to baseline expression levels for these genes, this finding would make them strong candidates for severe disease molecular markers. Furthermore, this study validates findings from mouse models of cerebral malaria not previously shown in humans (TLR2, TLR4, FOS, GZMB), and demonstrates that genes previously shown to govern human uncomplicated infection also determine the extent of severe symptoms (TLR8, CR1, C1q).

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2256W

Cross-ethnic gene co-expression network analysis in human adipose reveals novel triglyceride genes. *D. Weissglas[1], BE. Haas[1], S. Horvath[1,2], KH. Pietiläinen[3,4], RM. Cantor[1], E. Nikkola[1], A. Rissanen[5], M. Civelek[6], I. Cruz-Bautista[7], L. Riba[8], J. Kuusisto[9], J. Kaprio[5,10,11], T. Tusie-Luna[8], M. Laakso[9], CA. Aguilar-Salinas[7], P. Pajukanta[1].* 1) Department of Human Genetics, David Geffen School of Medicine at UCLA, Los Angeles, USA; 2) Department of Biostatistics, David Geffen School of Medicine at UCLA, Los Angeles, USA; 3) Department of Medicine, Division of Endocrinology, University of Helsinki, Helsinki, Finland; 4) Institute for Molecular Medicine FIMM, University of Helsinki, Helsinki, Finland; 5) Department of Psychiatry, Helsinki University Central Hospital, Helsinki, Finland; 6) Department of Medicine, David Geffen School of Medicine at UCLA, Los Angeles, USA; 7) Department of Endocrinology and Metabolism, UNAM, INCMNSZ, Mexico City, Mexico; 8) Department of Molecular Biology and Genomic Medicine, UNAM, INCMNSZ, Mexico City, Mexico; 9) Department of Medicine, University of Eastern Finland, Kuopio, Finland; 10) Department of Mental Health and Substance Abuse Services, National Institute for Health and Welfare, Helsinki, Finland; 11) Department of Public Health, University of Helsinki, Helsinki, Finland.

Background: Unfavorable serum lipid levels, such as elevated serum triglyceride (TG) levels are a well-known risk factor for coronary heart disease (CHD). We hypothesized that searching for TG-correlated gene expression networks in human adipose tissue may reveal novel genes for high TGs, as human adipose tissue is the major storage for fat in the form of TGs. Methods and Results: We measured gene expression in adipose RNAs from two Finnish and one Mexican study sample using either microarrays or RNA sequencing and constructed gene co-expression networks (modules). Significant correlation between serum TG levels and biological modules was observed in each of these study samples. The most correlated TG-modules from each of the three study samples from two distinct populations had a significant overlap of 34 genes ($P<1x10^{-10}$). Eleven of the 34 genes have previous evidence for CHD, type II diabetes, or obesity. Thus, this study provides 23 novel candidate genes for TGs. We also observed more nonsynonymous variants (P=0.034) and overall variants (P=0.018) in these 34 shared genes in subjects with elevated TGs versus subjects with low TGs. Furthermore, two of the 34 genes (ARHGAP9 and LST1) reside in previously reported GWA loci for TG levels, suggesting them as plausible regional candidates underlying the TG GWA signals. Conclusions: This study presents a novel TG biological network with 34 genes shared across three study samples from two distinct populations, Finns and Mexicans. Furthermore, as we utilized two unique methods of measuring gene expression, microarrays and RNA sequencing, these results should be robust to the gene expression measurement method.

2257T

Discovery and replication of nuclear genes influenced by mitochondrial genetic variation. *J. Hall, W. Bush.* Center for Human Genetics Research, Vanderbilt University, Nashville, TN.

It is well known that most human cells contain nuclear and mitochondrial genomes. We also know that mutations in mitochondrial DNA can lead to disease, but the complex interplay between the two genomes and how variation in both genomes contributes to changes in gene expression is relatively underexplored. Previous research has revealed nuclear-mitochondrial interactions in several species, from Saccharomyces cerevisiae (yeast) to Zea mays (corn) to Drosophila melanogaster (fly). To further explore this topic we hypothesized that gene expression levels for some nuclear-encoded genes are correlated with variation in the mitochondrial genome. We estimated the proportion of variance for gene expression traits that is due to all available nuclear and mitochondrial SNPs simultaneously using a mixed linear model (MLM). Using 1000 Genomes and HapMap genotype data (4,099,887 nuclear SNPs; 1,141 mitochondrial SNPs; 129 samples) and previously published gene expression data for Phase II HapMap samples (10,760 genes), we used MLMs to determine if genetic similarity is related to similarity in gene expression, and to partition the proportion of variance in gene expression due to nuclear and mitochondrial components. All models were adjusted for both nuclear and mitochondrial principal components to control for ethnic ancestry. After false discovery rate correction, we identified 175 nuclear genes (1.62%) that are significantly influenced by mitochondrial genetic variation, and we estimate that roughly 2% to 5% of their variance in gene expression is explained by mitochondrial SNPs. Functions of some significant genes include ATB binding, formation of the mitochondrial matrix, mitochondrial transport, response to oxidative stress, metal ion binding, and electron transport. Using an independent set of 165 additional multiethnic individuals with gene expression data from HapMap phase III, we replicate 8 of these associations ($p < 0.05$), though we expect many effects failed to replicate due to differences in frequency of mitochondrial variants across ethnicities. Genes showing significant replication p-values include PLN, NUCB1, LENG4, FSTL4, and GFRA2. Notably, a mitochondrial anchor protein (AKAP1) shows marginal evidence of replication. The results of this analysis illustrate that cumulative genetic variants in the mitochondria may illicit a response from the cellular environment, either directly or indirectly inducing changes to the expression of these genes.

2258F

Polygenic Inheritance of Paclitaxel-Induced Peripheral Neuropathy Driven by Axon Outgrowth Gene Sets. *A. Chhibber[1,2], J. Mefford[3], S.A. Pendergrass[7], R.M. Plenge[4,5,6], M.D. Ritchie[7], E.A. Stahl[4,5,6], J.S. Witte[2,3], D.L. Kroetz[1,2].* 1) Bioeng. & Therapeutic Sci, Univ Cal San Francisco, San Francisco, CA; 2) Institute for Human Genetics, University of California, San Francisco, San Francisco, CA, USA; 3) Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, CA, USA; 4) Program in Medical and Population Genetics, Broad Institute, Cambridge, MA, USA; 5) Division of Genetics, Brigham and Women's Hospital, Boston, MA, USA; 6) Division of Rheumatology Immunology and Allergy, Brigham and Women's Hospital, Boston, MA, USA; 7) Center for Human Genetics Research, Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, TN, USA.

Peripheral neuropathy (PN) is a common and often dose-limiting toxicity for patients treated with paclitaxel, with grade 2 or greater toxicity affecting up to 25% of patients. While diabetes or previous exposure to other chemotherapeutics with similar toxicity increase susceptibility to paclitaxel-induced neuropathy, for the vast majority of individuals there are no known risk factors that predispose patients to the adverse event. Further, pathogenesis for paclitaxel-induced neuropathy is unknown, though studies in animal and cell models have provided some potential mechanisms. Determining whether there is a heritable component to paclitaxel induced PN would be valuable in guiding clinical decisions and may provide insight into treatment of and mechanisms for the toxicity. Using genotype and patient information for the paclitaxel arm of CALGB 40101, a Phase III clinical trial comparing efficacy of single-agent paclitaxel with the current standard regimen as adjuvant therapy for breast cancer in women, we estimated the variance in maximum grade of PN and dose at first instance of PN explained by all autosomal SNPs as well as SNPs in six gene sets selected based on prior knowledge regarding possible mechanisms of the pathogenesis of paclitaxel-induced PN. DNA was isolated and genotyped using the Illumina 610-Quad platform. Following QC, a total of 859 genetic Europeans and 104 African Americans were included in our analyses. The whole genome and pathway-specific heritability analyses were conducted using the GCTA software tool[1]. For the pathway-specific analysis, we extracted all SNPs 10 kb upstream and downstream of autosomal genes in six specific Gene Ontology (GO) sets. Autosomal SNPs from the whole genome or GO SNP sets were then used to generate a relatedness matrix for Caucasians and African Americans, and subsequently used to estimate variance explained for each trait of interest by linear mixed modeling. While whole genome estimates of heritability were not significant, of the six GO sets selected the Axonogenesis GO Term set had significant estimates of heritability close to 20%. These results suggest that paclitaxel-induced neuropathy does in fact have a significant heritable component, and that this heritability is driven by genes involved in axon outgrowth. Disruption of axon outgrowth may be one of the primary mechanisms by which paclitaxel treatment results in PN in susceptible patients.[1]Yang el al.Nature Genetics(2011)43,519-525.

2259W

Association of adrenomedullin gene polymorphisms and lipid levels in Chinese population. *S. Chen, X. Lu, L. Wang, H. Li, J. Huang, D. Gu.* Fuwai Hospital, Beijing, China.

Objective: Dyslipidemia is an important risk factor of cardiovascular disease and is related to morbidity. The high heritability of circulating lipid levels is well established. Whereas the genetic structure of lipid metabolism has not been fully understood. The present study aimed to examine the associations between genetic variants of the adrenomedullin (ADM) gene and lipid levels in Chinese Han population. Methods: A total of 2313 Chinese without history of hypertension, diabetes, cardiovascular disease or cancer were recruited in the present study. Subjects who have current or recent (less than 2 weeks before screening visit) therapy of regulating lipids were excluded. Three tagging single nucleotide polymorphisms (SNPs) based on HapMap CHB data were selected and genotyped. Lipid levels were compared among genotypes of each SNP by general linear model adjusted for multiple covariates including age, gender and body mass index. Results: There were significant associations between the lipid levels and SNPs rs4910118 and rs7944706 after adjusting for covariates of age, gender and body mass index. For SNP rs4910118, total cholesterol level decreased with the number of minor T allele (P=0.03). Total cholesterol level was 192.1±1.0 mg/dl, 188.7±1.3 mg/dl and 185.2±3.4 mg/dl for CC, TC and TT genotypes, respectively. For SNP rs7944706, levels of total cholesterol and low density lipoprotein cholesterol were increased with the number of minor A allele. Total cholesterol level in subjects with GG, GA and AA genotypes was 186.7±1.5 mg/dl, 191.1±1.1 mg/dl and 194.4±1.6 mg/dl (P=0.0019), respectively. Level of low density lipoprotein cholesterol in subjects with GG, GA and AA genotypes was 111.9±1.3 mg/dl, 115.0±1.0 mg/dl and 118.8±1.5 mg/dl (P= 0.0019), respectively. But no association was found between SNP rs4399321 with lipid levels. Conclusion: Genetic variations of the ADM gene were associated with the lipid levels in Chinese population. This finding suggests ADM gene common variations may influence the lipid levels in Chinese. In addition, replications in other populations and further functional studies are required to confirm and interpret the association of ADM with lipids.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2260T

A High Throughput Molecular Assay for Detection of Renal Graft Acute Cellular Rejection Using Peripheral Blood. *M. Mikula[1], O. Zhukov[1], F.M. Hantash[1], N. Park[1], K. Zhang[1], J.R. Rion[1], P. Putheti[2], C. Snopkowski[3], A. Buller-Burckle[1], J. Popov[1], W. Sun[1], B. Crossley[1], M.P. Hernandez-Fuentes[4], S. Sacks[4], M. Suthanthiran[3], T.B. Strom[2], C.M. Strom[1], S.J. Naides[1].*1) Dept Molecular Genetics, Quest Diagnostics Nichols Inst, SanJuanCapistrano, CA; 2) Beth Israel Deaconess Medical Center, Boston, MA; 3) New York Presbyterian Hospital-Weill Cornell Medical Center, New York, NY; 4) King's College London, London, United Kingdom.

Purpose: To develop a molecular test for acute cellular rejection (ACR) in renal grafts. Methods: RNA was extracted from peripheral blood in PAXgene RNA tubes from healthy donors and from transplant patients undergoing renal biopsy for cause, then reverse transcribed for real-time PCR. Targeted markers included cytotoxic T lymphocyte markers (granzyme B, GZMB; perforin, PRF1), regulatory T lymphocyte transcription factor foxP3 (FOXP3), and T lymphocyte and monocyte expressed chemokine IP10 (CXCL10). Constitutively expressed T cell marker CD3E and pan-cell marker ABL1 were endogenous normalizers. Cycle threshold (Ct) values were compared to normal pooled control in 4 replicate reactions via the delta-delta-Ct method to give relative quantity (RQ) for each transcript. Multiple replicates confirmed acceptable intra- and interassay precision. Results: 121 healthy adults without renal, autoimmune or inflammatory disease were tested. Normal donor RQ values were Gaussian in distribution; reference ranges were 95th percentile confidence intervals. Of 21 patients undergoing renal graft biopsy for cause, all 3 with ACR, Banff IIb (n=2) or IIb-IIa (n=1), had elevations in at least 2 analytes normalized to CD3E and a moderate to high elevation in CXCL10 normalized to ABL1 (mean, 12.73 ±15.90). Four with acute tubular necrosis had elevations in 3 or 4 analytes normalized to CD3E, but only 2 had a mild elevation in CXCL10 normalized to ABL1 (mean, 1.28 ±1.15). Patients with antibody-mediated rejection (n=4), chronic cellular rejection (n=2), borderline rejection with Banff III histology (n=2), diabetic nephropathy (n=1), transplant nephropathy (n=1), minimal interstitial fibrosis and tubular atrophy (n=1), or normal histology (n=2) did not show elevations in any marker; 1 patient with antibody-mediated rejection had mild elevations in CXCL10, FOXP3, and GZMB when normalized to CD3E but not when normalized to ABL1. Conclusion: A high-throughput test using molecular biomarkers in peripheral blood is able to detect ACR and to differentiate ACR from other diagnoses known to cause renal graft loss.

2261F

Protein interaction networks reveal novel Autism risk genes within GWAS statistical noise. *C. Correia[1,2,3], Y. Diekmann[1], G. Oliveira[4,5,6], J.B. Pereira-Leal[1], A.M. Vicente[1,2,3], Autism Genome Project.* 1) Inst Gulbenkian Ciencia,Oeiras, Portugal; 2) Instituto Nacional de Saúde Dr. Ricardo Jorge, Lisboa, Portugal; 3) Center for Biodiversity, Functional & Integrative Genomics, Lisboa, Portugal; 4) Unidade Neurodesenvolvimento e Autismo, Centro de Desenvolvimento, Hospital Pediátrico (HP) do Centro Hospitalar e Universitário de Coimbra (CHUC), Coimbra, Portugal; 5) Centro de Investigação e Formação Clinica do HP-CHUC; 6) Faculdade de Medicina da Universidade de Coimbra.

Autism Spectrum Disorder (ASD) is a highly heritable neurodevelopmental disorder. However, genome-wide association studies (GWAS) thus far met limited success in the identification of common risk variants. This suggests that ASD, like most complex diseases, may result from the interaction of many variants with small individual risk, which cannot be detected in single-SNP analysis. Realizing this limitation, new strategies are needed to increase the power of GWAS analysis and reveal risk variants hidden within the statistical “noise”. One such approach is network-based analysis, which has been successfully applied in the analysis of high-throughput expression and sequencing data. In this study we applied a network-based strategy to the Autism Genome Project (AGP) GWAS, combining family-based association data from 2588 ASD families, genotyped using the Illumina 1M SNP, with Human Protein-Protein interaction (PPI) data compiled from public databases. By evaluating sets of genes selected at different gene-wise *P*-value thresholds for network connectivity, we found that genes associated with ASD at the range of GWAS statistical noise are functionally connected beyond random expectation. This observation was replicated in GWAS for six other complex diseases tested. As a proof of principle, using curated lists of candidate genes for each disorder, we further demonstrated that the selection of functionally connected genes based on the largest connected component (LCC) contains relevant disease biology, outperforming the selection of top GWAS genes. Further analysis of the network obtained using the AGP GWAS revealed enrichment in biological processes such as focal adhesion and Nerve Growth Factor signaling pathways. By overlapping the LCCs obtained from two autism GWAS and six other unrelated disease datasets, we highlighted eighteen genes present in both autism samples but in none of the other datasets. These genes are mostly novel candidates, and point toward processes such as axonogenesis and synaptic plasticity. By showing that relevant disease genes are more functionally related, our results demonstrated that network analysis is an effective strategy to extract biological information from GWAS at much higher significance levels. With this approach, we supported a role of cell-adhesion and NGF signaling in ASD, and highlighted novel susceptibility genes that should be further explored for causal variants and therapeutic targeting.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2262W

Systems biology in a non-human primate model: Transcriptomes of normal development in the vervet monkey. *A.J. Jasinska[1], N. Tran[1], C. Blum[1], J. DeYoung[1], O-W. Choi[1], Y. Huang[1], P. Thomas[1], E. Eskin[2,7], K. Dewar[3], J. Wasserscheid[3], G.M. Weinstock[4], W. Warren[4], M. Jorgensen[5], R. Cantor[1,7], R. Woods[6], G. Coppola[1,6], N. Freimer[1].* 1) Center for Neurobehavioral Genetics, Semel Institute for Neuroscience & Human Behavior, University of California, Los Angeles,; 2) Department of Computer Science, The University of California, Los Angeles; 3) Research Institute of the McGill University Health Centre, McGill University and Génome Québec Innovation Centre, and Departments of Human Genetics and Experimental Medicine, 740 Dr-Penfield Avenue, Montreal, PQ, Canada H3A 1A1; 4) The Genome Institute, Washington University School of Medicine, St Louis, MO; 5) Department of Pathology, Section on Comparative Medicine, Wake Forest School of Medicine, Winston-Salem, NC; 6) Dept of Neurology, The University of California, Los Angeles; 7) Department of Human Genetics, The University of California, Los Angeles.

Characterization of transcriptomic variation is emerging as a critical tool for understanding how quantitative trait loci (QTL) contribute to complex phenotypes. Human transcriptomic studies are limited by factors such as the feasibility of invasive tissue collection or variable environmental exposures that can be readily overcome in non-human primate (NHP) models. We characterized transcriptomic variation across multiple tissues and developmental stages and between individuals in 36 vervet monkeys from the Vervet Research Colony extended pedigree. We conducted RNA sequencing across early (7, 90 days, and one year) and later (2.5, 3, and 4+ years old) developmental time points in 6 individuals (3 males, 3 females) at each stage in five tissue types: caudate (brain), two endocrine tissues (pituitary and adrenal) and two peripheral tissues serving as a source of biomarkers (blood and fibroblasts). Comparisons between sexes, across tissue and across developmental stage led to identification of genes and co-expression gene networks showing sex-specific pattern or acting in specific tissues at specific developmental stage. Extending a previous microarray-based approach (Jasinska *et al. Human Molecular Genetics*, 2012) we identified numerous biomarker transcripts, showing high correlations in expression pattern between blood and/or fibroblasts and either adrenal, pituitary or caudate. For example, for 58 transcripts showing highly concordant expression patterns between blood and adrenal we can use blood expression-profiling from large, genotyped vervet population and pedigree samples to obtain insight into the activity of these genes in the adrenal tissue, which is not accessible on a large scale. We are now analyzing whole genome sequencing (WGS) data from all 36 study individuals to identify both cross-tissue and tissue-specific expression QTL (eQTL). Our preliminary data (based on ~8.5 million called variants) indicate that we are able to differentiate the effects on expression of age, tissue, or sex from those due to inter-individual differences. We are now scaling up this study, which ultimately will profile the vervet developmental transcriptome data from birth to adulthood in a larger sample for nearly 100 distinct tissues, providing a unique discovery resource for systems biology.

2263T

Estrogen pathway polymorphisms in relation to Primary Open Angle Glaucoma in the GLAUGEN Study and the NEIGHBOR Consortium. *S. Loomis[1], L.R. Pasquale[1,2], B.L. Yaspan[3], J.H. Kang[2], D. Gaasterland[6], T. Gaasterland[7], R.K. Lee[8], P.R. Lichter[9], Y. Liu[10,12], C.A. McCarty[11], S.E. Moroi[9], L. Olson[3], J.S. Schuman[13], K. Singh[14], D. Vollrath[15], G. Wollstein[13], D.J. Zack[5], R.R. Allingham[10], M.A. Pericak-Vance[8], R.N. Weinreb[16], K. Zhang[16], M.A. Hauser[10,12], J.E. Richards[9], J.L. Haines[3], J.L. Wiggs[1], the NEIGHBOR Consortium.* 1) Ophthalmology, Massachusetts Eye and Ear Infirmary, Boston, MA; 2) Channing Division of network Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA; 3) Center for Human Genetics Research, Vanderbilt University, Nashville, TN; 4) Department of Ophthalmology, University of North Carolina, Chapel Hill, NC; 5) Wilmer Eye Institute, Johns Hopkins University Hospital, Baltimore, MD; 6) Eye Doctors of Washington, Chevy Chase, MD; 7) Scripps Genome Center, University of California at San Diego, San Diego, CA; 8) Bascom Palmer Eye Institute and Human Genomics, University of Miami Miller School of Medicine, Miami, FL; 9) Department of Ophthalmology and Visual Sciences, University of Michigan, Ann Arbor, MI; 10) Department of Ophthalmology Duke University Medical Center, Durham, NC; 11) Essentia Institute of Rural Health, Duluth, MN; 12) Department of Medicine, Duke University Medical Center, Durham, NC;; 13) Department of Ophthalmology, UPMC Eye Center, University of Pittsburgh, Pittsburgh, PA; 14) Department of Ophthalmology, Stanford University, Palo Alto, CA; 15) Department of Genetics, Stanford University, Palo Alto, CA; 16) Department of Ophthalmology and Hamilton, University of California, San Diego, SD, CA.

PURPOSE: Altered estrogen metabolism has been implicated in the pathogenesis of primary open-angle glaucoma (POAG). We assessed the association between estrogen pathway related single nucleotide polymorphisms (SNPs) and specific POAG features. METHODS: We included 3,272 POAG cases and 3,473 controls from the combined Glaucoma Genes and Environment (GLAUGEN) study and NEI Glaucoma Human Genetics Collaboration (NEIGHBOR) consortium in this analysis. We assembled all SNPs from the estrogen metabolism pathway that were available on the Illumina 660W-Quad array platform. We assessed whether this aggregate of SNPs was associated with POAG overall and then by gender using the Pathway Analysis by Randomization Incorporating Structure (PARIS) analysis software package in our combined dataset. In secondary analysis we assessed associations between the estrogen SNP pathway and POAG stratified by intraocular pressure (IOP) $\geq$ 22 mm Hg (HPG) or IOP < 22 mm Hg (NPG) at diagnosis. In addition, we determined which SNPs in the pathway accounted for any significant associations. RESULTS: The estrogen SNP pathway was not associated with POAG, NPG or HPG overall ($p \geq 0.11$). Furthermore the estrogen SNP pathway was not associated with POAG in males (permuted p=1.0) but it was associated with POAG among women (permuted p=0.02). Among women, the estrogen SNP pathway was associated with HPG only (permuted p=0.004). The association between the estrogen SNP pathway and HPG in women was driven by polymorphisms in the following genes: COMT, ESR1, AKR1C4, ESR2, CYP1A2, CYP3A4, SRDSA1 and STS. The COMT SNPs showed the strongest association with POAG across the spectrum of IOP (permuted p<0.001 for HPG and p=0.01 for NPG) among women. CONCLUSIONS: The estrogen SNP pathway was strongly associated with HPG in women and not in men. These data provide insight regarding how common variants in the estrogen SNP pathway contribute to POAG in women.

2264F

Possible roles of miRNA to regulate effective signaling network in controlling cell fate. *M. Nomura[1], S. Kondo[1], T. Nagashima[2], Y. Suzuki[3], M. Okada-Hatakeyama[1].* 1) RIKEN RCAI, Yokohama, Kanagawa, Japan; 2) Center for Cancer Research, Tohoku University, Sendai, Miyagi, Japan; 3) Department of Medica Genome Sciences, University of Tokyo, Kashiwa, Chiba, Japan.

MCF-7 breast cancer cell line is a suitable system to study mechanisms of cell fate controlled by growth factors. Through an activation of shared ErbB receptor signaling pathways, the cell is proliferated in response to epidermal growth factor (EGF) stimulation, while it is differentiated in response to heregulin (HRG). In the previous studies, we showed that early responsive mRNA expression is quantitatively regulated by amplitude and duration of upstream signals, and the qualitative difference in mRNA expression induced by these factors were only observed at late time-points. Notably the HRG-mediated differentiation process is associated with delayed expression of transcription factors that positively and negatively regulate upstream cellular events. These coordinated efforts in a signal-transcription network could be seen in time varying changes of molecular functional groups. To elucidate such cliques in the network, we identified about 400 differentially expressed mRNAs in the EGF/HRG-stimulated MCF-7 cells in five time points and made a signal-transcription network using public databases. In addition to mRNA analysis, we measured miRNA time-course expressions and calculated correlation coefficient to consider temporal effects of miRNAs on their target mRNA expressions. Taking together, we will discuss how expression of miRNA is related to reorganization of cliques in the signal-transcription network in time and stimulus dependent manner.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2265W

Unraveling the molecular mechanisms underlying the teratogenicity of Topamax(Topiramate). *S.K. Rafi, A.J. Olm-Shipman, I. Saadi.* Department of Anatomy and Cell Biology, University of Kansas Medical Center, Kansas City, KS.

Topiramate is a structurally novel anti-epileptic drug (AED) with broad efficacy in epilepsy. It was initially approved by the Federal Drug Administration in 1996 “as an adjunctive therapy for adults with partial onset seizures”, and later, “for prevention of migraine headaches”. Topiramate is also prescribed (off-label) for bipolar disorders, bulimia nervosa, antipsychotic-induced weight gain, and alcohol abuse. Recently, topiramate use during pregnancy has been linked to a significantly increased risk of orofacial clefts (OFC). Approximately 30,000 children are born to epileptic mothers each year. Consequently, in March 2011, the FDA released a warning about the use of topiramate during pregnancy. However, the exact molecular mechanism of the teratogenicity of topiramate is unknown. We want to ascertain these molecular perturbations due to topiramate use to determine whether the beneficial or off-target effects lead to the observed teratogenicity. If the off-target effects are to blame, then future AEDs can be designed to possibly avoid these adverse consequences. The effect of topiramate involves inhibition of several ion channels, including the voltage-gated Na+, Ca+, and K+ ion channels, such as AMPA / GluR5 kainate receptors and ligand-gated non-NMDA glutamate receptors as well as variable modulation of GABA-activated ion channels and intracellular GABA concentration. Although GABA signaling has been implicated in the pathogenesis of OFC, our Ingenuity Pathway analysis of the gene expression microarray data, previously collected by Lamb et al. (Science, 2006), indicates no significant down-regulation of the GABA signaling pathway following topiramate treatment of MCF-7 cells. Instead, the two most significantly up-regulated canonical pathways are TGF-β signaling and retinoic acid receptor activation. Individually, topiramate significantly affects the expression (>1.5 fold) of a large number of other known OFC-related genes (26 up-regulated, 8 down-regulated). We are now looking at the effect of topiramate on actin cytoskeletal remodeling and on phosphorylation changes downstream of several cell-signaling cascades. These studies are likely to unravel critical molecular clues for the increased occurrence of OFC and other teratogenic effects following topiramate use in pregnancy.

2266T

The CARTaGENE Genomics Project: Systems Genetics of Cardiometabolic Phenotypes. *Y. Idaghdour, J-P. Goulet, J. Hussin, A. Hodgkinson, E. Gbeha, J-C. Grenier, T. de Malliard, V. Bruat, E. Hip-Ki, C. Boileau, Y. Payette, P. Awadalla.* Sainte-Justine Research Center, University of Montreal, Montreal, QC, Canada.

An area of fundamental biomedical research that merges population and quantitative genomics is the identification of transcriptional and other intermediate biomarkers for disease susceptibility and disease status. These biomarkers are quantitative traits whose architecture is modulated through mechanisms that can incorporate genetic and environmental cues. But what are the relative magnitudes of these effects, and how do gene expression profiles correlate with cardiometabolic phenotypes? To address these questions, we generated whole-genome genotypic data and whole-blood RNA-Seq gene expression profiles from 800 individuals having high and low cardiovascular disease risk scores from an aging cohort of 20,000 participants in Québec, Canada. For each participant, clinical material and medical information spanning a wide range of medically relevant phenotypes, including data from a well characterized and extended pedigree, were collected. We documented latent structure in gene expression profiles and transcriptional signatures that best correlate with cardio-metabolic phenotypes. Furthermore these analyses revealed the spectrum of allelic expression and regulatory variation associated with the clinical traits studied.

2267F

Association of insulin-like growth factor-1 gene polymorphism in Kuwaiti Arab patients with rheumatoid arthritis. *M. Haider[1], G. S. Dhaunsi[1], S. S. Uppal[2].* 1) Dept Pediatrics, Fac Med, Kuwait Univ, Safat, Kuwait; 2) Dept Medicine, Fac Med, Kuwait Univ. Safat, Kuwait.

Insulin-like Growth factor-1 (IGF-1) regulates cellular proliferation, differentiation, survival and regeneration, and low plasma levels of IGF-1 have been suggested to contribute towards pathogenesis of rheumatoid arthritis (RA). In view of the biological significance of IGF-1, we investigated a possible association of RA with polymorphism of a 192-bp allele which is cytosine-adenosine repeat located 1kb upstream from the IGF-1 gene transcription site and known to regulate serum IGF-1 levels. Blood samples were collected from 52 healthy controls (HC) and 68 RA patients to measure the levels of IGF-1 and to isolate genomic DNA. Polymorphism of IGF-1 gene was examined using polymerase chain reaction. Severity, duration, and activity of the disease were recorded for all RA patients. We observed that 97% of all the subjects that participated in this study showed the presence of a 192-bp allele of IGF-1 gene. All healthy controls exhibited the presence of 192-bp wild type allele. All non-carriers of the 192-bp allele had RA. Gender correlated significantly with allele frequencies as 14% of the male and only 2% of the female RA patients were non-carriers of 192-bp allele. Plasma IGF-1 levels were significantly lower ($p<0.01$) in RA patients as compared to HC, and all RA patients that were non-carriers of 192-bp allele had a significantly high disease activity score. Our data from Kuwaiti Arabs demonstrates an association of IGF-1 gene polymorphism with risk of developing RA in non-carriers of IGF-192-bp allele.

2268W

Variations within ARID5B, CEBPE, and IKZF1 and Risk of Childhood Leukemia. *L. Hsu[1], A. Chokkalingam[1], C. Metayer[1], J. Wiemels[2], P. Buffler[1].* 1) School of Public Health, University of California, Berkeley, Berkeley, CA; 2) Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, CA.

Leukemia is the most common childhood malignancy, accounting for 31% of all cancers diagnosed in children younger than 15 years old. Two genome-wide association studies found that genes involved in B-lymphocyte development, including IKZF1 (7p12.2), ARID5B (10q21.2), and CEBPE (14q11.2) are associated with ALL predisposition among the Caucasian population. Here, we attempted to validate selected single nucleotide polymorphisms (SNPs) from previous GWAS in both Caucasian and Hispanic populations in the California Childhood Leukemia Study (CCLS). In addition, potential gene and environment interactions between candidate genes and early life infection experiences (early daycare attendance, birth order, and ear infection) were assessed. The study population is comprised of 525 non-Hispanic White and Hispanic ALL cases and 775 controls from the CCLS. We examined the associations between 8 top SNPs identified in previous GWAS and childhood ALL risk and gene environment interactions. Logistic regression assuming a log-additive genetic model of inheritance was used to estimate odds ratios (OR) associated for each SNP, adjusting for age, sex, race and Hispanic status. In the Caucasian population, results for SNPs in ARID5B, CEBPE, and IKZF1 were consistent in terms of both significance and effect direction with previous GWAS findings after accounting for multiple testing. In the Hispanic population, although a similar relationship with disease risk was observed for all 5 ARID5B SNPs, the associations were not consistent for the SNPs in CEBPE and IKZF1. We also examined the associations within ALL subtypes. 5 ARID5B SNPs showed stronger association with ALL when the analysis was confined to B-cell ALL and B-cell hyperdiploidy ALL. CONCLUSIONS: The results from our study confirm that previous GWAS-identified SNPs in B-cell development genes are associated with childhood ALL risk in the Caucasian population. However, only SNPs on ARID5B are associated with ALL risk in the Hispanic population. The findings indicate that susceptibility in the Hispanic population differs from that of the non-Hispanic population, due possibly to differences in genetic structures and/or differences in environmental exposure profiles that modulate the effects of risk variants. Further investigations are needed to both fine-map the gene regions for identification of the causal loci, and to identify environment factors that may modulate the effects of these loci.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2269T

Contribution of common *PCSK1* genetic variants to obesity in a sample of 8,359 subjects from two multi-ethnic American populations. *H. Choquet[1], J. Kasberger[1], A. Hamidovic[2], M. Fornage[3], E. Jorgenson[1].* 1) Ernest Gallo Clinic and Research Center, University of California, San Francisco, Emeryville, CA; 2) Department of Preventive Medicine, Northwestern University, Chicago, IL; 3) Brown Foundation Institute of Molecular Medicine, University of Texas Health Science Center, Houston, TX.

Common *PCSK1* variants (notably rs6232 and rs6235) have been shown to be associated with obesity in European and Asian populations. To determine whether common *PCSK1* variants contribute to obesity in American populations, we conducted association analyses in 8,359 subjects from two multi-ethnic American studies: the Coronary Artery Risk Development in Young Adults (CARDIA) study and the Multi-Ethnic Study of Atherosclerosis (MESA). By exploring the full *PCSK1* locus, we observed that no variant remained significant after correction for multiple testing. Furthermore, we found no significant association for the rs6232 and rs6235 *PCSK1* variants with obesity by combining all subjects in a meta-analysis. However, the rs6232 and rs6235 variants have been shown to be associated with obesity in CARDIA study (OR=1.65 [1.08-2.51], *P*=0.02; OR=1.22 [1.02-1.46], *P*= 0.03 for the rs6232 and rs6235 respectively). In addition, we observed a nominal effect (OR=0.88 [0.78-0.99], *P*=0.03) for rs6235 in MESA study but in a direction inconsistent with the previous study. By evaluating the contribution of the two above-mentioned *PCSK1* variants in each ethnic group, we found that only the rs6232 variant was associated with BMI (*P*=0.006) and obesity (*P*=0.018) in European-American subjects from the CARDIA study. Moreover, the rs6232 variant was found to increase the obesity incidence during the 20 years of follow-up (HR=1.53 [1.07 - 2.19], *P*= 0.019) in European-American subjects from the CARDIA study. Alternatively, only the rs6235 variant was associated with BMI (*P*=0.028) and obesity (*P*= 0.018) in African-American subjects from the CARDIA study. These results indicate that common *PCSK1* variants (notably rs6232 and rs6235) contribute modestly to obesity in multi-ethnic American populations. Further, these results suggest that the association of rs6232 and rs6235 variants with obesity may be ethnicity-specific. However, multiple replication studies in multi-ethnic American populations are needed to confirm our findings.

2270F

Association between IRF6 gene and non-syndromic cleft lip and/or palate in Brazil. *T. Felix[1], L.T. Souza[1], T.W. Kowalski[1], I.L. Monlléo[2], E.M. Ribeiro[3], J. Souza[4], G.F. Leal[5], A.C. Fett-Conte[6], V. Gil-da-Silva-Lopes[7].* 1) Servico de Gen Medica, Hosp Clinicas Porto Alegre, Porto Alegre, RS, Brazil; 2) Medical Genetics Sector, State University of Alagoas, Maceio, AL, Brazil; 3) Medical Genetics Sector, Hospital Infantil Albert Sabin, Fortaleza, CE, Brazil; 4) Medical Genetics Sector, Assistance Center for Cleft Lip and Palate (CAIF), Curitiba, PR, Brazil; 5) Center for Face Defects (CADEFI), Recife, PE, Brazil; 6) Genetics Service, Faculty of Medicine of São José do Rio Preto, São José do Rio Preto, SP, Brazil; 7) Department of Medical Genetics, University of Campinas, Campinas, SP, Brazil.

Oral cleft (OC) is the most common craniofacial birth defect in humans. Cleft lip and palate (CL/P) have a multifactorial inheritance including both genetic and environmental factors. IRF6 gene is strongly expressed in the leading edge ectoderm of the palatal shelves prior to and during formation of the secondary palate. Studies have shown that variants in IRF6 gene are responsible for 12% of OC cases. The rs2235371 (V274I) polymorphism is in linkage disequilibrium with rs642961 polymorphism. This polymorphism was previously suggested to cause disruption of the binding site of transcription factor AP-2α, and to directly influence the risk of CL/P. The aim of this study was to evaluate the association of rs2235371 and rs642961 IRF6 gene polymorphisms in CL/P in Brazil. The SNPs were analyzed using TaqMan assay (Applied Biosystem). Statistical analysis was performed with the transmission disequilibrium test (TDT) using Family Based Association Test (FBAT) software. We analyzed 171 case-parent triads. Minor allele frequency of rs2235371 A allele was 0.080 and 0.178 for rs642961A allele. TDT analysis showed an overtransmission of G allele rs2235371 (p=0. 002) however we did not find association of rs642961 or combined haplotypes. This data confirms previous findings of association of rs2235371 G allele in the clefting population.

2271W

Validation of the *PPP1R12B* as a candidate gene for childhood asthma susceptibility. *M.B. Freidin[1], A.V. Polonikov[2].* 1) Research Institute of Medical Genetics, Tomsk, Russian Federation; 2) Kursk State Medical University, Kursk, Russian Federation.

Genome-wide association studies (GWAS) are a powerful tool for revealing positional candidate genes of complex diseases and traits. A number of such the studies have been performed in childhood and adult bronchial asthma (BA), which allowed disclosing dozens of novel candidate genes. However, given the high genetic heterogeneity of BA, a validation of the GWASs in independent populations is of critical importance. Recently, the *PPP1R12B* gene was revealed as a new candidate gene for childhood asthma in GWAS in Russians of West Siberia using Illumina 610QUAD chip (Mol Biol. 2011;45(3):464-72). We now set out to validate this finding in Russian population of Kursk. One hundred and fourteen patients with childhood BA (age-of-onset up to 16 years) and 279 healthy controls were genotyped using same Illumina chip. After quality assessment, total of 26 markers embracing the *PPP1R12B* gene region were analyzed using linear models approach to measure the associations between the disease and markers. Four markers (rs17438212, rs12734001, rs3767423, and rs3817222) were found to be significantly associated with childhood BA after correction for multiple testing. They represent two relatively distinct linkage disequilibrium (LD) blocks with two markers in each. The odds ratios for the associations varied between 1.78 and 2.04 (Padj = 0.029 - 0.03). The direction of association and the magnitude of the effect for the rs12734001 and rs3817222 markers were in accordance with initial finding in Russians of West Siberia. A haplotype comprised common alleles of four SNPs was protective against asthma, while two haplotypes containing at least one of pair of rare alleles from every LD block increased risk of the disease. Thus, the present data confirmed that the *PPP1R12B* gene is associated with childhood BA in Russians. Further studies are required to provide the functional basis for the association.

2272T

Association of the SNP rs 9939609 of the fat mass and obesity associated gene (FTO)with the risk for obesity in children from Yucatan, Mexico. *L. Gonzalez-Herrera, N. Mendez-Dominguez, F. Herrera-Sanchez, G. Uicab-Pool, V. Hernandez-Escalante, L. Ruiz-Ortiz, G. Storey-Montalvo, J. Zavala-Castro.* Campus de la Salud, Univ Autonoma de Yucatan, Meridal, Yucatan, Yucatan, Mexico.

Obesity in children is the first cause of morbidity in Yucatán, Mexico. Genetic variation plays a major role in determining the susceptibility to the obesogenic environment. Previous studies have identified the common variant rs9939609 in the FTO gene, which represents a T to A change. Significant associations of the SNP rs9939609 with BMI and with the risk for obesity have been suggested in homozygotes, who weigh about 3-4 kg more. Each risk allele A increases the BMI by the equivalent to ~0.40-0.66 kg/m2 in BMI. The central role of FTO might be through an effect on cerebrocortical insulin sensitivity as homozygous have a reduced insulin response in the brain. However, some other studies have reported lack of association of FTO rs9939609 with risk of obesity. In this study we evaluated the association of the rs9939609 of FTO gene with the risk for obesity in children from Yucatan, Mexico. We included 97 obese children with body mass index (BMI) > 25kg/m2 as cases and 73 healthy children with BMI < 25 kg/m2; in order to perform a case-control association study. Total cholesterol, HDL-cholesterol, LDL-cholesterol, and triglycerides were measured in fasting conditions. Genomic DNA was extracted and the SNP rs9939609 was determined by real time PCR with Taqman probes (ID: C_ 30090620_10 Applied Biosystems). Genotype and allele frequencies between cases and controls were compared using SNPstats software. Genotype and allele frequencies were distributed according to Hardy-Weinberg expectations (p>0.05) in cases and controls. The three rs9939609 genotypes were similarly distributed between obese and non-obese children (p>0.05). The homozygous TT wildtype genotype was the most frequent, whereas the homozygous AA genotype was the least frequent in the studied population. Significant differences were not found when comparing genotype and allele frequencies between cases and controls for any allele nor genotype for the rs9939609 (p>0.05), suggesting that this SNP is not associated with the genetic risk for child obesity in the population of Yucatan. Lipid profile was not significant different between obese and non-obese children, except for triglycerides, which was higher in obese children. The mean levels of the measured lipids were not correlated with any rs9939609 genotype (p>0.05) in the studied population. The SNP rs9939609 is not associated with the genetic risk for obesity, nor with the lipid profile in obese children from Yucatan, Mexico.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2273F

Shared genetic loci between leprosy and inflammatory bowel disease in Korean patients. *Y. Jung[1], M. Hong[1], J. Baek[1], S. Yang[2], K. Song[1].* 1) Biochem and Molecular Biol, Univ of Ulsan College of Medicine, Seoul, South Korea; 2) Dept Internal Medicine, Univ Ulsan Col Medicine, Seoul, Korea.

Crohn's disease (CD) and ulcerative colitis (UC) are inflammatory bowel diseases of unknown cause. It has been suggested that genetic and environmental factors cause dysregulated mucosal immune responses to the gut flora in susceptible individuals. Mycobacterium avium subspecies paratuberculosis has been implicated in CD. It has shown to cause bowel inflammation in ruminants. It is of interest that significant overlap of susceptibility genes is present between CD and leprosy, an infectious disease caused by Mycobacterium leprae. Recent genome-wide association study of leprosy conducted in the Chinese population identified eight susceptibility loci including CCDC122, LACC1(C13orf31), NOD2, TNFSF15, HLA-DR, IL23R, RAB32, and RIPK2 (Zhang et al. 2009, 2011). Of those, LACC1, NOD2, TNFSF15, HLA-DR, and IL23R were shown to be associated with CD susceptibility in Caucasians, and TNFSF15 and HLA-DR in East Asians, respectively. The goal of this study was to test a potential overlapping of the rest of the leprosy susceptibility genes including CCDC122, LACC1 and RIPK2 with CD or UC in Korean patients. Five single nucleotide polymorphisms from the 3 loci including CCDC122, LACC1, and RIPK2 were genotyped in 631 patients with CD, 636 patients with UC, and 597 healthy controls. Statistically significant association with CD was found at LACC1 (rs3764147, P = 2.4 x 10-4, OR = 1.36 (95% CI = 1.16-1.61) under the allelic model). Rs42490 of RIPK2 showed significant association with UC (rs42490, P = 0.009, OR = 1.62 (95% CI = 1.12-2.33) under the recessive model). Our data suggest that LACC1 and RIPK2 variants might be associated with pathogenesis of CD and UC in Koreans, respectively. Our data may support pathobiological roles of intestinal flora for developing intestinal bowel diseases in genetically susceptible individuals.

2274W

Fatness related allelic variant near insulin receptor substrate 1 (IRS-1) gene is associated with low low-density lipoprotein cholesterol in childhood and a better cardiometabolic profile in their adulthood: The Bogalusa Heart Study. *S. Li[1], W. Chen[1], E.N. Smith[2], S.S. Murray[3], C. Fernandez[1], J. Xu[1], S.R. Srinivasan[1], N.J. Schork[4], G.S. Berenson[1].* 1) Tulane Center for Cardiovascular Health and Department of Epidemiology, Tulane School of Public Health and Tropical Medicine, New Orleans, LA; 2) Department of Pediatrics and Rady's Children's Hospital, School of Medicine, University of California at San Diego, La Jolla, CA; 3) Department of Pathology, School of Medicine, University of California at San Diego, San Diego, CA; 4) The Scripps Research Institute and the Scripps Translational Science Institute, La Jolla, CA.

Recent genome-wide association studies have identified variants near insulin receptor substrate 1 (IRS-1) gene that are associated with body fatness, type 2 diabetes, and various metabolic traits in adults. However, limited information is available in this regard in children. We examined the effects of one variant near IRS-1, rs1399627, which is in perfect linkage disequilibrium with the variant reported to be associated with body fatness, on metabolic profiles in childhood (average age=12.0 years) as well as in adulthood (average age=36.5 years) in a community-based, longitudinal cohort of the Bogalusa Heart Study (n=1,196). The fatness increasing allele (T) of rs1399627 was associated with lower levels of low-density lipoprotein (LDL) cholesterol in children (P<0.001); associated with lower levels of LDL cholesterol (P=0.07), triglycerides (P=0.002), insulin (P=0.06), and glucose (P=0.05), and higher levels of high-density lipoprotein (HDL) cholesterol (P= 0.07) and adiponectin (P=0.007) in adulthood. Further, the T allele was associated with a better metabolic profile, with 20.7% of non-T carriers versus 27.6% of T-carriers having no metabolic abnormalities (according to levels of HDL cholesterol, LDL cholesterol, triglycerides, glucose, systolic and diastolic blood pressures, C-reactive protein, and adiponectin) in adults (P=0.03). In conclusion, fat-increasing allelic variants near IRS-1 gene are associated with lower levels of LDL cholesterol in childhood and a better metabolic profile in adulthood.

2275T

Genetic Variants Associate with Nonalcoholic Fatty Liver Disease (NAFLD) in African and Hispanic Americans. *S.K. Musani[1], N.D. Palmer[2], M.F. Feitosa[3], M.A. Jhun[4], L.M. Yerges-Armstrong[5], R. Hernaez[6], J. Liu[1], B. Kahali[7], J.J. Carr[8], T.B. Harris[9], A.V. Smith[10], L.E. Wagenknecht[11], I.B. Borecki[3], E.K. Speliotes[7], L.F. Bielak[4] for the Genetics of Obesity-Related Liver Disease (GOLD) Consortium.* 1) Jackson Heart Study, University of Mississippi Medical Center, Jackson, MS; 2) Department of Biochemistry, Wake Forest School of Medicine Winston Salem, NC; 3) Department of Genetics, Washington University, Saint Louis, MO; 4) Department of Epidemiology, University of Michigan School of Public Health, Ann Arbor MI; 5) Department of Medicine, Division of Endocrinology, Diabetes and Nutrition - University of Maryland School of Medicine, Baltimore MD; 6) Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD; 7) Department of Internal Medicine, Division of Gastroenterology and Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI; 8) Departments of Radiologic Sciences, Internal Medicine-Cardiology and Public Health Sciences, Wake Forest School of Medicine, Winston-Salem, NC; 9) National Institute on Aging, National Institutes of Health, Bethesda, MD; 10) Icelandic Heart Association, Kopavogur IS-201, Iceland; 11) Division of Public Health Sciences, Wake Forest University School of Medicine, Winston-Salem, NC.

NAFLD is an obesity related condition that affects > 50% of individuals in some populations and is expected to become the leading cause of liver disease around the world by 2020. NAFLD is a spectrum of disease that includes hepatic steatosis (fat in liver cells), steatohepatitis (inflammation around the fat), and fibrosis/cirrhosis (scarring of the liver). NAFLD is more prevalent in individuals of Hispanic than European than African ancestries. Association of genetic polymorphisms in or near five genes with NAFLD has been replicated in individuals of European ancestry. Here we investigate whether these polymorphisms are associated with NAFLD in African and Hispanic Americans. As in individuals of European ancestry, hepatic steatosis was quantified using computed tomography and then inverse normally transformed. We used an additive genetic model to examine the association of hepatic steatosis with five single nucleotide polymorphisms (SNPs) in or near *PNPLA3*, *PP1R3B*, *NCAN*, *GCKR* and *LYPLAL1* in five cohorts of African Americans (n=3,124) and one cohort of Hispanic Americans (n=849) controlling for age, age^2, sex, number of alcoholic drinks, principal components for population stratification, and relatedness in the family studies, using regression. We combined results across studies using fixed-effects meta-analysis in METAL. We also estimated heritability of hepatic steatosis in the family studies using SOLAR. Heritability of hepatic steatosis ranged from 0.20 - 0.34 in these cohorts. Variants in or near *PNPLA3* (p=5.1E-18), *NCAN* (p=4.3E-03), *GCKR* (p=5.1E-02), *PPP1R3B* (p=3.3E-06) in African, and *PNPLA3* (p=2.0E-07) and *PPP1R3B* (p=2.5E-02) in Hispanic Americans had statistically significant, directionally consistent effects as those in the European ancestry discovery study. Total variance explained by the five SNPs across these five loci was 5.13% vs. 4.76% vs. 4.24% in individuals of Hispanic, European, and African ancestries. Fine mapping of these loci in African Americans highlights missense variants at *PNPLA3* and *GCKR*and better defines the association region at *LYPLAL1*. Some genetic variants that predispose to NAFLD are shared across ancestries. Missense variants identified in *PNPLA3* and *GCKR*during cross-ancestry analyses are likely functional.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2276F

Genetic association to ERAP1 in psoriasis is confined to disease onset after puberty and not dependent on HLA-C*06. *P. Nikamo[1], J. Lysell[1], L. Padyukov[2], I. Kockum[3], M. Ståhle[1].* 1) Department of Medicine, Dermatology Unit, Karolinska Institutet, Stockholm, Sweden; 2) Department of Medicine, Rheumatology Unit, Karolinska Institutet, Stockholm, Sweden; 3) Department of Medicine, Neuroimmunology Unit, Karolinska Institutet, Stockholm, Sweden.

The genetic background of psoriasis is strong and several susceptibility genes have been identified with HLA-Cw*0602 remaining the strongest candidate. Precisely how HLA-C contributes to psoriasis is however unclear. Evidence for an interaction between HLA-C and endoplasmic reticulum aminopeptidase 1, ERAP1, confined to individuals carrying the HLA-C risk allele was recently presented. ERAP1 plays a crucial role in MHC class I peptide processing and is involved in cell surface shedding of proinflammatory cytokine receptors, altogether suggestive of a potential involvement in psoriasis pathogenesis. Psoriasis displays wide clinical variation and genetic heterogeneity likely contributes substantially to clinical diversity. Therefore, stringent stratification for sub-phenotypes may be crucial in genetic analyses. Age of onset is an obvious phenotype and separating psoriasis into type I (onset <40 years of age) and Type II (> 40 years of age) has been useful. To sharpen the age dependent phenotype, we compared genotypes for ERAP1 (rs26653, rs30187, rs27524) and HLA-C*06:02 in healthy controls and cases stratified for onset of psoriasis <10, 10-20, 20-40 and >40 years of age. This approach revealed that association to ERAP1 was confined to cases with onset between 10-20 years (OR 1.59, CI 1.28-1.98, p=0.00008) and no association was detected in cases with onset below 10 years, reflecting genetic heterogeneity within the childhood psoriasis population. In contrast to earlier findings, association to ERAP1 was neither dependent on nor interacting with HLA-C*06:02. ERAP1 SNP rs26653, previously not reported in psoriasis, is non-synonymous, has suggestive functional consequences and herein displays strong association to disease.

2277W

On the association of common polygenic variation with body mass index across adolescent development: A longitudinal twin study. *R.E. Peterson, B.T. Webb, E.C. Prom-Wormley, J.L. Silberg, L.J. Eaves, H.H. Maes.* Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, Richmond, VA.

There have been few genetically informative longitudinal studies examining the stability of body composition over time. However, there is some evidence to suggest additive genetic effects associated with BMI are relatively stable across adolescence as indicated by genetic cross-time correlations upwards of 0.80. In contrast, the environmental cross-time correlations have been shown to decrease over time and common environment has been shown to be important in early but not late adolescence or adulthood. Furthermore, genome-wide association studies (GWAS) of BMI using large-scale adult samples have yielded 32 robustly associated variants. Of which 23 were associated with BMI in a cross-sectional analysis of children and adolescents. Further research needs to address when in development these variants become important for predicting BMI. Given the dramatic increase of obesity prevalence in both children and adults in developed countries and the numerous negative consequences associated with elevated body weight, future research is warranted to understand the genetic and environmental contributions to BMI trajectories across adolescence. Therefore, the purpose of this research is to utilize a developmental twin study design to 1) test for changes in the relative contributions of genes and environment, 2) test for shared genetic and environmental liability across time, 3) identify factors driving change versus maintaining stability and 4) evaluate association of validated BMI-SNPs across adolescence. BMI was calculated from weight and height collected on up to four waves and ages ranged from 8 to 18 in 2,794 twin participants from the Virginia Twin Study of Adolescent Behavioral Development (ABD). Preliminary results suggest that SNPs in FTO and TNN13K are associated with BMI from age 11 through development and a SNP in MC4R is associated from age 17. Understanding obesity development will aid in identifying obesogenic vulnerability time-points and facilitate targeted prevention and treatment efforts.

2278T

Absence of the chromosome 16p11.2 duplication in metabolically normal, constitutionally thin individuals with extremely low BMI. *J. Hager[1], N. Germain[2], E. Meugnier[3], B. Galusca[2], H. Vidal[3], B. Estour[2].* 1) Human Genetics, Nestle Institute of Health Sciences, Lausanne, Vaud, Switzerland; 2) Dept. of Endocrinology, Centre Hospitalier Universitaire, Saint Etienne, France; 3) UMR INSERM U-449/INRA-1235, Faculty of Medecine R Laennec, Lyon, France.

We recently described a phenotype termed constitutional thinness (CT), for individuals with extreme low BMI (<17 kg/m2) but normal metabolic profile without any cognitive or psychiatric abnormalities (e.g. anorexia nervosa or other abnormal food-intake behaviors). Constitutionally thin individuals have normal food intake are remarkably resistance to weight gain. Jacquemont et al. recently reported the association of a deletion/duplication variant on chromosome 16p11.2 with extreme high and low BMI's respectively (Nature, 6 October 2011, Vol. 478: 97-102). Given the similarly low BMI in CT individuals it is important to determine if the reported duplication associated with low BMI is also present in the CT phenotype. We used genome-wide genotype data to screen the 16p11.2 area for the reported duplication CNV in constitutionally thin and normal weight individuals. Apart from one normal weight individual, none of the individuals showed evidence for a duplication in the reported chromosomal area. We also evaluated the expression profiles of the genes in the duplication interval in adipose tissue. Again, contrary to reported expression profiles by Jacquemont et al. we did not find any evidence for differential over-expression of genes in the duplication area, substantiating the absence of any duplication in the CT individuals. In summary we did not find evidence that CT individuals carry duplications of the chromosomal region 16p11.2. Constitutionally thin individuals are characterized by a stable very low BMI, comparable to the BMI reported by Jacquemont et al. but are metabolically normal and don't show any of the psychiatric abnormalities that constituted the selection criteria their study. Jacquemont et al. ascertained individuals primarily for a variety of cognitive and psychiatric phenotypes rather than BMI. Indeed this duplication was shown previously to be associated with both autism and schizophrenia. It may be that the association with low BMI of the individuals reported by Jacquemont et al. is secondary to the psychiatric profile and abnormalities in feeding behavior rather than a susceptibility to low BMI per se. In conclusion we believe that our additional findings are important and need to be considered when reporting on associations with BMI. Also it highlights the importance of careful phenotypic description to distinguish extremes in weight distribution due to abnormal food-intake behavior from those due to differences in metabolic state.

2279F

Lack of association of genetic variants in *PROX1*, *ADCY5*, *SLC2A2*, *DGKB-TMEM195*, *ADRA2A*, *CRY2*, *FADS1*, *IGF1* and *C2CD4B* with glucose- and insulin-related traits in Chinese adolescents. *R.C.W. Ma, C. Tam, G. Jiang, J. Ho, V. Lam, Y. Wang, H.M. Lee, A. Kong, W.Y. So, J. Chan.* Department of Medicine and Therapeutics, The Chinese University of Hong Kong, Hong Kong, Hong Kong.

Objective - In this study, we aim to examine the associations of previously reported genetic variants for fasting glucose or insulin levels in Chinese adolescents.

Methodology - Nine single nucleotide polymorphisms (SNPs) from *PROX1*, *ADCY5*, *SLC2A2*, *DGKB-TMEM195*, *ADRA2A*, *CRY2*, *FADS1*, *IGF1* and *C2CD4B* were genotyped in 975 adolescents [age mean ± SD = 15.5 ± 1.9 years, % males = 46.6] that participated in a health screening program. Associations of SNPs at additive model with glucose-and insulin-related quantitative traits including fasting plasma glucose and insulin (FPG and FPI), as well homeostasis model assessments of insulin resistance and beta-cell function (HOMA-IR and HOMA-β) were tested by multivariate linear regression adjusted for age, sex and body mass index.

Results - We did not observe any significant association between SNPs and quantitative traits, except that the A-allele of rs11071657 in *C2CD4B* revealed trends of reducing FPI (P = 0.092) and HOMA-IR (P = 0.044). In addition, five and seven SNPs showed similar direction of associations for FPG and HOMA-β to earlier reports, respectively.

Conclusions - Our findings do not support association of the reported loci with glucose- and insulin-related traits in Chinese adolescents. *Acknowledgement* - This work was supported by a CUHK Direct Grant (2041567) and the Innovation and Technology Fund (ITS/487/09FP).

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2280W

Combined Effects of Genetic Variants identified from Genome-wide Association Studies Increased Risk of Type 2 diabetes in Chinese. *C. Tam, J. Ho, Y. Wang, V. Lam, M. Ng, R. Ma, W. So, J. Chan.* Department of Medicine & Therapeutics, The Chinese University of Hong Kong, Hong Kong, Hong Kong.

Background - Type 2 diabetes (T2D) is a complex disease resulting from the interaction of multiple genetic variants and environmental factors. Recent meta-analysis of genome-wide association studies (GWAS) identified novel loci increasing risk for T2D, which were widely reproduced in European and Asian populations. In the present study, we aim to investigate their univariate and combined genetic effects on T2D in Chinese population.

Methodology - We genotyped 14 single nucleotide polymorphisms (SNPs) from β-cell function-related genes including *NOTCH2*, *ADAMTS9*, *IGF2BP2*, *WFS1*, *CDKAL1*, *JAZF1*, *SLC30A8*,*CDKN2A/B*, *HHEX*, *TCF7L2*, *KCNQ1*, *KCNJ11*, *TSPAN8/LGR5*, *HNF1B* in 6103 T2D patients and 2061 healthy controls with age greater than 16 years. Combined genetic score was calculated by summing the score of risk allele for each SNP under additive genetic model. Associations between T2D and genetic variants were tested by logistic regression adjusted for age, sex and body mass index.

Results - We observed consistent and significant associations of *NOTCH2*, *IGF2BP2*, *WFS1*, *CDKAL1*, *JAZF1*, *SLC30A8*, *CDKN2A/B*, *HHEX*, *TCF7L2*, *KCNQ1* and *KCNJ11* with T2D risk, yielded odds ratios ranged from 1.10 to 2.15 ($1.8x10^{-15} < P < 0.042$). Moreover, the 14 SNPs exhibited joint effect on increasing T2D risk (per-allele OR (95% CI) = 1.15 (1.11 - 1.18); $P = 7.5x10^{-21}$). Subjects carrying 16 or more risk alleles had an OR of 3.45 (95% CI, 2.41 - 4.94) compared to those carrying 11 or less risk alleles.

Conclusion - In summary, our study support common variants in *NOTCH2*, *IGF2BP2*, *WFS1*, *CDKAL1*, *JAZF1*, *SLC30A8*, *CDKN2A/B*, *HHEX*, *TCF7L2*, *KCNQ1* and *KCNJ11* loci independently or jointly increase risk for T2D in Chinese population.

Acknowledgement - This work was supported by the Hong Kong Governments Research Grant Committee Central Allocation Scheme (CUHK 1/04C), Research Grants Council Earmarked Research Grant (CUHK4727/0M), the Liao Wun Yuk Diabetes Fund and the Innovation and Technology Fund (ITS/487/09FP).

2281T

AGRP SNPs and gender effects on obesity in African Americans. *E. Shah[1], A. Patterson[1], U. Al-Alem[2], R.A. Kittles[1,2,3].* 1) Institute of Human Genetics, University of Illinois at Chicago, Chicago, IL; 2) Division of Epidemiology and Biostatistics, School of Public Health, University of Illinois at Chicago, Chicago, IL; 3) Section of Hematology/Oncology, Department of Medicine, University of Illinois at Chicago, Chicago, IL.

Background: Obesity is a complex metabolic disorder that results from a sustained imbalance between food intake and energy expenditure. African Americans (AAs) have the highest prevalence of obesity compared to Hispanics and non-Hispanic Whites. Multiple genetic studies have focused on body mass index (BMI) in diverse populations. Here we aim to identify genetic effects on BMI in candidate genes and/or previous GWAS findings near the *AGRP, FTO, NEGR1, BDNF, JAZF1 and LGR4* gene regions in individuals of west African descent.**Method**: We examined eighteen (18) SNPs previously associated with obesity in whites and West African populations in a total of 1,393 AAs in a two-stage analysis [691 males in our first stage and 702 AAs (278 males and 424 females) in our second stage]. BMI was calculated from weight and height. Regression analyses were used to test for association of SNPs and BMI, controlling for age, gender, and west African genetic ancestry. **Results**: We report that SNP rs11575892, in the Agouti Related Protein gene (*AGRP*), was associated with BMI in our stage one, stage two and pooled analyses (P<0.004). In addition, two other *AGRP* SNPs, rs503891 and rs34123523, were associated with BMI only among females (P=0.02). **Conclusion**: Polymorphisms in the *AGRP* gene have been shown to be involved in the development of body fatness in different populations. Our findings suggest that several *AGRP* SNPs are involved in BMI in AAs. In addition, we detected gender differences in *AGRP* SNP effects on BMI. The gender differences observed in our study may be related to the differences in composition and other measures of adiposity, including serum leptin levels, sex steroids, and body fat patterning, which have been shown to vary between the sexes.

2282F

Pleiotropic effects of obesity-susceptibility loci on metabolic traits: a meta-analysis of up to 37,874 individuals. J. van Vliet-Ostaptchouk[1,2,3], M. den Hoed[4], J. Luan[4], J-H. Zhao[4], K.K. Ong[4], P.J. van der Most[2], A. Wong[5], R. Hardy[5], D. Kuh[5], M.M. van der Klauw[1,6], M. Bruinenberg[6], B.H.R. Wolffenbuttel[1,6], N.J. Wareham[4], H. Snieder[2,6], R.J.F. Loos[4,7]. 1) Department of Endocrinology, University Medical Center Groningen, University of Groningen, Groningen, Netherlands; 2) Unit of Genetic Epidemiology and Bioinformatics, Department of Epidemiology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands; 3) Netherlands Consortium for Healthy Aging, the Netherlands; 4) Medical Research Council Epidemiology Unit, Institute of Metabolic Science, Cambridge, CB2 0QQ UK; 5) MRC Unit for Lifelong Health and Ageing, University College London, WC1B 5JU, UK; 6) LifeLines Cohort Study, University Medical Center Groningen, University of Groningen, The Netherlands; 7) Genetics of Obesity and Related Metabolic Traits Program, Mount Sinai School of Medicine, New York, NY 10029, USA.

Obesity is a key risk factor for a number of metabolic diseases, including type 2 diabetes, dyslipedimia and cardiovascular disease. Although the exact biological mechanisms linking obesity to these co-morbidities are not fully understood, a set of common genetic factors with pleiotropic effects (i.e. affecting multiple traits) might in part explain the observed associations. We aimed to investigate whether the currently established obesity-susceptibility loci identified through genome-wide associations studies for body mass index (BMI) and waist-to-hip-ratio (WHR) influence metabolic traits independently from obesity-related phenotypes. We systematically assessed associations of the 32 BMI and 14 WHR loci, individually and combined in two genetic predisposition scores (GPS-BMI, GPS-WHR), with glycaemic traits, blood lipids, and blood pressure (BP) and examined whether these associations were influenced by the level of obesity by performing meta-analyses of data of up to 37,874 individuals of European ancestry from six population-based studies. The meta-analyses revealed associations of both the individual obesity-susceptibility loci and the genetic predisposition scores with metabolic traits that were not driven by the obesity-related phenotypes. We observed significant associations of the BMI-increasing alleles at five BMI-loci with higher concentrations of LDL and total cholesterol levels (QPTCL), lower 2-hr glucose levels (RBJ, QPTCL), lower levels of HDL- (SLC39A8, MTCH2), LDL- (FLJ35779), and total cholesterol (FLJ35779), and reduced diastolic BP (SLC39A8) (all Ps<2.5x10-4), independent of their effect on BMI. WHR-increasing alleles at two WHR-loci were associated with higher proinsulin levels (GRB14) and lower levels of fasting glucose (CPEB4) (Ps<2.3x10-4), independent of both BMI and WHR. A higher GPS-BMI was associated with increased levels of insulin, pro-insulin and triglycerides and with lower HDL-cholesterol levels (Ps<2.3x10-3), but these associations were abolished after adjusting for BMI. The associations of the GPS-BMI with lower 2-hr glucose and systolic and diastolic BP (P<1.9x10-3) became apparent after adjustment for BMI. The GPS-WHR was associated with lower HDL-cholesterol and higher triglyceride levels independently of WHR and BMI (Ps<2.9x10-3). Conclusions: These results provide evidence that obesity-susceptibility loci are associated with metabolic traits independent of their effect on obesity, suggesting pleiotropic effects.

2283W

Investigation of celiac disease susceptibility variants in the Irish population using the immunochip. C. Coleman[1], E. Quinn[1], A.W. Ryan[1], J. Conroy[2], S. Ennis[2], V. Trimble[1], R. McManus[1]. 1) Institute of Molecular Medicine, Trinity College Dublin, Dublin, Ireland; 2) Health Sciences Centre, University College Dublin, Ireland.

Celiac disease (CD) is a common, complex and chronic immune-mediated disease triggered in genetically sensitive individuals by the ingestion of gluten. Susceptibility is prevalent in European populations at a frequency of approximately 1%. HLA-DQA1 and HLA-DQB1 alleles are necessary but not sufficient to cause disease and recent genome wide association studies (GWAS) have identified many associated non-HLA risk loci. A large proportion of the causative susceptibility variants likely remain to be identified due to insufficient capture of genetic variation using existing methodologies. The Immunochip is a custom Illumina Infinium high-density array which addresses these issues, consisting of 196,524 common and rare polymorphisms designed to perform deep replication and fine mapping of established GWAS significant loci for 12 major autoimmune and inflammatory diseases, including celiac disease. Trynka *et al* (Nat Genet, 2011) reported strong association with several existing and new susceptibility loci in a study of celiac disease using the immunochip. In order to replicate these loci in an Irish case control CD study, we genotyped 926 individuals (461 cases and 465 controls) using the immunochip. After QC filters, 101,498 SNPs remained with which to conduct analysis. Association analysis was performed using PLINK and the preliminary results show nominal association in our sample of several of the strongest findings from the Trynka *et al* study, as well as several previously unreported associations. We are currently genotyping the most associated variants in an extended study population. Other non-HLA variants previously reported to be strongly associated with CD show no evidence of association in our study.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2284T

Risk alleles associated with ANA production in healthy population or transition to autoimmunity. E. Wakeland[1], P. Raj[1], E. Rai[1], QZ. Li[1], D. Karp[1], P. Doshi[1], N. Olsen[2], KM. Sivils[3], J. James[3], J. Kelly[3], B. Lauwerys[4]. 1) Dept Immunology, Univ Texas SW Med Ctr, Dallas, TX.5323 Harry Hines Blvd, Dallas, Texas 75390-9093, USA; 2) Division of Rheumatology, Penn State Hershey Medical Center, 500 University Drive, Hershey PA 17033; 3) Oklahoma Medical Research Foundation, 825 North East 13th Street, Oklahoma City, Oklahoma 73104, USA; 4) Service de Rhumatologie, SSS/IREC/RUMA, Cliniques Universitaires Saint-Luc & Université catholique de Louvain Avenue Hippocrate 10, bte B2.5390, 1200 BRUXELLES, Belgium.

Susceptibility to SLE is a complex, genetic disease, impacted by more than 30 susceptibility loci and a variety of unknown environmental factors. The production of antinuclear antibody (ANA) is a universal feature of SLE, but is not unique to SLE patients since more than 5% of healthy individuals are also ANA positive. We hypothesize that ANA positive normal individuals have a subset of autoimmune susceptibility alleles and that they have increased risk for development of autoimmunity. We genotyped 1726 ANA negative controls, 318 ANA positive controls and 525 SLE patients using the Immunochip, which contains 196,524 SNPs that provide a dense coverage of susceptibility loci associated with a variety of autoimmune diseases. Genotypes were called on Genome Studio and SNP analysis was done with Golden Helix. Our analysis of this data replicated several GWAS signals in SLE genes and also discovered new genes reaching genome-wide ($p \leq 10^{-8}$) or suggestive association ($p \leq 10^{-5}$). Association test statistics indicated that SLE susceptibility genes could be grouped on the basis of their role in driving the production of ANAs versus driving the transition from ANA positive to SLE. We found that genes of adaptive immune pathway, for example LAMB4, LAMB1, LY9, SLAMF7, LYN, IKZF1, ETS, LRRK2, CCR3, SLC30A7, TNFSF4, BANK1, RTEL1, TNFAIP3, were associated with the transition to ANA positive phenotypes in normal individuals, while variations in IRF5, ITGAM, STAT4, TNFAIP3, XKR6, KIR2DL3, FUT2, HERC2, were associated with the transition, from ANA positive to full SLE phenotype. Additional genes and associations will be discussed during presentation.

2285F

Interleukin-18 Genetic Polymorphism and Kawasaki disease in Han Chinese children. H. Chi[1,4], MR. Chen[1,4], NC. Chiu[1,4], FY. Huang[1], HW. Chan[2], SC. Chang[2], CL. Lin[2], YJ. Lee[1,2,3,5]. 1) Department of Pediatrics, Mackay Memorial Hosp, Taipei, Taiwan; 2) Department of Medical Research, Mackay Memorial Hospital, Taipei; 3) Department of Medicine, Institute of Biomedical Sciences, Mackay Medical College, New Taipei City; 4) Mackay Medicine, Nursing and Management College, New Taipei City; 5) Department of Pediatrics, Taipei Medical University, Taipei, Taiwan.

Background: Interleukin 18 (IL18) stimulates interferon-γ production in Th1 cells. Therefore we designed a case-control study to investigate the association between the IL18 gene and KD. Methods: The patients were 190 unrelated children (112 boys) with KD. Their age at the diagnosis was 1.9 ± 1.7 years. The controls were 1211 subjects. dbSNPs rs187238 (-137G>C) and rs1946518 (-607T>G) in the promoter region of the IL18 gene were genotyped. Statistical analysis: We assessed the Hardy-Weinberg equilibrium, estimated the frequencies of haplotypes, and tested pairwise linkage disequilibrium (LD) between the SNPs in both KD and control groups. Statistical differences in genotype, allele, carrier, and haplotype distributions between patients and controls were compared. Odds ratios and 95% confidence intervals were calculated. A Pc value < 0.05 was considered statistically significant. Results: The genotypes of both SNPs were in Hardy-Weinberg equilibrium in controls (rs187238, P = 0.37; rs1946518, P = 0.83). The frequencies of each allele of the SNPs or haplotypes was not significantly different between male and female controls (data not shown). There was no significant difference in the distribution of genotype, allele or carrier of each SNP between the patients and controls. SNP rs187238C was 55 (14.5%) in patients vs. 271 (12.1%) in controls (P = 0.19). rs1946518G was 187 (49.2%) in patients vs. 1095 (48.8%) in controls (P = 0.89). The 2 SNPs were closely linked (D', 0.961; r2, 0.121) and formed 4 haplotypes. Among them, the GG haplotype (rs182738-rs1946518) was the most frequent (48.7%). The distribution of these haplotypes in patients was not significantly different from that in controls (P > 0.05). Conclusion: The IL18 gene was not associated with KD in Han Chinese children.

2286W

Association between IL1A, IL1B and TNFA polymorphisms and glaucoma in a Brazilian population. M.B. Oliveira[1], J.P.C Vasoncellos[2], V.P. Costa[2], M.B. Melo[1]. 1) Center of Molecular Biology and Genetic Engineering, University of Campinas; 2) Department of Ophthalmology, Faculty of Medical Sciences, University of Campinas.

Glaucoma is a progressive atrophy of the optic disc characterized by loss of retinal ganglion cells, which leads to a corresponding visual field defect. It's the major cause of irreversible blindness worldwide. Among several forms of glaucoma, the most prevalent is primary open angle glaucoma (POAG). POAG is a complex disease with several risk factors including intraocular pressure (IOP) increase and positive family history. The treatment of glaucoma is based on IOP reduction, which can be achieved by topical medication and/or surgical procedure in order to avoid disease progression. Linkage analysis, association studies and candidate genes approaches have been performed in order to identify disease-causing genes and variants of susceptibility associated with POAG. Among them there are genes that codify for cytokines (IL1A, IL1B and TNFA), which have been associated in previous studies with higher risk for the development of GPAA in some populations. Besides, these molecules can interfere with the wound healing process that is vital to the success of surgical procedures performed in glaucoma treatment. This case control study involved approximately 100 unrelated POAG patients and 100 healthy controls from Brazil for the determination of polymorphisms in the IL1A, IL1B and TNFA genes. Comprehensive ophthalmic evaluation was performed and genomic DNA was obtained from case and control groups. Five single nucleotide polymorphisms (SNPs): IL1A (-889C/T; rs1800587), IL1B (-511C/T; rs16944), IL1B (3953C/T; rs1143634), TNFA (-238G/A; rs361525) and TNFA (-308G/A; rs1800629) were determined through direct sequencing. The association related above was evaluated using chi-square test. No significant difference was observed in the allele and genotype frequencies for IL1A, IL1B and TNFA SNPs between POAG patients and controls. Among patients with POAG, an analysis comparing genotypes and clinic data as intraocular pressure, vertical cup to disc ratio and number of surgical procedures necessary to IOP control were performed and no statistical differences among the groups were observed. In a Brazilian population sample no association was observed between disease risk alleles in the IL1A, IL1B and TNFA genes and POAG as well as with the severity of glaucoma. A similar study with additional and larger cohorts of patients using also other population groups is necessary to further substantiate the observation.

2287T

Evaluation of new classification criteria for HLA-DRB1 alleles based on three amino acid positions (11, 71 and 74) in Japanese RA patients; disease susceptibility and severity. S. Yoshida[1,2], K. Ikari[1], K. Yano[1], Y. Toyama[2], A. Taniguchi[1], H. Yamanaka[1], S. Momohara[1]. 1) Institute of Rheumatology, Tokyo Women's Medical University, Shinjuku, Tokyo, Japan; 2) Department of Orthopaedic Surgery, Keio University School of Medicine, Shinjuku, Tokyo, Japan.

Background. Rheumatoid arthritis (RA) is a complex polygenic disease of unknown etiology that is characterized by the appearance of progressive joint damage. Though a lot of susceptibility genes to RA have been identified, the human leukocyte antigen (HLA) locus are unequivocally the strongest genetic risk for RA. To date, several studies have suggested that shared epitope (SE, particular sequence of amino acids spanning positions 70-74) are associated with susceptibility to and severity of RA. Recently, however, Raychaudhuri et al. reported that three amino acid positions (11, 71 and 74) can almost completely explain the association between HLA and susceptibility to RA (1). In the present study, we investigated the relevance of this new HLA classification in terms of RA susceptibility and severity in Japanese patients.

Methods. DNA samples from 2068 Japanese RA patients and 752 controls were used for the study. Among them, Sharp/ van der Heijde score (SHS) of hands at 5-year disease duration, which represents joint damage, were measured in 865 patients. Typing of HLA-DRB1 was performed using the AlleleSEQR HLA-Sequencing-Based Typing Kit. Association between RA susceptibility and HLA-DRB1 alleles was examined by the relative predispositional effect method and differences in SHS among copies of the susceptible HLA-DRB1 alleles were analyzed by linear regression analysis.

Results. The result of effect estimates for the three amino acids associated with susceptibility to RA was similar to the results from Raychaudhuri et al. The major difference between SE and susceptible haplotypes in new classification was *0901, which is non-SE allele but one of the strong susceptible alleles in Asian. Though the susceptible haplotypes in new classification were associated with SHS, the single amino acid (Val11 or Leu11) was more explainable for radiographic severity.Conclusion. We confirmed the new classification criteria of HLA-DRB1 in Japanese RA patients. Further studies are needed on the effect of the three amino acid (positions 11, 71 and 74) on radiographic damage.

Reference. 1. Raychaudhuri S. et al. Five amino acids in three HLA proteins explain most of the association between MHC and seropositive rheumatoid arthritis. Nat Genet. 44:291-6. 2012.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2288F

Genetic susceptibility to leptospirosis in the Azorean island of São Miguel population. L.M. Esteves[1], S.M. Bulhões[1], M.J. Brilhante[1], C.C. Branco[1,5], F.M. Mota[2], C. Paiva[3], M.L. Vieira[4], L. Mota-Vieira[1,5]. 1) Genetics & Molecular Pathology Unit, Hospital of Divino Espirito Santo of Ponta Delgada (Azores), Portugal; 2) Infectious Diseases Department, Hospital of Divino Espirito Santo of Ponta Delgada (Azores), Portugal; 3) Internal Medicine Department, Hospital of Divino Espirito Santo of Ponta Delgada (Azores), Portugal; 4) Leptospirosis and Lyme Borreliosis Group, Unit of Medical Microbiology, Instituto de Higiene e Medicina Tropical, Centro de Recursos Microbiológicos, Universidade Nova de Lisboa, Portugal; 5) Azores Genetics Research Group, Instituto Gulbenkian de Ciência, Lisbon, Portugal.

In the majority of infectious diseases only a proportion of individuals exposed to the pathogen become infected and develop clinically evident disease. This inter-individual variability is in part determined by the host immune responses to pathogen infection. Here, we investigate how host genetic diversity is involved on the susceptibility/resistance to leptospirosis, a disease well spread in the Azores (Portugal), an archipelago located on the North Atlantic Ocean. The strategy consisted of a population-based case-control study. The cases included 105 individuals diagnosed with leptospirosis by clinical history and laboratory analysis, and the control population consisted of 470 DNA samples of unrelated healthy blood donors from the Azorean São Miguel Island. Twelve genes - *IL10*, *IL1α*, *IL1β*, *IL6*, *IL12RB1*, *CISH*, *TLR2*, *TLR4*, *TLR9*, *CD14*, *LTA* and *TNF* - involved in the innate immune response to pathogens were analyzed by two methods: single (real-time PCR TaqMan® allelic discrimination) and multiplex SNP (SNaPshot minisequencing) genotyping. In a total of 25 SNPs, statistical analysis revealed 3 single risk alleles associated with the disease susceptibility: rs622502 (*CISH*, RR=1.553, 95% CI 1.056-2.250), rs5743836 (*TLR9*, RR=1.182, 95% CI 1.150-1.182) and rs4986790 (*TLR4*, RR=1.068, 95% CI 1.040-1.068). We also observed two SNPs - rs16944 (*IL1β*, RR=0.748, 95% CI 0.578-0.952) and rs2569191 (*CD14*, RR=0.075, 95% CI 0.037-0.143) - that confer resistance to leptospira infection. In order to access if heterozigosity or homozigosity influence this susceptibility, we performed a genotype risk analysis. Results demonstrate that, for rs622502 (*CISH*), carriers of C allele have an increased risk (RR=1.592, 95% CI 1.092-2.254) for the infection. On the other hand, for rs2569191 (*CD14*), carriers of T allele show a resistance effect (RR=0.202, 95% CI 0.099-0.383). Currently, we are studying the association of single and combined SNPs with disease severity and clinical outcome. Overall, these data suggest that, in São Miguel population, at least, *CISH*, *TLR9* and *TLR4* genes play an important role in host susceptibility to leptospirosis. (lmotavieira@hdes.pt, Funding by Fundação Calouste Gulbenkian, ref P-99888: SDH49, and by DRCTC from Government of Azores).

2289W

A replication study supports *NOD2* and *CCDC122* as leprosy susceptibility genes. H. Salomão[1], W.L. Silva[2], P. Medeiros[2], V.M. Fava[1], I.M.F. Dias-Baptista[2], M.L.C. Virmond[2], M.O. Moraes[3], M.T. Mira[1,4], A.C. Pereira-Latini[2]. 1) School of Medicine, Pontifícia Universidade Católica do Paraná - PUCPR, Curitiba, Paraná, Brazil; 2) Instituto Lauro de Souza Lima, Bauru, São Paulo, Brazil; 3) Laboratório de Hanseníase, Instituto Oswaldo Cruz, FIOCRUZ, Rio de Janeiro, Brazil; 4) School of Health and Biosciences, Pontifícia Universidade Católica do Paraná, Curitiba, Brazil.

Leprosy is a chronic infectious disease caused by *Mycobacterium leprae* that affects 230.000 new individuals worldwide every year. A previous genome-wide study conducted in a Chinese population sample revealed association between leprosy and five non-HLA genes: *CCDC122, C13orf31, NOD2, TNFSF15* and *RIPK2*. Here, we present the results of an independent, two-step replication study including all five candidate genes. First, we genotyped 36 tag *SNPs* capturing the entire information of all candidate genes in 179 individuals distributed in 51 trios composed by an affected child and both parents. All trios were recruited at the Prata Colony, an isolated, highly endemic former leprosy located at the outskirts of Brazilian Amazon. Positive signals of association were replicated in an independent case-control sample of 402 patients and 413 controls recruited at the city of Rondonópolis, Mato Grosso state, central Brazil. Fluorescence-based TaqMan technology was applied to produce genotypes of all individuals. Family-based association analysis was performed as implemented in the software FBAT. Linkage disequilibrium estimations were carried out using the Haploview software. The statistical software R for Windows, version 2.14.0, was used for case-control analysis. All markers were in Hardy Weinberg equilibrium. The family-based study revealed association between leprosy and susceptibility alleles G of rs8057341 (p=0.01) and T of rs4942254 (p=0.04), located at *NOD2*and *CCDC122*, respectively. Both signals were replicated in the case-control sample (AA genotype of rs8057341: p=0.02; OR=0.56, 95% CI: 0.34-0.94; CC genotype of rs4942254: p=0.03; OR=0.65, 95% CI: 0.44-0.97). The association between leprosy and *NOD2* marker rs8057341 is a perfect replication of the results observed in the original Chinese study. Our results support the hypothesis of a role for variants of *NOD2* and *CCDC122* in the complex molecular mechanism controlling leprosy susceptibility. * HS and WLS authors contributed equally for the study. Funding: DECIT-MS/CNPq. Grant 576051/2008-0.

2290T

Positive association of Apolipoprotein E4 polymorphism with Recurrent pregnancy Loss in Iranian patients. M.T. Akbari[1,2], N. Asgari[2]. 1) Department of Medical Genetics, Tarbiat Modares University, Tehran, Iran; 2) Tehran Medical Genetics Laboratory, Taleghani street, No. 251,Tehran, Iran.

Numerous lines of evidence implicate Apolipoprotein E (APOE) in lipid metabolism during pregnancy. Hence, a role for its polymorphism has been envisaged in recurrent pregnancy loss (PRL) considering major structural and functional differences between different genotypes of the three major alleles of E2, E3 and E4. To investigate the role of different APOE genotypes in Iranian patients, we performed a case-control study of 81 women with two or more pregnancy losses who did not have any other known risk factors including anatomic anomalies of the reproductive system, infections, immunologic factors, hormonal imbalances, chromosomal abnormalities and environmental factors; and the control group were 81 women with at least two healthy children and no PRL in their reproductive history. DNA was extracted from the peripheral blood following written consent and APOE genotyping was carried out by amplifying exon 4 of the gene and subjecting it to digestion by HhaI restriction enzyme. Genotyping was concluded by analyzing different fragment sizes produced, which resulted in finding significantly higher frequency of combined E3/E4 & E4/E4 genotypes in the patients (almost 20%) compared to the normal controls (2.4%). The genotypes were confirmed by DNA sequencing. Allelic frequency for E4 was 13.5% in the patients and only 1% in the control group. Our findings is in line with a number of similar studies carried on other populations. Therefore, Apo E4 polymorphism seems to be contributing to the thrombophilic risk factors as a background to RPL.

2291F

Investigation of rs4236601 and rs4977756 SNPs in a Primary Open-Angle Glaucoma Brazilian population. H.F. Nunes[1], M.B. Oliveira[1], .JP.C. Vasconcelos[2], V.P. Costa[2], M.B. de Melo[1]. 1) CBMEG, UNICAMP, campi1-Center of Molecular Biology and Genetic Engineering, University of Campinasnas, Brazil; 2) 2Department of Ophthalmology, Faculty of Medical Sciences, University of Campinas.

Primary open-angle glaucoma (POAG) is a chronic neurodegenerative disease that leads to progressive damage to retinal ganglion cells resulting in visual field loss. Glaucoma is recognized as the main cause of irreversible blindness worldwide. Although the pathophysiology of glaucoma is not well understood, positive family history is one of the most important risk factor for glaucoma development. Two single nucleotide polymorphisms (SNPs) have been widely associated with glaucoma, although their frequencies vary among different populations. The risk allele rs4236601 is located in a LD block at chromosome 7q31 that contains two genes, CAV1 and CAV2 and rs4977756 is located in the CDKN2B-AS1 gene. The goal of this study was to investigate the distribution of these single nucleotide polymorphisms (SNPs) as a glaucoma risk factor in a sample of Brazilian POAG patients. A case control study involving 100 POAG patients and 100 control subjects was genotyped by direct sequencing. The frequency of each of the three possible genotypes of rs4236601 and rs4977756 were compared between POAG patients and controls and a significant difference for the rs42366601 was detected (p=0.00012). The minor allele frequency of rs4236601 in our study group from Brazil was 51.3% in the POAG patients and 31,5% in control subjects. Significant difference was detected in allele frequency between the POAG patients and controls (p=0.002). For the rs4977756 in the CDKN2B-AS1 gene, no significant difference was observed (p=0.9). The minor allele frequency was next to 50% in both groups. The rs4236601 is located between CAV1 and CAV2 genes. However, the mechanism through which alterations in these genes might cause glaucoma is unknown. It has been hypothesized that they might alter nitric oxide or transforming growth factor-β(TGF-β) signal transduction although no glaucoma causing mutations have been identified in either CAV1 or CAV2 genes. Overall, these data and observations imply that a POAG risk allele at the 7q31 locus is possibly strongly associated with glaucoma in a sample of the Brazilian population.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2292W

Role of genetic variation at the *FTO* gene in genetic susceptibility to osteoarthritis. K. Panoutsopoulou[1], S. Metrustry[2], S. Doherty[3], L.L. Laslett[4], R.A. Maciewicz[5], D.J. Hart[2], W. Zhang[3], K.R. Muir[6], M. Wheeler[3], arcOGEN. Consortium[7], C. Cooper[8,9], T. Spector[2], F.M Cicuttini[10], G. Jones[5], N.K. Arden[8,9], M. Doherty[3], E. Zeggini[1], A. Valdes[2]. 1) Wellcome Trust Ctr, Wellcome Trust Sanger Institute, Cambridge, CB10 1HH, UK; 2) Dept of Twin Research, King's College London, St Thomas' Hospital, Westminster Bridge Rd, London SE1 7EH, UK; 3) Academic Rheumatology, Clinical Sciences Building, Nottingham City Hospital Hucknall Road, Nottingham, NG5 1PB, UK; 4) Menzies Research Institute, University of Tasmania, Hobart, Australia; 5) Respiratory & Inflammation Research Area, AstraZeneca, Charnwood R&D, Bakewell Road, Loughborough, Leicestershire, LE11 5RH, UK; 6) Health Sciences Research Institute, Warwick Medical School University of Warwick, Gibbett Hill Road, Coventry, CV4 7AL, UK; 7) www.arcogen.org.uk; 8) NIHR Musculoskeletal Biomedical Research Unit, University of Oxford, Oxford, UK; 9) MRC Lifecourse Epidemiology Unit, University of Southampton, Southampton General Hospital, Southampton, SO16 6YD; 10) Department of Epidemiology and Preventive Medicine, School of Public Health and Preventive Medicine, Monash University, Alfred Hospital, Melbourne, VIC 3004, Australia.

Variation in the *FTO* gene is known to play an important role in susceptibility to obesity which is a major risk factor for the development of osteoarthritis (OA). A recent study by the arcOGEN Consortium established an association of rs8044769 at the BMI-associated gene *FTO* and OA reaching almost genome-wide significance. The discovery and replication studies did not match cases and controls for BMI as the discovery dataset employed population-based controls lacking BMI information. A sensitivity analysis in a subset of samples from the discovery set showed attenuation of the signal after BMI adjustment compared to the unadjusted analysis. In this study we have thoroughly investigated the role of the *FTO* polymorphism rs8044769 in risk of knee and hip OA in cases and controls from 7 independent cohorts from the UK and Australia matched for BMI category. We genotyped rs8044769 in 892 normal weight, 2036 overweight (25<=BMI<30) and 2300 obese knee OA patients, 1162 normal weight, 1722 overweight and 1374 obese hip OA patients and 2372 normal weight, 1931 overweight and 890 obese healthy controls. Association analyses were carried out adjusting for gender and the results were meta-analysed. In these samples the *FTO* SNP is strongly associated with overweight status, and overweight status is strongly associated with both knee and hip OA. When samples were stratified by BMI category meta-analysis of the association between the *FTO* SNP and OA resulted in an odds ratio (OR) =0.97 (95%CI 0.82-1.15) for the association between hip OA in normal weight individuals, OR=1.04 (95%CI 0.95-1.14) in overweight (including obese) and OR=1.02 (95%CI 0.76-1.37) for obese individuals showing significant heterogeneity (p<0.01) among cohorts for the latter comparison. For knee OA the results were OR=1.03 (0.87-1.21) in normal weight individuals, OR=1.01 (0.94-1.09) in overweight people and OR=1.05 (0.87-1.26) in obese individuals showing near significant (p<0.06) heterogeneity. No significant associations (p<0.05) were observed at any of the BMI-stratified meta-analyses. Our results indicate that the association between *FTO* and OA is due to its effect on obesity/overweight and that it has no direct effect on OA risk when adjusted for BMI category. This is consistent with what is known of the biology of the *FTO* gene.

2293T

Association of COX-2, IL-1β and IL-1 RN gene polymorphisms with susceptibility to chronic periodontitis in North India population. G. Prakash[1], S. Ajay[2], K.K Gupta[3], J. Dixit[4], D. Bali[5], B. Mittal[1]. 1) Department of Genetics, SGPGIMS, Lucknow, India; 2) Department of Biochemistry, SPPIDMS, Lucknow, India; 3) Department of Periodontics, SPPIDMS, Lucknow, India; 4) Department of Periodontics, Faculty of Dental Sciences, CSMMU, Lucknow, India; 5) Department of Periodontics and Implantology, DAV Dental College & Hospital, Yamuna Nagar, Haryana, India.

Introduction: Periodontitis is a chronic inflammatory disease resulting in the destruction of the tissue and alveolar bone supporting the teeth and leading ultimately to tooth loss. Studies have shown the involvement of inflammatory genes like IL-1, IL-4, IL-6, IL-10, TNF-α, TLR-2, TLR-3 and COX-2 in etiology of chronic periodontitis. Genetic variants in these inflammatory genes are known to affect the respective protein functions, thus may modulate susceptibility of chronic periodontitis. Therefore, in present study we investigated the role of COX-2 (-765 G>C rs20417, +8473 T>C rs5275, -1195 G>A rs689466), IL-1β -511 C>T (rs2853550) and IL-1 RN 86bp VNTR (rs2234663) with risk of chronic periodontitis in North India population. Materials and Methods: The genotyping of COX-2 (-765 G>C, +8473 T>C, -1195 G>A), IL-1β -511 C>T and IL-1 RN 86bp VNTR polymorphisms were carried out in 154 healthy controls and 152 chronic periodontitis patients. Cases were clinically diagnosed by detection of chronic inflammatory changes in the marginal gingiva, presence of periodontal pocket and loss of clinical attachment. All subjects were genotyped by polymerase chain reaction-restriction fragment length polymorphism method (PCR-RFLP). Difference in the allele and genotype frequencies were assessed by Chi-square test (P<0.05). The risk of periodontitis associated with allele and genotype was calculated as Odd ratio (OR) with 95% Confidence Intervals (CI) using SPSS ver.15.0. Results: We did not find significant association of COX-2 (-765 G>C, +8473 T>C, -1195 G>A), IL-1β -511 C>T and IL-1 RN 86bp VNTR with risk of chronic periodontitis at genotypic and allelic level (P value at allelic level for COX-2 -765, 0.254, COX-2 +8473, 0.647, COX-2 -1195 0.374, IL-1β, 0.897and IL-1 RN, 0.561). Even, when subjects were stratified on the basis of gender, no gender-specific risk was observed (P>0.05). In case only analysis, no significant interaction of COX-2 -765 G>C (P=0.502), +8473 T>C (P=0.919), -1195 G>A (P=0.110), IL-1β -511C>T (P=0.193) and IL-1 RN (P=0.771) were observed in patients with tobacco usage habits. Conclusion: The present study shows that COX-2 (-765 G>C, +8473 T>C and -1195 G>A), IL-1β -511 C>T and IL-1 RN 86bp VNTR polymorphism do not play any significant role in genetic susceptibility to chronic periodontitis in North India population. Financial support from ICMR, New Delhi, India.

2294F

A variant of filaggrin (FLG) are associated with susceptibility to food allergy in the Japanese population. T. Hirota[1], M. Ebisawa[2], T. Imai[3], Y. Suzuk[4], N. Shimojo[5], Y. Kohno[5], M. Kubo[1], M. Tamari[1]. 1) CGM, RIKEN Yokohama, Yokohama, Kanagawa, Japan; 2) Clinical Research Center for Allergy and Rheumatology, Sagamihara National Hospital, National Hospital Organization, Kanagawa, Japan; 3) Department of Pediatrics, Showa University School of Medicine, Tokyo, Japan; 4) Departments of Public Health, Chiba University Graduate School of Medicine, Chiba, Japan; 5) Departments of Pediatrics, Graduate School of Medicine, Chiba University, Chiba, Japan.

[Background]: Food allergy (FA) is an adverse health effect arising from a specific immune response that occurs reproducibly on exposure to a given food. FLG encodes an important epidermal protein abundantly expressed in the outer layers of the epidermis. Loss-of-function variants within the FLG gene are associated with atopic dermatitis and other atopic diseases including FA. But genetic influences of FLG on FA in the Japanese population are unclear. **[Objective]**: To investigate the association between filaggrin loss-of-function variants and food allergy and food allergy related phenotypes in the Japanese population. **[Methods]**: Eight variants (R501X, 3321delA, S1695X, Q1701X, S2554X, S2889X, S3296X, K4022X) in FLG gene, which have been already reported in the Japanese populations, were genotyped by Taqman and invader methods. We performed an association study of FA using a total of 532 subjects with FA and 234 controls in the Japanese population. Among the FA case, a total of 60 subjects have food-induced anaphylaxis. We conducted an association analyses by Fisher's exact test.**[Results]**: We found that S2889X variant was associated with FA under an allelic model in the Japanese population (P = .0042; OR, 3.94; 95% CI, 1.39-11.15). In further analyses of patient subgroups, we observed a strong association between S2889X variant and FA patients who had a history of food-induced anaphylaxis (P = .000035; OR, 10.08; 95% CI, 3.15-32.19).**[Conclusion]**: S2889X variant in the FLG gene might be involved in FA in a loss of function manner. It has also been suggested that the variant plays a role in food-induced anaphylaxis. Further studies are needed for better understanding of pathophysiology of food allergy.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2295W

AKT1 fails to replicate as a longevity-associated gene in Danish and German nonagenarians and centenarians. *M. Nygaard[1,2], M. Soerensen[1,2], F. Flachsbart[4], J. Mengel-From[1,2], Q. Tan[1,2], S. Schreiber[4], A. Nebel[4], K. Christensen[1,2,3], L. Christiansen[1,2].* 1) Epidemiology, Institute of Public Health, University of Southern Denmark, Odense, Denmark; 2) Department of Clinical Genetics, Odense University Hospital, Odense, Denmark; 3) Department of Clinical Biochemistry and Pharmacology, Odense University Hospital, Odense, Denmark; 4) Institute of Clinical Molecular Biology, Christian-Albrechts-University, Kiel, Germany.

Despite the fact that approximately 25% of variation in human lifespan is attributable to genetic factors, polymorphisms in only a few candidate genes have so far been consistently found to be associated with longevity. Recently, *AKT1* has been suggested as a third universal longevity gene, following the well-established apolipoprotein E (*APOE*) and forkhead box O3 (*FOXO3*) genes. To further elucidate this potential role of *AKT1*, case-control studies involving a large and unique sample of 2996 long-lived individuals (nonagenarians and centenarians) and 1840 younger controls of Danish and German ancestry were performed. In addition to the previously described candidate variation, rs3803304, this study was expanded to include a number of additional SNPs to ensure better coverage of the known allelic variance in *AKT1* and to facilitate a more exhaustive exploration of the possible relevance of *AKT1* variations in longevity.

The main finding of this study was a lack of formal replication of the association between rs3803304 and longevity as previously found in three independent populations. However, restricting the analysis to centenarians revealed a trend of supportive evidence in Danish men (p=0.03), with the same direction of effect as identified before (OR=0.49).

Furthermore, six additional SNPs were tested in both the Danish and German study population. Of note, one of these, rs2494731, is in perfect linkage disequilibrium with the *AKT1* SNP rs2498804, which was recently found to be associated with lifespan in a population of Dutch nonagenarians. Nevertheless, we did not replicate an association with longevity, whether applying a cross-sectional approach or taking advantage of the longitudinal data with more than 12 years of follow-up available for the Danish nonagenarian sample.

In conclusion, despite a substantial power, we did not confirm *AKT1* as a longevity-associated gene in populations of Danish and German nonagenarians and centenarians. This finding does not necessarily rule out a potential relevance of *AKT1* in longevity since population genetic differences and incomplete coverage of allelic variation in *AKT1* may explain the lack of replication. However, all in all our results do suggest that *AKT1* is not a universally acting longevity gene.

2296T

Association of _HLA-DRB1_ and _HLA-DQB1_ genes with asthma in the Spanish population. M. Pino-Yanes[1,2,3], A. Corrales[1,2], M. Acosta-Herrera[1,2,10], J. Cumplido[4], F. Sánchez-García[5], R. González[6], J. Figueroa[7], I. Sánchez-Machín[6], O. Acosta[8], A. Sánchez-Palacios[7], J.C. García-Robaina[9], J. Villar[1,10], E.G. Burchard[3], T. Carrillo[4], C. Flores[1,2]. 1) CIBER de Enfermedades Respiratorias, Instituto de Salud Carlos III, Madrid, Santa Cruz de Tenerife, Spain; 2) Research Unit, Hospital Universitario N.S. de Candelaria, Tenerife, Spain; 3) Department of Medicine, University of California San Francisco, San Francisco, CA; 4) Allergy Unit, Hospital Universitario Dr. Negrín, Gran Canaria, Spain; 5) Immunology Unit, Hospital de Gran Canaria Dr. Negrín, Gran Canaria, Spain; 6) Allergy Unit, Hospital del Tórax, Complejo Hospitalario Universitario NS Candelaria, Tenerife, Spain; 7) Allergy Unit, Hospital Universitario Insular de Gran Canaria, Gran Canaria, Spain; 8) Neumology Unit, Hospital Universitario de Canarias, Tenerife, Spain; 9) Allergy Unit, Hospital Universitario N.S. de Candelaria, Tenerife, Spain; 10) Multidisciplinary Organ Dysfunction Evaluation Research Network (MODERN), Research Unit, Hospital Universitario Dr. Negrin, Gran Canaria, Spain.

Before the advent of genome-wide association studies (GWAS), *HLA-DRB1* and *HLA-DQB1* genes were independently associated with asthma in several candidate gene association studies. After years of oblivion, the importance of these genes in the genetic of asthma has been recently corroborated by GWAS studies and one GWAS meta-analysis. In spite of these evidences, the classic alleles underlying these associations have not been determined. Here we aimed to test the reproducibility of *HLA-DRB1* and *HLA-DQB1* associations with asthma in a Spanish sample, and to explore the classic alleles involved in the disease. For this purpose, 22 single nucleotide polymorphism (SNPs) with ability to predict classic alleles from *HLA-DRB1* and*HLA-DQB1* were genotyped using TaqMan®, KASPar® and SNaPshot® in 313 training samples typed for the classic alleles of these genes. After checking the adequate prediction of classic alleles at 2 digit-resolution, these 22 SNPs were genotyped in a case-control sample of 1865 unrelated participants consisting of 606 asthmatics and 1259 controls. Results replicated the association of individual SNPs with asthma after population stratification adjustments ($1.12 \leq OR \leq 1.94$, $2.8 \times 10^{-7} \leq p\text{-value} \leq 0.048$). In addition, the imputation of classic alleles allowed identifying both protective and risk alleles in both genes. Taken together, our results support the important role of these genes in asthma susceptibility.

Supported by FUNCIS 23/07 and grants from the Spanish Ministry of Science and Innovation FIS PI08/1383, FIS PI11/00623 and EMER07/001 to CF. MPY was supported by Fundación Canaria Doctor Manuel Morales (Spain).

2297F

Turning the GWAS upside down: A PheWAS approach in studying human disease. S. Hebbring[1,2], S. Schrodi[1], Z. Ye[1], Z. Zhou[1], D. Page[2], M. Brilliant[1]. 1) Marshfield Clinic, Marshfield WI; 2) University of Wisconisin Madison, Madison WI.

In classical genetics, there are two common experimental approaches: “reverse” and “forward” genetics. Forward genetics is the process of going from phenotype to genotype and has been popularized in linkage and genome wide association studies (GWAS). Reverse genetics is the process of going from genotype to phenotype and has been a powerful tool in studying the genetics of animal models. In an attempt to apply a reverse genetics approach to study human disease, we screened 114 putative loss-of-function nonsense polymorphisms in 112 genes, along with four positive control SNPs, with 4841 unique diagnostic codes (ICD9 codes) in 4241 Marshfield Clinic patients as part of the Personalized Medicine Research Project. As anticipated, many of the positive control SNPs were significantly associated with expected ICD9 codes. Previous studies have shown rs2200733 is associated with atrial fibrillation utilizing forward genetic approaches (i.e., GWAS). In this study, ICD9 code for atrial fibrillation (427.31) was one of the top four ICD9 codes associated with rs2200733 (p=0.00024). Interestingly, positive control SNP rs17234657, previously shown to be associated with Crohn's disease, was not associated with Crohn's (43 reported cases with ICD9 555), but the variant allele increased the risk for other inflammatory conditions such as chronic rhinitis (ICD9 472.0, p=0.0012), gastritis (ICD9 535.41, p=0.0040), and sinusitis (ICD9 461, p=0.0091). Aside from the positive control SNPs, novel associations were observed for many of the nonsense SNPs screened. Strong associations were observed for rs28502153, a nonsense SNP within GAB4. Rs28502153 was associated with numerous muscular-skeletal phenotypes affecting the limbs (ICD9 755, p=4.1e-5), fingers (ICD9 736.2, p=0.00083), joints (ICD9 728.3 and 728.5, p=0.0003 and p=0.0031 respectively), and posture (ICD9 781.92, p=0.0093). GAB4 is believed to interact with GRB2, a growth factor receptor binding protein necessary in development. In summary, these results demonstrate that genotype-phenotype relationships previously characterized using forward genetics can be observed when studied in the reverse direction while characterizing other related, yet novel, relationships. Second, with proof-of-principle established, reverse genetics maybe a powerful and complementary approach in studying human genetics. This is the first known attempt to systematically screen nonsense polymorphisms for disease risk.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2298W

Assessment of Bias in an EMR-derived Cohort. *W.S. Bush, L.C. Dumitrescu, D.C. Crawford.* Ctr Human Gen, Vanderbilt Univ, Nashville, TN.

Large-scale biological repositories linked to electronic medical records (EMRs) are quickly becoming feasible alternatives to more traditional population-based cohorts for conducting genetic epidemiological studies. The ability to rapidly obtain high-quality DNA from a large number of patients, and then link data from high-throughput genetic assays to patient-specific EMR data, while meeting IRB guidelines for confidentiality, provides the needed infrastructure for research in pharmacogenetics toward individualized medicine. However, extension of a hospital-based approach to etiologic research may be more susceptible to bias if the genetic variant is in some way associated with healthcare utilization, diagnosis, or treatment patterns at that clinic. The most appropriate controls to represent the source population of the cases are not easily defined in this context, and the possible differences in cases and controls may introduce confounding, reduce statistical power, or increase the false positive rate. In this study, we compare the risk estimates for established cancer risk loci from published population-based cohort studies to the risk estimates generated by the EMR-derived cross-sectional study conducted using BioVU, the Vanderbilt DNA Biobank.

2299T

Defining the use of electronic medical records in genetic studies of multiple sclerosis. *M.F. Davis[1], J.L. McCauley[2], J.C. Denny[3], J.R. Oksenberg[4], S.L. Hauser[4], M.A. Pericak-Vance[2], J.L. Haines[1].* 1) Center for Human Genetics Research, Vanderbilt University Medical Center, Nashville, TN; 2) John P. Hussman Institute for Human Genomics, University of Miami, Miller School of Medicine, Miami, FL; 3) Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN; 4) Department of Neurology, School of Medicine, University of California, San Francisco, CA.

Multiple sclerosis (MS) is a neurodegenerative disease with complex genetic underpinnings and clinical heterogeneity. The HLA locus was identified as a critical genetic determinant of the disease in the 1970s, but no other genes or loci were discovered until the recent GWAS era. Over 80 genome-wide significant and confirmed loci exist in the published literature for MS, but these loci account for a small proportion of the genetic contribution of MS. It is possible that these loci affect the expression of MS rather than contributing significantly to the burden of risk for developing the disease. As yet, previous studies have lacked access to long-term clinical information, and few studies have analyzed the genetics of the various clinical aspects of MS. At Vanderbilt University, we have the unique opportunity to access the Synthetic Derivative (SD), a de-identified database of over 2 million electronic medical records (EMR). Records in the SD begin in 1991 and contain all clinic notes, problem lists, letters between physicians, prescribed medications, ICD-9 billing codes, and imaging reports for patients at Vanderbilt. The SD provides us access to the records of 6,022 patients with MS and we are undertaking a study to characterize the clinical information available from the EMR that can be used for genetic studies of MS. We are investigating the range and depth of information that can be extracted, especially age of onset, progression (EDSS, 25-foot timed walk), first neurological symptom, and response to treatments, with the goal of performing genetic analyses using these sub-phenotypes to better understand the genetic heterogeneity of MS. We have developed algorithms to extract this information. Currently, we have been able to extract the subtype of MS for 4,800 individuals, EDSS score for 1,700 individuals, 25-foot timed walk score for 3,500 individuals and age of diagnosis for 3,000 individuals. Through the BioVU resource at Vanderbilt, we have DNA samples for 1,000 individuals with MS and 2,500 control subjects and have genotyped them on the Immunochip for genetic analyses of these traits. The dataset is approximately 90% Caucasian and 10% African-American. Genetic analyses are currently ongoing. Through these efforts we have seen that EMRs provide a rich resource of detailed clinical data, which will be critical to further uncovering the genetic component of MS.

2300F

The NINDS Repository: A Public Resource of Biomaterials for Genetic, Mechanistic and Biomarker Discovery for Neurological Disorders. *A. Scutti[1], M. Self[1], K. Reeves[1], J. Santana[1], S. Heil[1], C. Pérez[1], A. Nasuti[1], R. Corriveau[2], C. Tarn[1].* 1) NINDS Repository, Coriell Institute for Medical Research, Camden, NJ; 2) National Institute for Neurological Disorders and Stroke-NIH, Bethesda, MD.

The burden of neurological disorders is a serious health concern and presents a great challenge to healthcare systems globally. In most cases, the causes of neurological disorders are multifactorial, involving interactions of genetic and environmental factors. In neurodegenerative disorders such as Huntington's disease, Parkinson's disease, and Amyotrophic Lateral Sclerosis (ALS), clear mechanisms of disease are not well understood and the lack of reliable biomarkers for early diagnosis hampers timely intervention. Studies aimed at identifying mechanisms of disease onset will likely require large numbers of samples and a variety of biomaterials from affected and unaffected individuals. The National Institute of Neurological Disorders and Stroke (NINDS) Repository is a public resource established to provide a centralized resource of biological samples (DNA, cell lines, fibroblasts, induced pluripotent stem cells [iPSCs], plasma, serum, CSF, and urine) and corresponding clinical and phenotypic data from a diverse population of neurological patients and neurologically normal controls. To date, biomaterials from more than 38,000 individuals with cerebrovascular diseases (>12,000), Parkinson's disease (>7,000), motor neuron diseases (>2,700), epilepsy (>5,000), Tourette syndrome (>3,700), dystonia (>390), Huntington's disease (>30) and neurologically-normal controls (>7,000) have been successfully banked in the NINDS Repository. More than 23,000 samples are currently available through the NINDS Repository web catalog. To aid high-throughput gene discovery, the NINDS Repository offers disease- and neurologically-normal control panels of genomic DNA in 96-well plate formats and custom plate design services. Additionally, genotype data from over 4,000 NINDS Repository samples are available through the NIH's database of Genotypes and Phenotypes (dbGaP). Furthermore, the NINDS Repository establishes and banks fibroblast cell cultures and iPSCs from individuals with neurodegenerative disorders and known genetic mutations. Recently, the NINDS Repository began expanding its collection to include other types of biofluids such as serum, cerebrospinal fluid, and urine, with the goal of facilitating biomarker discovery in neurodegenerative diseases. By establishing a centralized, publicly accessible resource of biospecimens and associated clinical data, the NINDS Repository facilitates discovery of new genetic and molecular signatures related to neurological disorders.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2301W

Mutation of *SLC45A3* in familial and sporadic Menière's disease. *C.A. Campbell[1], B.J. Boese[2], C.C. Della Santina[3], C. Li[4], N.C. Meyer[1], L.T. TeGrootenhuis[1], J. Webster[5], D.A. Stephan[6], H. Najmabadi[7], A. Danashi[7], A.P. May[8], K. Fredrickson[2], B.J. Gantz[1], J.P. Carey[3], L.B. Minor[3], M.R. Hansen[1], T.T. Harkins[9, 10], R.J.H. Smith[1].* 1) Otolaryngology - Head and Neck Surgery, University of Iowa, Iowa City, IA; 2) 454 Life Sciences, Branford, CT; 3) Otolaryngology-Head & Neck Surgery, The Johns Hopkins University, Baltimore, MD; 4) Biostatistics, Harvard School of Public Health, Boston, MA; 5) Neurogenomics Division, Translational Genomics Research Institute, Phoenix, AZ; 6) The Institute For Individualized Health, Palo Alto, CA; 7) Genetics Research Center, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran; 8) Fluidigm Corporation, South San Francisco, CA; 9) Roche Applied Science, Indianapolis, IN; 10) Present address: Life Technologies, Beverley, MA.

Menière's disease (MD) is a complex idiopathic disorder of the inner ear characterized by the symptom triad of vertigo, sensorineural hearing loss and tinnitus. The incidence of MD varies around the world and it is estimated to affect 1-2 per 1,000 Caucasian individuals. Most cases of MD are sporadic although 5-14% of individuals will report a family history in which MD segregates in an autosomal dominant fashion with incomplete penetrance. To date, a genetic contribution to sporadic and/or familial MD has not been identified reflecting the challenges that complex diseases present to causative gene discovery.

We previously performed a genome-wide linkage scan followed by targeted exon capture and pyrosequencing of the candidate interval (1q32) to identify a deletion in *SLC45A3* which segregated with disease in a Chilean family. We hypothesized a gene identified in the rare familial form of MD would also be involved in the sporadic form of MD.

As a test of the common-disease-rare-variant hypothesis we performed candidate gene sequencing in a large MD cohort (n=250) using a Fluidigm Access Array followed by pyrosequencing as well as Sanger sequencing, and identified the same deletion in two sporadic patients. We subsequently screened 320 sporadic MD patients for the deletion and identified three additional individuals with the identical 36-bp deletion. We found the encoded protein, Prostein to be expressed in the cochlea and endolymphatic sac where it likely plays a role in osmotic regulation.

In summary, through the use of two complementary human genetics approaches - 1) segregation analysis with targeted-sequence capture and massively parallel sequencing; and, 2) candidate gene sequencing in a large MD cohort as a test of the common-disease-rare-variant hypothesis - we identified the first gene causally related to Menière's disease, *SLC45A3*. This study is exciting as we have identified a deletion in *SLC45A3* found in individuals with MD across five populations. The approach is broadly applicable to other complex diseases. (This research was supported in part by a grant from the American Otological Society (RJHS).).

2302T

Next generation sequencing identifies genes dysregulated in otosclerotic stapes bone and provides a candidate gene set for filtering of whole exome sequence data in familial cases. *J.L. Ziff[1], J. Lavy[2], H. Powell[2], S. Khalil[2], K.P. Steel[3], S.R. Saeed[1,2], S.J. Dawson[1].* 1) UCL Ear Institute, University College London, United Kingdom; 2) Royal National Throat, Nose and Ear Hospital, London, United Kingdom; 3) Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Cambridge, United Kingdom.

Otosclerosis is a common acquired form of conductive hearing loss, characterised by abnormal bone remodeling within the otic capsule, leading to fixation of the stapes bone in the middle ear. Evidence suggests that it is unique to humans and occurs more frequently in females and individuals of Caucasian origin. The aetiology of otosclerosis is unknown; it is considered a complex disorder but also has rare monogenic forms. A number of investigative strategies have been utilised to identify genes involved in the disease process, however, to date no definitive causative genes have been identified. To improve our understanding of the disease pathology, we compared the transcriptomes of 8 otosclerotic and 4 control stapes samples using Illumina HiSeq 2000 paired end RNA sequencing. In subsequent analysis Z ratios were calculated for each gene to distinguish significant changes in gene expression between the diseased and control tissue. In total, 176 genes were found to be significantly differentially expressed in the otosclerotic stapes ($p<0.05$) and to have a fold difference >1.5. Of these 97 were up-regulated and 79 down-regulated. Pathway analysis shows that the dysregulated gene set is enriched for genes involved in TGFβ signalling, collagen fibril formation and calcium binding. Whole exome sequencing has also been performed on 10 individuals across 4 different families, each exhibiting monogenic inheritance of otosclerosis: - two non-immediate family members with otosclerosis from each of two unrelated Caucasian British families; three immediate family members with otosclerosis from a large Caucasian Portuguese family with multiple consanguineous marriages; two affected and one unaffected siblings from a family of Caribbean origin. The candidate gene set obtained from RNA sequencing will be utilised to filter the exome sequencing data with the aim of identifying the pathogenic variant in each of the 4 families.

2303F

Dissection of healthy ageing by whole genome and exome sequencing of supercentenarian genomes. *X. Estivill[1], N. Badarinarayan[2], P. Rosenstiel[2], M. Bayés[3], S. Ossowski[1], C. Tornador[1], H. Blanche[4], G. Escaramís[1], M. Gut[3], S. Heath[3], F. Flachsbart[2], A. Nebel[2], M. Lathrop[3], A. Franke[2], M. Gratacòs[1], S. Schreiber[2,5], I. Gut[3], D. Trujillano[1].* 1) Genes & Disease Program, Center Genomic Regulation, Barcelona, Catalonia, Spain; 2) Institute of Clinical Molecular Biology, Christian-Albrechts-University of Kiel, Kiel, Germany; 3) National Center of Genomic Analyses Barcelona, Catalonia, Spain; 4) CEPH, Paris, France; 5) Department of General Internal Medicine, University Hospital S.-H., Kiel, Germany.

Ageing is the major risk factor for the development of chronic disorders, including cancer, cardiovascular, osteoarticular and neurodegenerative diseases. In Europe it is expected that there will be 1.6 million citizens 100+ years old by 2050, a dramatic increase with respect to the nearly 40,000 in 1996. Several individuals have been reported to live over 110 years (supercentenarians). Since the prevalence of supercentenarians is very low, it can be considered a “rare” phenotypic condition, which could provide important insights on key factors of healthy ageing and disease protection. We hypothesize that extraordinary long-lived individuals are gifted with a combination of common and rare protective genetic factors. Thus, in this project we wanted to identify both novel and known genetic variants that influence lifespan in exceptionally old individuals. To reach this goal we carried out whole-genome and whole-exome sequencing of four supercentenarians of European origin (one Spanish, one French and two German). The high depth of sequencing achieved by combining the Illumina and Solid platforms allowed us to confidently mine the data not only for single nucleotide variants (SNV) of interest, but also novel SNV discovery involved in healthy aging. The selected SNVs are being tested for association in a larger cohort of centenarians and controls by means of direct genotyping for common SNVs and targeted resequencing for novel and rare SNVs. In addition the data allowed us a precise detection of indels and structural variants. We identified hundreds of potentially disease-protective variants in our cohort of centenarians. Our analysis pipeline focused on identifying novel, rare, and common SNVs affecting genes involved in pathways associated to age-related diseases. Targeted-resequencing 222 genes enriched in novel and rare SNVs in a validation cohort of 1,000 centenarians and controls from the same geographic regions, and genotyping in the validation cohort a total of 120 common SNVs, that show altered allelic frequencies in our initial sample set when compared to the normal population, allow to draw a network of genes, and variants that show protective effects. Whole-genome and whole-exome sequencing of exceptionally old individuals provides a high level of resolution of SNV discovery to dissect the molecular basis of healthy aging. The results of association studies in centenarian samples should further substantiate these findings.

2304W

Whole Exome Sequencing Cases: Finding and Testing with External Controls. *A. Hendricks on behalf of the UK10K Statistics Group.* Wellcome Trust Sanger Institute, Hinxton, Cambridge, United Kingdom.

It is common for researchers undertaking whole exome sequencing (WES) studies to sequence disease cases and to seek externally available controls. Finding an appropriate control set to use in case-control analysis is an open and important question that is composed of two parts: finding a suitable set of controls and developing appropriate analysis methods and protocols. Here, we explore the ability of various control sets and analysis methods to produce an unbiased distribution of test statistics using a large set of sequenced samples from the UK10K project (www.uk10k.org). Within UK10K we are sequencing 10,000 individuals: 4,000 cohort participants using low depth (6x) whole-genome sequencing (WGS) and 6,000 disease cases using WES (50x). Currently, over 2,400 cohort participants and over 3,000 cases have gone through sequencing, quality control (QC), and variant calling. These data provide a unique opportunity to assess case-control analysis over a variety of samples, traits, possible confounders such as population stratification, sequencing types (WGS and WES), and sequencing attributes such as depth and quality. We started by analyzing single nucleotide variants in WES cases and controls and investigated tests restricted to fairly homogenous samples to assess the impact of simple QC filtering. We first filtered variants for depth and genotype quality on the individual level and then applied sample-wide variant filters. We found that after applying filters on minor allele frequency, call rate, and average quality by depth we produced an observed test statistic distribution similar to the expected distribution. Our particular set of filters controlled for inflation even when case status was assigned so that confounding factors such as batch-specific differences in depth were present. Further, we investigated using WGS cohort samples as population controls for WES cases. Preliminary results suggest that after applying individual level variant filters as well as sample-wide variant filters the observed distribution of test statistics is similar to the expected distribution. Although further research is necessary, this suggests that publicly available WGS cohort samples such as 1000 Genomes or UK10K could possibly be used as controls for WES cases. Finally, while we have focused initially on single variant tests, we will expand our comparisons and development to locus-based analyses.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2305T

Whole-exome sequencing study of four families with bipolar disorder. *E. Elhaik[1,2], M. Pirooznia[1], F.S. Goes[1], J. Parla[5], R. Karchin[3], A. Chakravarti[2], P.P. Zandi[1], R.W. McCombie[5], J.B. Potash[4].* 1) Department of Mental Health, Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD; 2) McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD; 3) Medical Research Center, University of Iowa, Iowa City, IA; 4) Department of Biomedical Engineering and Institute for Computational Medicine, Johns Hopkins University, Baltimore, MD; 5) Cold Spring Harbor Laboratory, Woodbury, NY.

Background: Bipolar disorder (BP) is a common mental disorder often associated with lifelong disability and premature mortality. We are conducting a whole-exome study of BP using next generation sequencing to examine the whole exomes of up to 100 multiplex BP families (with at least 6 individuals from 2-3 generations of each family), 1,800 BP cases, and 1,800 controls with the goal of identifying rare and common genetic variants associated with the disease. **Methods**: Exome sequencing was performed in a pilot sample of 22 individuals from four multiplex BP families using solution-based capture and paired-end sequencing on the Illumina GA II. Alignment and variant calling were performed with BWA, SAM tools, and GATK. SNVs were annotated with the SIFT and PolyPhen tools. Families were analyzed separately for the segregation of functionally relevant variants with disease. **Results**: We identified a single common (MAF=0.2) deleterious splice site variant (rs8373) in a zinc-finger protein gene (ZFP91) that segregated with all affected relatives and none of the unaffected married-in relatives in all four families. A family-based test indicated the variant was significantly associated with BP in these families (p=0.0026). ZFP91 is involved in the non-canonical nuclear factor kB (NF-kB) signaling pathway, which regulates the canonical NF-kB pathway. The non-canonical pathway is associated with adaptive immunity and protection against inflammation and apoptosis. **Conclusions**: Our initial analysis of four multiplex families with bipolar disorder revealed a common splice-site polymorphism in ZFP91 that segregates with disease in all pedigrees. Mutations inhibiting the non-canonical NF-kB, such as the one identified here, have been shown to induce apoptosis and inflammation due to the continuous activation of the complementary pathway. This variant was imputed in the Psychiatric GWAS Consortium (PGC) mega-analysis of Bipolar Disorder, but it was not significantly associated with illness. However, if the current finding can be replicated in other sequenced families, it may provide evidence of a potential inflammatory etiology in bipolar disorder. Such replication efforts are ongoing.

2306F

Whole exome sequencing of five consanguineous families with juvenile myoclonic Epilepsy. *B. ouled amar bencheikh[1,2], J. Gauthier[1], F. Lahjouji[2], I. Jouan[1], P. La chance touchette[1], K. Khaldi[2], H. Lamghari[2], C. Bourassa[1], D. Spiegelman[1], A. Dionne laporte[1], H. Daoud[1], F. Hamdan[1], H. Belaidi[2], O. Ouazzani[2], P. Dion[1], P. Cossette[1], G. Rouleau[1].* 1) Center of Excellence in Neuroscience of University, Centre de recherche du CHUM, Montréal, Québec, Canada; 2) Service de Neurophysiologie Clinique, Hopital des Spécialites, CHU Ibn Sina, Rabat, Maroc.

Juvenile myoclonic epilepsy (JME) is one of the forms of genetic (Idiopathic) generalized epilepsies (GGEs), which in turn account for 40 to 60% of all epilepsies. GGEs are characterized by a complex genetic mode of transmission most likely due to phenotypic, genetic heterogeneity and incomplete penetrance of the disease. Linkage analyses have had limited success in the identification of genes involved in GGE/JME. The recent development of Next generation sequencing methods is currently facilitating the understanding of Mendelian and complex diseases. We selected 13 patients from five consanguineous families of juvenile myoclonic epilepsy (JME) compatible with an autosomal recessive transmission and performed exome sequencing. Previously identified loci and causative genes for JME have been excluded by linkage analysis. The targeted exome capture was done using the Agilent SureSelect V4 kit and the sequencing was done on an Illumina HiSeq2000 system. The variants were called and annotated using the GATK software and the ANNOVAR program, respectively. Called variants were filtered against dbSNP, 1000 Genomes Project and In-house database (data from in-house exome resequencing projects). We considered all variants homozygous shared between all affected individuals from each family. Only non-synonymous, splice site and coding InDels were included in the analysis. Our best candidate variants were then validated by Sanger sequencing on both strands. The most promising variants segregating with the disease in these consanguineous families will be validated by functional studies. The exome sequencing approach using consanguineous families has great potential to identify new genes of JME which in turn would lead to new insights for GGE and open new avenues for the development of novel therapeutic targets.

2307W

GEnetics of Nephropathy - an International Effort (GENIE): CNV Analysis. *Y. Meng[1], R.M. Salem[1,2,3], A.J. McKnight[7], E.P. Brennan[8,9], N. Sandholm[4,5,6], C. Forsblom[4,5], J. Florez[1,3,10], C. Godson[8,9], P-H. Groop[4,5,11], A.P. Maxwell[7,12], J.N. Hirschhorn[1,2,3] for the GENIE Consortium.* 1) Population and Medical Genetics, Broad Inst, Cambridge, MA, USA; 2) Endocrine Research Unit, Department of Endocrinology, Children's Hospital, Boston, MA, USA; 3) Department of Medicine, Harvard Medical School, Boston, MA, USA; 4) Folkhälsan Institute of Genetics, Folkhälsan Research Center, Biomedicum Helsinki, Helsinki, Finland; 5) Division of Nephrology, Department of Medicine, Helsinki University Central Hospital, Helsinki, Finland; 6) Department of Biomedical Engineering and Computational Science, Aalto University, Finland; 7) Nephrology Research, Centre for Public Health, Queen's University of Belfast, UK; 8) Diabetes Research Centre, Conway Institute, School of Medicine and Medical Sciences, University College Dublin, Dublin, Ireland; 9) Mater Misericordiae Hospital, Dublin, Ireland; 10) Center for Human Genetic Research, Massachusetts General Hospital, Boston MA, USA; 11) Baker IDI Heart and Diabetes Institute, Melbourne, Australia; 12) Regional Nephrology Unit, Level 11, Tower Block, Belfast City Hospital, Belfast, UK.

Diabetic nephropathy (DN) is a leading cause of kidney failure. GWA studies of SNPs have been performed for DN, including by our consortium (GENIE), but the role of copy number variation in DN has not been investigated. Copy number variant (CNV) analysis, genotyped on Illumina HumanOmni1-Quad, was performed in 1,726 samples from the UK-ROI cohort within the GENIE consortium. To assess whether CNVs are associated with DN, we estimated copy number at known copy number polymorphisms (CNPs) as well as novel CNVs using Birdsuite and performed association analysis using PLINK. All participants in UK-ROI have type 1 diabetes (T1D), and are classified as normoalbuminuric controls (at least 15 years of T1D duration), end stage renal disease (ESRD), or DN (either macroalbuminuria or ESRD). A total of 1,399 samples passed quality control. Three contrasts were defined for association testing: DN vs. controls, ESRD vs. controls, and ESRD vs. all. We tested association with individual CNPs, association with global CNV burden, and regional associations with rare (≤1%) CNVs, analyzing CNVs occurring within defined genomic or genic regions. No individual CNP showed significant associations after correcting for multiple testing. In the burden analysis, subjects with DN had a greater global burden of CNVs than controls ($p \leq 0.05$), including an excess of lower-frequency (≤5%) and rare (≤1%) duplications, and rare (≤1%) duplications overlapping genes. No significant associations were seen for burden of deletions, and the signal for duplications was not significant after correction for multiple testing. Rare CNVs (≤1%) within 20 kb of each gene in the genome were analyzed, testing for an excess of CNVs in cases. Significance was assessed by 1 million permutations using a 1-sided test. We identified three genic regions on chromosomes 2, 8, and X with an excess of CNVs in DN cases ($p \leq 10^{-4}$, 14-17 CNVs in cases vs. 0 CNVs in controls in each region); one genic region on chromosome 11 with an excess of CNVs in ESRD cases ($p \leq 10^{-4}$, 6 CNVs in cases vs. 0 CNVs in controls). Replication and validation studies, as well as pathway-based analyses, are being initiated. In summary, we performed the first copy number association analyses of DN, and preliminary results have highlighted four gene regions with a possible excess of copy number variation in patients with DN compared with controls.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2308T

Whole genome sequencing identifies coding variants influencing canonical lipid species in Mexican American families. *C. Bellis[1], P.J Meikle[2], J.M Weir[2], J.B Jowett[2], T.M Teslovich[3], G. Jun[3], S. Kumar[1], M. Almeida[1], J.M Peralta[1], E.E Quillen[1], C. Fuchsberger[3], A.R Wood[4], T.M Frayling[4], P. Cingolani[5], T.W Blackwell[3], R. Sladek[5], G. Atzmon[6], J. Laramire[7], S. Lincoln[7], D.M Lehman[8], G. Abecasis[3], M.C Mahaney[1], T.D Dyer[1], L. Almasy[1], R. Duggirala[1], J. Blangero[1], J.E Curran[1].* 1) Department of Genetics, Texas Institute for Biomedical Research, San Antonio, TX; 2) Baker IDI Heart and Diabetes Institute, Melbourne, VIC, Australia; 3) University of Michigan, Ann Arbor, Michigan, USA; 4) University of Exeter, Exeter, United Kingdom; 5) McGill University, Montreal, Quebec, Canada; 6) Albert Einstein School of Medicine, Bronx, New York, USA; 7) Complete Genomics Inc., Mountain View, California, USA; 8) University of Texas Health Center at San Antonio, San Antonio, Texas, USA.

The lipidome encompasses the complete universe of fundamental lipid species. It represents a potential gold mine of clinically relevant phenotypes that may be better predictors of disease risk than those that are commonly studied. Additionally, the biologically simpler nature of such lipid species presents the hypothesis that determinants may reside closer to the causal action of genes than more complex integrated lipid measures, such as total cholesterol level, making them valuable phenotypes for finding genes involved in lipid metabolism. Implementation of a targeted capture method involving liquid chromatography electrospray ionization-tandem mass spectrometry provided precise identification and quantification of 356 lipid measures in 1202 Mexican American individuals from large pedigrees in the San Antonio Family Heart Study (SAFHS). Genome-wide association analyses revealed several significant QTL localizations for these canonical lipid phenotypes, however by harnessing the power of whole genome sequence (WGS) data the chance of identifying a functional variant influencing lipid metabolism is significantly increased. We currently have lipidomic profiles and WGS data available for 596 Mexican American individuals, and will be increasing this to 2,000 in the coming months. In this study, we focus on the potential effects of obligately functional protein altering variants. Analysis of the WGS data identified 44,985 such non-synonymous coding variants, of which 11,484 are predicted to be highly deleterious (based on PolyPhen prediction, having a score > 0.7). Preliminary variant-specific and gene-specific burden association results have identified several rare coding variants apparently influencing lipid species including a highly deleterious rare variant in the GPIHBP1 gene previously known to be involved in severe hypertriglyceridemia. In this gene, an S144F mutation is highly associated with multiple forms of lysophosphatidylcholine. These results illustrate the utility of WGS analysis for direct identification of functional targets influencing lipid measures.

2309F

Identifying non-synonymous variants exhibiting pleiotropic effects on preeclampsia and CVD risk traits. *M.P. Johnson[1], S.P. Brennecke[2], J.M. Peralta[1], T.D. Dyer[1], M.A. Almeida[1], J.M. Proffitt[1], C.E. East[2], A.G. Comuzzie[1], M.C. Mahaney[1], L. Almasy[1], J.E. Curran[1], J. Blangero[1], E.K. Moses[1,3].* 1) Department of Genetics, Texas Biomedical Research Institute, San Antonio, TX, USA; 2) Department of Perinatal Medicine, Royal Women's Hospital & Department of Obstetrics and Gynaecology, University of Melbourne, Parkville, Victoria, Australia; 3) Centre for Genetic Epidemiology and Biostatistics, University of Western Australia, Perth, Western Australia, Australia.

Preeclampsia is a serious pregnancy disorder promoting morbidity and mortality to both mother and child. Delivery of the fetus is the only means to resolve severe symptoms. Women with preeclamptic pregnancies demonstrate increased risk for later life cardiovascular disease (CVD) and good evidence suggests these two syndromes to share several risk factors and pathophysiological mechanisms. We have previously identified preeclampsia susceptibility loci on chromosomes 2q, 5q and 13q using a Caucasian Australian family-based study design. To identify preeclampsia candidate genes at our 2q locus we initially employed an exon-centric Sanger sequencing strategy with follow up genotyping in an extended set of Australian preeclampsia families. Based on our gene-centric multiple testing criteria we identified a significant association (p=2.7E-03) for a missense *LCT* variant (N1639S) with preeclampsia susceptibility. We replicated this association (p=2.0E-02) in an independent Caucasian Australian case-control population. Furthermore, in a large Mexican American familial cohort the N1639S variant exhibited pleiotropic effects with several CVD-related risk traits: AGE products (p=7.3E-03), GR (p=8.9E-03), IL-6 protein (p=8.0E-03), and soluble VCAM1 (p=4.2E-02). More recently, we have employed a next-generation exome sequencing strategy to identify family-specific loci segregating in preeclamptic women only. The exomes of 7 women (3 preeclamptics, 4 controls) from an Australian family contributing to our 5q locus were sequenced using Illumina's TruSeq Exome assay. As a first pass, we prioritized the identification of non-synonymous variants within the 1-LOD localization support interval of the 5q locus. We identified a missense variant (A2803T) in the *GPR98* gene to segregate in the preeclamptic women but not in the unaffected women. Additional next-generation exome sequencing in the Mexican American families identified three *GPR98* missense variants (P1987L, Y2232C, N2345S) to be significantly associated with HDL cholesterol levels (p<1.24E-06). Given that hypertension is a defining feature of preeclampsia it is also of major interest to us that a recent GWAS has reported a significant association with diastolic blood pressure for a SNP at the *GPR98* locus in African Americans. Taken together, these data present for the first time empirical evidence of possible shared genetic mechanisms underlying both preeclampsia and other CVD-related risk traits.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2310W

Analysis of Coding Variants Identified from Exome Sequencing in Diabetic and Non-diabetic Nephropathy Genes in African Americans. *J.N. Cooke[1, 3, 4], N.D. Palmer[2, 3, 4], M.C.Y. Ng[3,4], P.J. Hicks[2], J.A. Bonomo[1, 3, 4], J.M. Hester[3, 4, 5], C.D. Langefeld[6], B.I. Freedman[7], D.W. Bowden[2, 3, 4, 8].* 1) Molecular Medicine and Translational Science, Wake Forest School of Medicine, Winston-Salem, NC; 2) Department of Biochemistry, Wake Forest School of Medicine, Winston-Salem, North Carolina; 3) Center for Genomics and Personalized Medicine Research, Wake Forest School of Medicine, Winston-Salem, North Carolina; 4) Center for Diabetes Research, Wake Forest School of Medicine, Winston-Salem, North Carolina; 5) Program in Molecular Genetics and Genomics, Wake Forest School of Medicine, Winston-Salem, North Carolina; 6) Department of Biostatistical Sciences, Wake Forest School of Medicine, Winston-Salem, North Carolina; 7) Department of Internal Medicine - Section on Nephrology, Wake Forest School of Medicine, Winston-Salem, North Carolina; 8) Department of Internal Medicine - Section on Endocrinology, Wake Forest School of Medicine, Winston-Salem, North Carolina.

Type 2 diabetes-associated end-stage renal disease (T2D-ESRD) and other non-diabetic forms of nephropathy disproportionately affect African Americans. Prior studies have identified common genetic variants contributing to these diseases; however, novel exome sequencing techniques allow detection of coding variants that may confer additional genetic risk. We identified putative functional coding single nucleotide variants (SNVs) in genes with compelling evidence for association with nephropathy in the NHLBI Exome Sequencing Project database. SNVs were genotyped in 5244 African Americans (n=3438 cases with T2D- or non-T2D ESRD, n=1806 controls) and 1532 European Americans (n=613 T2D-ESRD cases, n=919 controls). Logistic regression analyses were performed to assess the association with T2D-ESRD, non-T2D ESRD, and overall nephropathy; admixture and APOL1 risk status covariates were incorporated in African American analyses. In African American T2D-ESRD analyses, four SNVs were associated (P<0.05; additive model unless specified): rs9481410 in C6orf167 (P= 0.024 [OR=0.89; 95% CI 0.81-0.99]), rs116147257 in APOL3 (P=0.00066 [OR=0.67; 95% CI 0.53-0.84]), rs7285158 in APOL2 (P=0.011 [OR=1.16; 95% CI 1.03-1.29]), rs2239785 in APOL1 (P=0.0012 [OR=0.83; 95% CI 0.74-0.93]). Two SNVs were associated with non-T2D ESRD: rs4925583 in OR2L13/OR2L8 (P=0.0068 [OR 0.84; 95% CI 0.73-0.95]) and rs7285167 in APOL2 (P=0.021 [OR=1.17; 95% CI 1.02-1.34]). Three SNVs replicated association in European American T2D-ESRD: rs4925583 (P=0.047 [OR= 0.73; 95% CI 0.53-1.00]), rs9481410 (P=0.020 [OR=1.22; 95% CI 1.03-1.45]), rs7285167 (P=0.0077 [OR=0.95; 95% CI 0.79-1.15]). Additionally, variants in CNDP1 and OR2L13/ORAK2 were associated with T2D-ESRD in European Americans (rs3747154, P=0.00094, OR=4.95 [95% CI 1.61-15.23]; rs4478844, P=0.0069 [recessive], OR=1.53 [95% CI 1.12-2.08], respectively) but were not associated in African Americans. Importantly, results from this study suggest the presence of newly identified variants in the APOL region as well as variants in other putative nephropathy genes may contribute to diabetic, non-diabetic, and/or all-cause nephropathy in African Americans and/or European Americans. It is therefore apparent that studies of coding variants compliment and extend common variant analyses, contributing insight for some loci, though odds ratios were not striking and recessive models, the likely model for rare variants, are difficult to test.

2311T

Assessing the role of low-frequency variants in type 2 diabetes risk using exome sequencing of 2,000 Danish individuals. *K.E. Lohmueller[1], T. Sparso[2], Q.B. Li[3], E. Andersson[2], N. Grarup[2], A. Albrechtsen[4], T. Korneliussen[5], K. Kristiansen[5], T. Hansen[2,6], G. Sanchez[1], I. Hallgrímsdóttir[1], J. Wang[3], R. Nielsen[1,5], O. Pedersen[2,7,8,9].* 1) Integrative Biology, University of California, Berkeley, CA; 2) The Novo Nordisk Foundation Center for Basic Metabolic Research, University of Copenhagen, Copenhagen, Denmark; 3) BGI-Shenzhen, Shenzhen, China; 4) Centre of Bioinformatics, University of Copenhagen, Copenhagen, Denmark; 5) Department of Biology, University of Copenhagen, Copenhagen, Denmark; 6) Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark; 7) Faculty of Health Sciences, Aarhus University, Aarhus, Denmark; 8) Hagedorn Research Institute, Gentofte, Denmark; 9) Institute of Biomedical Science, University of Copenhagen, Copenhagen, Denmark.

In an effort to better understand the role that low-frequency variants play in the genetic basis of type 2 diabetes, obesity, and related metabolic phenotypes, we sequenced the exomes of 2,000 Danish individuals. Half the individuals (the cases) suffered from type 2 diabetes, moderate obesity (BMI > 27.5 kg/m^2), and hypertension. The others (the controls) were healthy individuals selected from two Danish population cohorts. Data were generated using Agilent All Exon v.2 capture kits followed by Illumina sequencing. Each individual exome had an average depth of 56x. The high coverage per exome allowed accurate identification of coding region single nucleotide variants (SNVs). After applying stringent quality control filters, we identified >90,000 SNVs with minor allele frequency >0.1% within the exons. Of these, 13% were not present in dbSNP build 134. We applied several gene-based tests to detect genes showing different patterns of genetic variation in cases and controls. However, the results of these tests suggest that there are, at most, a limited number of genes containing low-frequency variants (minor allele frequency between 0.1% and 5%) having a strong effect on diabetes risk. Interestingly, the genes showing the largest effects using the sequence kernel association test (SKAT) also contained SNVs that have nominally significant differences in allele frequency between cases and controls (P<0.05), indicating that, at least in the present study, some of the effects captured by gene-based burden tests may also be observed in marginal (single marker) associations. Overall, our results point to two complementary conclusions for understanding the genetics of type 2 diabetes and for implementing exome sequencing studies for complex traits. First, current exome sequencing studies based on a few thousand individuals may be underpowered to detect associations of moderate effect. Such effect sizes may be typical, even when considering low-frequency variants. Second, genetic heterogeneity may be a considerable impediment of exome sequencing studies if causal variants are located in different genes across different individuals. We will discuss strategies to overcome these challenges.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2312F

Whole exome sequencing identifies a novel mutation within the FOXP3 gene in 2-generational family with inflammatory bowel disease (IBD).

D.T. Okou[1], K. Mondal[2], L.J. Kobrynski[3], J.G. Mulle[2], D. Ramachandran[2], V.C. Patel[2], P. Bose[2], J.P. Waters[1], P. Sampath[4], D.J. Cutler[2], M.E. Zwick[2], S. Kugathasan[1, 2, 5]. 1) Dept Pediatric Gastroenterology, Emory Univ School of Medicine, Atlanta, GA; 2) Dept Human genetics, Emory Univ School of Medicine, Atlanta, GA; 3) Dept Pediatric Allergy/Immunology, Emory Univ School of Medicine, Atlanta, GA; 4) Dept Rheumathology, Emory Univ School of Medicine, Atlanta, GA; 5) Children's Health Care of Atlanta, Atlanta, GA.

Early onset Inflammatory Bowel Disease (IBD) has many distinct clinical phenotypes including Crohn's disease and ulcerative colitis. Primary immunodeficiencies can result in chronic bowel inflammations that can mimic IBD with similar clinical, endoscopic and pathological features. We describe an early onset IBD presentation in a 7-year old boy with growth failure, chronic active colitis and severe eczematous dermatitis, who has been resistant to all conventional medications including immunosuppressive and biologics. His two older male siblings have a variety of autoimmune manifestations characterized by diabetes mellitus type 1, psoriasis and severe eczematous dermatitis. His mother has psoriasis and had a colectomy for medically refractory IBD. Because of the variety of immune mediated manifestations within the family, a potential genetic disorder was suspected and the differential diagnosis included X-linked IPEX syndrome. IPEX is a rare disorder caused by mutation in the FOXP3 gene and characterized by the triad of enteropathy, dermatitis and endocrinopathy. Since the presentation was not classical for IPEX, but a strong monogenetic cause was suspected, we performed exome sequencing of the entire family (both parents along with 3 boys). We identified a previously not reported and novel missense variant in exon 6 (c.694A>C, p.C232G) of FOXP3. The affected mother is heterozygous and all male offspring are hemizygous for the same variant. This variation, changes a Cytosine to Guanine at ChrX:49112217(hg19), which is highly conserved among all vertebrates and as well as in placental mammals. The mutation does not affect the FOXP3 protein expression but could interfere with the suppressive function of CD4+CD25+ regulatory T cells (Tregs). No other deleterious variants were identified that could explain the IBD and various other autoimmune disorders in this family. FOXP3 encodes a DNA-binding protein of the forkhead/winged-helix family and is the central controller of the Tregs development. To date, 40 different FOXP3 mutations have been described, and we add a new mutation to the literature. Conclusions: We report a novel mutation within FOXP3 which is associated with a milder form of IPEX syndrome. This may have important implications for IPEX biology as well as IPEX-directed therapeutic development. Our study represents a successful elucidation of a Mendelian disorder using whole exome sequencing.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2313Wh

Identification of rare genetic variants that contribute to diabetic nephropathy in type 2 diabetes through family-based targeted exome sequencing. *M.G. Pezzolesi[1, 2], J. Jeong[1], A.M. Smiles[1], J. Skupien[1,2], J.C. Mychaleckyj[3], S.S. Rich[3], J.H. Warram[1], A.S. Krolewski[1,2].* 1) Section on Genetics and Epidemiology, Joslin Diabetes Center, Boston, MA; 2) Department of Medicine, Harvard Medical School, Boston, MA; 3) Center for Public Health Genomics, University of Virginia School of Medicine, Charlottesville, VA.

Genome-wide linkage analysis in extended pedigrees from the Joslin T2DN Family Collection previously identified several chromosomal loci that contribute to variation in urinary albumin excretion and variation in renal function in patients with type 2 diabetes (T2D). To identify susceptibility genes that contribute to these diabetic nephropathy (DN) sub-phenotypes, we recently initiated a family-based resequencing project to identify rare variants that underlie these linkage signals. One-hundred twenty-six DN cases from 42 families with an excess of renal disease (3.0 cases per family) were selected for resequencing of the coding region of all protein-coding genes at four genomic regions with evidence for linkage of urinary albumin excretion levels (chr. 5q, 7q, 21p, and 22q) and two genomic regions linked to variation in renal function (chr. 2q and 7p). Target enrichment of 3,966 exons from 361 genes across these six regions was performed using a custom target enrichment library followed by next-generation sequencing. The resulting data achieved greater than 70X on-target read depth and exceeded 20–fold sequence coverage of greater than 80% across the target regions. Multi-sample variant calling was implemented using GATK. To date we have completed sequencing and analysis of 63 DN cases from 21 of the families selected for sequencing. In this dataset, we identified 2,733 non-reference variants, including 192 novel variants not annotated in dbSNP or present in data from the 1000 Genomes Project. Non-synonymous and/or frameshift variants were identified in 247 genes across the six linkage regions. To prioritize the variants detected across these loci, we restricted our analysis to novel functional variants segregating in the sequenced samples. In doing so, novel candidate DN genes emerged in two distinct loci. Preliminary data from our study highlights the utility of combining evidence from family-based linkage analysis with targeted exome sequencing to uncover rare genetic variants that contribute to DN. Similar analyses in all 126 DN cases are currently underway and will be completed soon; we anticipate that this approach will allow us to discover novel genes that contribute to variation in urinary albumin excretion and renal function in T2D and rare variants that contribute to the linkage signals identified in the Joslin T2DN Family Collection.

2314T

Exome sequencing and functional data identify PRDM1 and NPD52 as susceptibility factors in Crohn's disease. *A. Franke[1], D. Ellinghaus[1], H. Zhang[2], S. Schreiber[1], M. Parkes[2].* 1) Inst Clinical Molec Biol, Kiel, Germany; 2) Addenbrooke's Hospital, University of Cambridge, Gastroenterology Research Unit, Cambridge, UK.

For most of the 71 Crohn's disease genetic susceptibility loci identified to date the causal variants are unknown. Furthermore,more than 70% of the heritability of Crohn's disease is yet to be accounted for. Here, we interrogated 42 exomes from unrelated Crohn's disease cases together with genome-wide association (GWA) data, to directly identify disease-associated causative genetic variants. Rare missense mutations were identified in PRDM1. These increase T cell proliferation and cytokine secretion upon activation and also increase expression levels of adhesion molecules responsible for T cell migration into intestinal tissues. The risk allele of a previously identified common PRDM1 SNP correlated with reduced PRDM1 expression in both ileal biopsies and peripheral blood mononuclear cells. All functional data for PRDM1 are in accordance with previous murine studies, where Prdm1 deletion led to enhanced T cell proliferation and activation and eventually to a Crohn's disease-like phenotype. We identify a new association between Crohn's disease and a common missense variant Val248Ala in NDP52. This impairs NDP52 regulatory function to inhibit proinflammatory NF-κB signalling and affects protein stability in the context of toll-like receptor pathways. NDP52 is an adaptor for selective autophagy of both intracellular bacteria and specific signalling molecules. Our results thus implicate NDP52 as a novel molecular link between selective autophagy and regulation of proinflammatory signalling in Crohn's disease.

2315F

Exome sequencing for Crohn's disease pedigrees. *B.-S. Petersen[1], Y. Zeissig[2], M. Kohl[2], S. Schreiber[3], S. Zeissig[3], A. Franke[1].* 1) Institute of Clinical Molecular Biology, Christian-Albrechts-University Kiel, Germany; 2) Department of General Pediatrics, University Hospital Schleswig Holstein, Kiel, Germany; 3) Department of Internal Medicine, University Hospital Schleswig Holstein, Kiel, Germany.

Crohn's disease (CD) is a complex, chronic inflammatory bowel disease. A variety of genetic and environmental factors likely play a role in causing CD. Genome-wide association studies and meta-analyses have so far identified 71 susceptibility loci for CD. These, however, explain only 23% of the heritability so far. To identify further sources of heritability, other approaches apart from SNP-based case-control association studies are needed, for example the identification of rare susceptibility variants with high penetrance by systematic resequencing studies. We have therefore carried out exome sequencing for several pedigrees with Crohn's disease for finding rare high penetrance variants and novel candidate genes involved in disease etiology. The project is aimed at the detection of novel mutations shared by several affected individuals of a pedigree as well as *de novo* mutations in sporadically occuring cases. In one trio, we identified a single *de novo* mutation as the cause of early-onset IBD in the male child: a novel hemizygous nonsense SNP in the X-linked inhibitor of apoptosis gene (*XIAP*). Mutations in *XIAP* are known to cause X-linked lymphoproliferative disease type 2 (XLP2) and have once previously been reported to result in early-onset IBD in a male patient. Our preliminary results further indicate that*de novo* mutations indeed seem to play an important role in sporadic early-onset cases and have revealed novel rare mutations shared by the affected individuals of a pedigree which may have an important impact on the development of the disease.

2316W

Whole-exome sequencing of patients with isolated biliary atresia. *E.A. Tsai[1,2], S. Sukhadia[2], C. Grochowski[2], C.-F. Lin[3], O. Valladares[3], L.-S. Wang[3], L. Leonard[2], G.D. Schellenberg[3], K.M. Loomes[4,5], B.A. Haber[6], M. Devoto[4,7], N.B. Spinner[2,3].* 1) Genomics and Computational Biology Graduate Group, University of Pennsylvania, Perelman School of Medicine, Philadelphia, PA; 2) Department of Pathology and Laboratory Medicine, The Children's Hospital of Philadelphia, Philadelphia, PA; 3) Department of Pathology and Laboratory Medicine, University of Pennsylvania, Perelman School of Medicine, Philadelphia PA; 4) Department of Pediatrics, University of Pennsylvania, Perelman School of Medicine, Philadelphia, PA; 5) Division of Gastroenterology, Hepatology and Nutrition, The Children's Hospital of Philadelphia, Philadelphia, PA; 6) Hepatology, Infectious Diseases Clinical Research Department, Merck Research Laboratories, North Wales, PA; 7) Division of Human Genetics, The Children's Hospital of Philadelphia, Philadelphia, PA.

Biliary atresia (BA) is a severe pediatric liver disease resulting in necroinflammatory obliteration of the extrahepatic biliary tree and presenting within the first few months of life. BA is the most frequent indication for pediatric liver transplantation in the United States. The etiology of BA is unknown, with evidence for infectious, environmental, and genetic factors described. Possible genetic factors include a BA associated locus on 10q24.2 revealed by GWAS, report of *JAG1* missense mutations in a small percentage of patients, and a BA associated deletion within 2q37.3 revealed by chromosomal microarray analysis. Animal models have suggested potential genes, but mutations have not been found in humans. We performed exome-sequencing on 30 Caucasian, non-syndromic BA patients to further investigate the genetic basis of BA. The exomes were captured using the Agilent All Exon V5 71MB capture technology and sequenced on the Illumina HiSeq platform. The 100bp paired end reads were aligned to GrCH37 using BWA and further processed using Picard and SAMtools. The variants were called using GenomeAnalysis ToolKit. There are 4,938 loss of function variants found in these samples including 1,098 nonsense mutations. Preliminary analysis has revealed several variants in candidate regions which are being validated. One patient had a heterozygous, missense mutation in exon 6 of *ADD3*, a gene that maps within 10q24.2 upstream of the BA-associated SNP rs17095355. There were no pathogenic mutations found in either *JAG1* or *GPC1*, the candidate gene in 2q37.3. The polymorphisms in these genes were not enriched in our patient population compared to the 1000 Genomes phase 1 allele frequency in Europeans. Further evaluation of candidate gene and novel gene variants will help delineate the complex genetic architecture of this multifactorial disease.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2317T

Identification of low-frequency and common variants associated with circulating levels of vitamin $_{B12}$ and folate. *N. Grarup[1], C.H. Sandholt[1], P. Sulem[2], T.S. Ahluwalia[1], T. Sparsø[1], A. Albrechtsen[3], L. Husemoen[4], D.R. Witte[5], Y. Li[6], A. Linneberg[4], R. Nielsen[3,7,8], T. Jørgensen[4,9,10], J. Wang[1,6,11], T. Hansen[1,12], U. Thorsteinsdottir[2,13], K. Stefánsson[2,13], O. Pedersen[1,14,15,16].* 1) Novo Nordisk Foundation Centre Basic Metabolic Research, University of Copenhagen, Copenhagen, Denmark; 2) deCODE Genetics, Reykjavik, Iceland; 3) Centre of Bioinformatics, Faculty of Science, University of Copenhagen, Copenhagen, Denmark; 4) Research Centre for Prevention and Health, Glostrup University Hospital, Glostrup, Denmark; 5) Steno Diabetes Center, Gentofte, Denmark; 6) BGI-Shenzhen, Shenzhen, China; 7) Department of Integrative Biology, University of California, Berkeley, CA, US; 8) Department of Statistics, University of California, Berkeley, CA, US; 9) Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark; 10) Faculty of Medicine, University of Aalborg, Aalborg, Denmark; 11) Department of Biology, Faculty of Science, University of Copenhagen, Copenhagen, Denmark; 12) Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark; 13) University of Iceland Faculty of Medicine, Reykjavik, Iceland; 14) Faculty of Health Sciences, Aarhus University, Aarhus, Denmark; 15) Hagedorn Research Institute, Gentofte, Denmark; 16) Institute of Biomedical Science, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark.

Background: Vitamin B_{12} (B_{12}) and folate are water-soluble micronutrients. Blood levels are influenced by processes which are dependent on numerous enzymes, cofactors and transporters. B_{12} and folate both play a role in one-carbon metabolism and deficiency results in common clinical features including a wide range of neurological and haematological manifestations. We aimed at elucidating the genetic background determining circulating levels of B_{12} and folate by combining Icelandic and Danish data. *Methods*:In the Icelandic study, 1176 individuals were genome sequenced to a depth of 10× and 22.9 million single-nucleotide variants (SNVs) were identified and imputed into chip-genotyped Icelanders using long-range phasing and the Icelandic genealogical database. The effective Icelandic sample size was 23,493 and 20,542 individuals with measurements of serum B_{12} and serum folate, respectively. In the Danish study, 16,197 coding SNVs, found by exome sequencing to a depth of 8× in 2,000 Danish individuals, were genotyped in ~8,500 individuals. Association analyses were done in the two studies and for overlapping SNVs meta-analyses were performed in total sample sizes of 31,786 and 28,970 for B_{12} and folate, respectively. *Results*: A total of 12 loci associated with circulating serum B_{12} levels at $P < 5 \times 10^{-8}$ were found and six of these were novel loci. The most statistically significant novel locus was for a non-synonymous variant with an allele frequency of 3.1% on chromosome 19p13 ($P = 8.4 \times 10^{-59}$). Other novel loci associated with circulating B_{12} levels at conventional genome-wide significance levels were located at chromosome 2q34, 4q30, 14q24, 11q13 and 22q12. For six loci (chromosome 6p12, 10p13, 11q12, 13q33, 19p13 and 19q13) previously reported to associate with vitamin B_{12} levels we confirmed the association. The combined effect of the associated loci explained 11.5% of the total variation in B_{12} in a sample of the general Danish population. Three novel loci associated with circulating levels of folate located on chromosome 1p36, 11p13 and 11p14 were found in analyses in up to 28,970 individuals. *Conclusion*: Association mapping of circulating B_{12} and folate levels for up to 22.9 million SNVs in up to 31,786 individuals revealed six novel loci for B_{12} and three novel loci for folate levels. Interestingly, eleven of the 12 lead SNVs associated with serum B_{12} are non-synonymous variants in genes coding for proteins known to be involved in B_{12} metabolism.

2318F

Whole-exome sequencing in age-related macular degeneration (AMD) using a highly-discordant phenotype and genotype design: evidence for novel loci underlying bilateral late AMD. *W.K. Scott[1], W.H. Cade[1], M.D. Courtenay[1], P.L. Whitehead[1], I. Konidari[1], W.F. Hulme[1], S.G. Schwartz[2], J.L. Kovach[2], G. Wang[1], A. Agarwal[3], J.L. Haines[4], M.A. Pericak-Vance[1].* 1) Dr. John T. Macdonald Foundation Department of Human Genetics and John P. Hussman Institute for Human Genomics, University of Miami, Miami, FL; 2) Bascom Palmer Eye Institute, University of Miami, Naples, FL; 3) Vanderbilt Eye Institute, Vanderbilt University, Nashville, TN; 4) Department of Molecular Physiology and Biophysics and Center for Human Genetics Research, Vanderbilt University, Nashville, TN.

Age-related macular degeneration (AMD) is the most common cause of irreversible vision loss in older adults in developed countries. Large-scale genome-wide linkage and association studies have identified nineteen genes with variants significantly associated with developing AMD. However, these genes do not explain all of the estimated genetic risk for the disease, as a portion AMD cases carry few of these risk factors, and some unaffected individuals carry many of them. It is believed that this “missing heritability” may comprise gene-gene and gene-environment interactions and rare variant effects not detected by association studies. To identify rare variants that influence the risk of AMD, we first calculated the genetic burden in cases and controls due to the 10 confirmed AMD loci. We then selected for whole exome sequencing (WES) the 20 bilaterally affected choroidal neovascularization cases with the lowest genetic burden and 20 controls with no drusen and the highest genetic burden. Exome capture was performed with the Agilent Sure Select All Exon V4 + UTRs Kit and sequenced on the Illumina HiSeq2000. Sequences were aligned with BWA and variants called with GATK and VQSLOD recalibration. Variants called with adequate depth (6 reads), and quality scores indicating high probability of a true variant call (VQSLOD > -3, PL > 99 for wild type homozygote) were retained for analysis. Gene-based tests of association were conducted in the RVASSOC framework, using a joint test of association by summing individual chi-square tests for all variants in a gene and assessing significance with a permutation test. Preliminary analysis did not detect a gene associated with AMD surpassing a Bonferroni-corrected threshold ($p < 2 \times 10^{-6}$). However, 34 genes met suggestive criteria for association ($p < 0.005$), including two (C3orf17 and PVRL2) with $p < 0.001$. These results indicate that novel biological pathways distinct from complement activation and angiogenesis may contribute to risk of development of bilateral CNV in cases with low burden of known AMD loci.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2319W

Deep resequencing of GWAS loci identifies a novel rare variant in RNF186 associated with Ulcerative colitis. *P. Goyette[1], M. Beaudoin[1], G. Boucher[1], C. Stevens[4], K.S. Sin Lo[1], A. Alikashani[1], M. Ladouceur[1], M.A. Rivas[3], S.R. Brant[5], R.H. Duerr[5], D.P. McGovern[8], M.S. Silverberg[9], M. D'Amato[7], A. Franke[10], S. Vermeire[11], G. Lettre[1,2], J.H. Cho[6], R.J. Xavier[4], M.J. Daly[4], J.D. Rioux[1,2], NIDDK and International IBD Genetics Consortia.* 1) Montreal Heart Inst, Montreal, Canada; 2) Universite de Montreal; 3) Analytic and Translational Genetics Unit (ATGU), Massachusetts General Hospital, Boston, MA, USA; 4) Broad Institute of Harvard and MIT, Cambridge, MA, USA; 5) Division of Gastroenterology, Hepatology and Nutrition, Department of Medicine, School of Medicine, and Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, Pennsylvania, USA; 6) Keck Biotechnology Resource Laboratory and Yale School of Medicine, New Haven, Connecticut, USA; 7) 18Karolinska Institutet, Department of Biosciences and Nutrition, Stockholm, Sweden; 8) 10The Medical Genetics Institute, Cedars-Sinai Medical Center, Los Angeles, California, USA; 9) 13Mount Sinai Hospital Inflammatory Bowel Disease Centre, University of Toronto, Toronto, Ontario, Canada; 10) 7Institute of Clinical Molecular Biology, Schittenhelmstr. 12, D–24105 Kiel, Germany; 11) Division of Gastroenterology, University Hospital Gasthuisberg, Leuven, Belgium.

Genome-wide association studies and follow-up meta-analyses in UC have identified 47 common disease-associated variants that account for only 16% of the expected genetic heritability. It has been suggested that some of the missing heritability may arise from unidentified high effect rare risk variants in known susceptibility regions. To test this, we have used a pooled high-throughput targeted sequencing approach in 200 UC cases and 150 healthy controls (HC), all of French Canadian descent, to study 57 genes in regions associated with UC from our recent meta-analysis of 3 independent genome-wide scans. We performed follow-up genotyping of 42 rare non-synonymous variants in independent case-control cohorts (totaling 12,153 UC cases and 17,575 HC) from the International IBD Genetics Consortium immunochip dataset, and identified significant association ($P<0.0012$) to previously reported rare variants in IL23R and CARD9, as well as novel independent associations with a coding variant in RNF186 ($P=9.11E–04$). Given the challenge inherent in achieving corrected significance thresholds for rare variants, even with large cohorts, we expect that some of the other variants that we identified and found to have nominal significance ($P<0.05$) are truly associated with UC. In fact with our target set of 42 variants included in follow-up genotyping, we would expect <1 SNP to exceed $P < 0.01$ and ~2 SNPs to exceed $P < 0.05$ by chance alone, whereas we observe 3 SNPs with $P < 0.01$ and 9 SNPs with $P < 0.05$, suggesting that there are additional true positives that have not met the more stringent threshold.

2320T

Traits of human disease significance segregating in the non-human primate vervet. *W. Warren[1], A.J. Jasinska[2], N. Tran[2], Y. Huang[2], K. Dewar[3], J. Kaplan[4], P. Minx[1], L. Hillier[1], R.K. Wilson[1], M. Nordborg[5], G.M. Weinstock[1], N.B. Freimer[2].* 1) The Genome Institute, Washington University School of Medicine, St Louis, MO; 2) Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles, CA; 3) McGill University, Montreal, Canada; 4) Wake Forest University, Wake Forest, NC; 5) Gregor Mendel Institute, Vienna, Austria.

Non-human primates (NHP) are important models for genetic investigations of traits relevant to human pathologies, due to their high conservation of molecular processes and physiology with humans. The vervet monkey (Chlorocebus aethiops sabaeus) is a NHP model with an exceptionally well characterized genetic background: a small number of West African vervets were brought to the Caribbean ~ 80 generations ago, and 57 Caribbean vervets were used to found the Vervet Research Colony (VRC), now a nine-generation deep extended pedigree of ~475 individuals. To investigate the genetic causes of numerous quantitative disease-related phenotypes, some not readily studied in humans, we generated the following tools: 1) Vervet genomic assembly, 2) Whole genome sequencing (WGS) in VRC individuals with deep multi-system phenotypic data, and 3) Population genetic analyses of African, Caribbean and VRC monkeys. We have produced a reference genome (2.8 Gb) from a VRC monkey with an N50 contig and scaffold length of 21kb and 2.87Mb, respectively. Analysis of the reference assembly is focused on comparative studies in relation to other primate genomes. The WGS of the VRC has included 723 past and current animals and is being used to map phenotypes collected on a whole pedigree scale including brain-related expression quantitative traits, brain structural measures, concentrations of neurochemical metabolites in cerebrospinal fluid, and systematic measures of stress responsivity. To elucidate the genetic history of Caribbean vervet populations and to evaluate the feasibility of large-scale association analysis among independent individuals, we performed WGS in 130 wild vervets from three Caribbean islands and five major vervet subspecies drawn from across Sub-Saharan Africa. At least one individual from each geographic group was sequenced at 10x coverage while other individuals representing diverse populations within each subspecies were sequenced at 4x coverage. The population genetic data have already yielded variation data relevant to our understanding of susceptibility to infection with simian immunodeficiency virus (SIV), of which the vervet is the most abundant natural host.

2321F

Association and replication of rare allele windows within cis-regulatory regions to nearby gene expression. *A. Fish, L. Wiley, W. Bush.* Center for Human Genetics Research, Vanderbilt University, Nashville, TN.

Expression Quantitative Trait Loci (eQTLs) have been studied extensively in multiple populations and across multiple tissues, however the vast majority of these analyses have focused on the effect of common alleles. It is unknown how much impact rare alleles have on the regulation of expression for nearby genes. Using 149 individuals from the 1000 Genomes Project and existing gene expression data, we defined cis-regulatory regions surrounding 11,465 genes to perform a rare variant eQTL analysis. Rare alleles (MAF < 0.05) were collapsed into 5KB windows for a sliding window analysis over the cis-regulatory region. Windows were evaluated using a linear regression model. Gene expression values were normalized to adjust for population-level gene expression differences. From the linear regression, we used a Student's t-test to compare population-adjusted mean expression differences between individuals with one or more rare alleles versus individuals carrying no rare alleles. From this screening analysis, we detected 315 windows reaching genome-wide significance, affecting the expression of 65 genes. Using an independent set of 163 additional multiethnic individuals from the 1000 Genomes Project coupled with gene expression data, we replicated 146 significant windows at $p < 0.05$. Of these significant windows, 126 contain known common eQTL signals and represent potential synthetic associations based on surrounding linkage disequilibrium patterns. Significant windows are enriched for lymphoblastoid chromatin immunoprecipitation (ChIP) peaks from the ENCODE project (53/315). 353 Rare alleles within these significant windows are shown to alter transcription factor binding from the HaploREG database. Additionally, 53 significant windows contain a miRNA binding site predicted by either miRanda or TargetScan. The remaining 31 windows with no notable annotations may represent novel regulatory elements. These results illustrate that cis-regulatory regions can harbor rare alleles altering gene expression, and that collections of these alleles can be detected using collapsing methods. Finally, the reported windows represent regions where rare alleles have known functional impact and thus may be excellent candidates for rare-variant analysis of disease phenotypes.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2322W

Asthma: An omics view through discordant monozygotic twins. *R. Chen[1], S. Runyon[2], J. Li-Pook-Than[1], K. Nadeau[2], M. Snyder[1]*. 1) Genetics, Stanford University, Stanford, CA; 2) Pediatrics, Stanford University, Stanford, CA.

Asthma, a complex disease that affects 11% of the population in the United States for all ages, or 13% in children under 18 years of age, has a clear hereditary component. However, many of the genetic findings to date are inconsistent, mainly due to the small sample size of available case-control studies, as well as phenotypic heterogeneity of the disease. The low mutation rate in monozygotic twins (estimated 4.6 x 10^{-10} per nucleotide per generation) makes them an extremely valuable population to study complex genetic disorders such as asthma. In our study we determined the whole genomes of 12 pairs of monozygotic twins discordant in allergic asthma (as well as 1 pair discordant in allergy and 2 concordant healthy pairs), and sequenced the whole transcriptome of peripheral blood mononuclear cells (PBMCs) from 20 (including the 12 pairs for whole genome sequencing) discordant monozygotic twins pairs. We observed a spectrum of genomic differences within each pair of monozygotic twin, including single nucleotide variants, insertions/deletions, copy number variants, structural variants, as well as different mobile element insertions. Whole transcriptome sequencing also revealed gene expression differences as well as allele-specific differences within each twin pair and between the asthmatic and normal twin groups. Our study provided a unique perspective to investigate genetic factors associated with this complex, common disease at the omics level.

2323T

Whole-genome deep-coverage sequencing of asthma genomes. *M.A.R. Ferreira, C.S. Tang, G. Chenevix-Trench.* Genetics and Computational Biology, Queensland Institute of Medical Research, Brisbane, Australia.

BACKGROUND. High-impact (nonsense, splice or missense) variants have not been systematically studied for association with asthma risk.

AIM. To screen all high-impact variants detectable in a small number of carefully selected asthmatic individuals (n=13) and prioritize for follow-up the set of uncommon variants most likely to be related to asthma.

METHODS. We sequenced with >30X coverage the genomes of 13 doctor-diagnosed asthmatics, using Illumina GAII or HiSeq. We also sequenced both parents (discordant for asthma) of one asthmatic and an additional 18 unrelated controls. Asthmatics were selected based on seven criteria: (a) never smoker; (b) positive skin prick test; (c) positive bronchial hyperresponsiveness test; (d) childhood-onset asthma; (e) recent asthma treatment; (f) positive family history; and (g) a high ECRHS symptoms score. Sequence reads were analysed using the GATK pipeline, with variants called across the 33 samples jointly and annotated with ANNOVAR.

RESULTS. 8,769,721 SNPs passed QC, of which 92% were known (dbSNP 135) and 8% were novel; the transition to transversion ratio was 2.18 and 2.06, respectively. We identified 31,397 high-impact variants in 11,211 genes. To prioritize variants for follow-up, we selected variants: (a) called in the affected parent and the affected offspring, but not the unaffected mother of the trio (2,567 variants, 1,855 genes); and (b) with a MAF<5% in the NHLBI GO Exome Sequencing and the 1000 Genomes Projects (328 variants, 310 genes); and (c) more common in the 13 asthmatics than in the 18 controls (246 variants, 238 genes); and, finally, (d) located in genes with more high-impact variants in asthmatics than controls (96 variants, 89 genes). One high-impact variant for each of the 89 genes is currently being genotyped in 500 cases and 500 controls. The most promising 30 variants will then be genotyped in a further 1,000 cases and 1,000 controls.

CONCLUSIONS. By sequencing 33 individuals, we identified 89 genes with high-impact uncommon variants that are plausibly related to asthma. These are being genotyped in a larger cohort.

2324F

Rare non-coding variant prioritization and burden testing in whole genome sequencing of complex traits. *K.J. Gaulton, M.A. Rivas, M. van de Bunt, C.M. Lindgren, M.I. McCarthy, the GoT2D consortium.* Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, Oxford, United Kingdom.

Whole genome sequencing studies of complex traits identify millions of non coding variants, however these studies will often be under-powered to detect association in single rare variants of modest effect and require prioritization for follow up or burden testing. A wealth of annotation data exists predictive of functional activity yet optimal strategies of how to use this information to enrich for trait-influencing variants remain unclear. Here we present one of the first systematic approaches to identify rare non-coding variants influencing complex traits, including accumulation and annotation of sequence variants with experimentally and computationally-derived data, identification of annotations relevant to trait biology, separating functional from benign variant effects, and defining units for burden testing. We used this approach to predict and test variants altering transcriptional regulatory elements for contribution to type-2 diabetes (T2D) using whole genome sequence data from the GoT2D project, which is employing a series of genetic techniques to identify variants influencing T2D. A preliminary freeze of 908 samples identified 11.3 million non coding variants, 60% of which are low frequency or rare (MAF<.05). We derived the relevance of experimental regulatory sites across >40 tissues at known T2D-associated variants and identified significant enrichment in several tissues including pancreatic islets (P<1x10–6), chorion (P=7x10–5) and skeletal muscle (P=3x10–4). Of 590k low frequency variants in these cell-type sites, a further 160k were filtered for functional effect by identifying allelic differences in predicted transcription factor (TF) binding for 531 vertebrate factors in JASPAR. We grouped this set of *in silico* regulatory variants in both gene- and TF-centric burden units and tested groups for T2D association using SKAT. Variants in CREB1 binding sites showed significant association to T2D across 531 tested TF groups (P=1.8x10–5). CREB has roles in adipogenesis, brain function and glucose homeostasis and plausibly may contribute to the pathophysiology of T2D. These results highlight the utility of integrating annotation data to identify trait-associated non-coding units harboring rare variants that are individually only nominally significant. More generally, this framework implemented in software CHAOS can be applied to any functional class to identify, group and test annotated sequence variants for contribution to complex traits.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2325W

Deep Sequencing of Genomic Regions Associated with Cleft Lip Susceptibility. *C. Harris[1], E. Leslie[2], R. Fulton[1], T. Miner[1], E. Appelbaum[1], C. Fronick[1], M. Mead[1], D. Larson[1], M. Adela Mansilla[3], A. Lidral[4], L. Moreno[5], J. Hecht[6], M. Marazita[7], A. Scott[8], T. Beaty[9], J. Murray[2], R. Wilson[1], G. Weinstock[1].* 1) Washington Univ Sch Medicine, The Genome Institute, Saint Louis, MO; 2) Department of Pediatrics, University of Iowa, Iowa City, Iowa; 3) The Genome Institute, Washington University, St Louis, Missouri, USA; 4) Departments of Pediatrics, Epidemiology, and Biological Sciences, University of Iowa, Iowa City, Iowa; 5) Department of Orthodontics-Dows Institute, University of Iowa, Iowa City, Iowa, USA; 6) Department of Pediatrics and Pediatric Research Center, University of Texas Medical School, Houston, Texas; 7) Center for Craniofacial and Dental Genetics, Department of Oral Biology, School of Dental Medicine, Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, Pennsylvania, USA; 8) Institute of Genetic Medicine, Johns Hopkins University, Baltimore, Maryland; 9) Department of Epidemiology, School of Public Health, Johns Hopkins University, Baltimore, Maryland, USA.

Nonsyndromic cleft lip and/or cleft palate (NSCL/P) is a group of common birth defects with complex etiology reflecting the action of multiple genetic and/or environmental factors. Genome wide association studies (GWAS) have been widely used for their unbiased approach to identifying genes controlling complex traits such as NSCL/P. One challenge following GWAS is identifying specific disease-causing etiologic variants at each associated locus tagged by presumably non-causal markers. Here we describe targeted capture and sequencing of thirteen candidate loci from NSCL/P GWAS in 1,800 case-parent triads of European and Asian ancestry. The choice of a case/parent trio design allows use of the TDT to detect the most overtransmitted (and presumably causal) variant(s), identification of rare variants could contribute to missing heritability commonly seen in GWAS studies and the detection of *de novo* variants that could have a high prior probability of being causal. The thirteen regions together make up 6.6 megabases of target space, which we covered at greater than 20X depth across a per-sample average of 83.2% of the targeted regions. To achieve this coverage, the samples were multiplexed with 96 samples per capture of the target region, and two Illumina Hi-Seq lanes were generated per capture-set. Preliminary analysis of 62 trios revealed several potentially interesting missense coding mutations that are homozygous in the children, but not so in any of the parents, and are also previously unknown or reported with very low global minor allele frequency in dbSNP. These occurred in genes such as *ABCA4*, *MAFB*, *NANS*, and *ARHGAP29*. In addition, *de novo* mutation analysis found several potentially regulatory variants outside of coding regions in the affected children. We will present analysis of both inherited and de novo mutations discovered in the target regions for additional families and report their association with the presence of NSCL/P for both specific variants and each targeted gene. Funding Sources: This study was funded in part by NIH R01 DE016148 (MLM), NIH U01 HG–005925 (JCM), NIH R01 DE-014581 (THB) and NIH U54 HG-003079 (RKW).

2326T

Exome sequencing identifies rare variants associated with the common pain disorder fibromyalgia. *H. Hor[1], E. Docampo[1], R. Rabionet[1], M. Gut[2], M. Bayés[2], I. Gut[2], A. Collado[3], S. Ossowski[1], X. Estivill[1].* 1) Genes and Disease Program, Center for Genomic Regulation (CRG) and Pompeu Fabra University, Barcelona, Spain; 2) Centro Nacional de Analisis Genomico, Barcelona, Spain; 3) Department of Rheumatology, Hospital Clínic, Barcelona, Spain.

Fibromyalgia (FM) is a complex and one of the most frequent pain disorders with a prevalence of 2% in the general population, mainly affecting women in their mid-forties. Previous candidate gene approaches, generally performed in a very limited number of samples, identified variants in a small number of genes that were associated with the disease, but these findings lacked replication in independent cohorts. We performed an exome sequencing study on well-characterized sporadic and familial cases of FM of Spanish origin to identify genetic variants that are causally involved or associated with the disease. While aiming at a total sample size of 100 exomes, we already analyzed a subset of 36 exomes of sporadic and familial FM cases by focusing on single nucleotide variants (SNV). Following exome sequencing and SNV calling, we merged the variants of all exomes and performed a prioritization of rare, non-synonymous (and splice-site) SNVs according to their minor allele frequency (MAF) and functional damaging potential. Finally, by focusing only on genes with at least two damaging variants among our samples, we were able to narrow down the list of variants to around 450 candidate genes. To further prioritize these variants, we performed a functional enrichment analysis and we surprisingly found an enrichment of genes related to Medical Subject Headings (MeSH) terms for pain related genes and pathways. Based on this, we selected 26 candidate variants for further validation by genotyping in 1200 FM cases and 1000 control samples of Spanish origin. Two SNVs in candidate genes on chromosome 3 (OR: 1.7) and 19 (OR: 1.5) were significantly associated with FM. We are currently recruiting additional FM samples of European origin to replicate our findings in a second independent cohort of samples. Meanwhile we are also increasing our exome sample size to reach 100 exomes, which will be similarly analyzed to obtain a candidate gene set that can be further analyzed by targeted resequencing in a case-control manner. In conclusion we show that exome sequencing has the power of detecting rare variants for complex disorders and even for a highly heterogeneous disorder such as fibromyalgia. Replication of our findings in other cohorts would convert this study into the identification of the first genetic variants associated with FM.

2327F

Application of whole genome sequencing to identify susceptibility loci for polycystic ovary syndrome. *G. Kosova[1], R.S. Legro[2], A. Dunaif[1], M. Urbanek[1].* 1) Division of Endocrinology, Metabolism and Molecular Medicine, Northwestern Univeristy Feinberg School of Medicine, Chicago, IL, USA; 2) Department of Obstetrics and Gynecology, Pennsylvania State College of Medicine, Hershey, PA, USA.

Polycystic ovary syndrome (PCOS) is a complex polygenic endocrine disorder common among reproductive age women. It is characterized by hyperandrogenemia, chronic anovulation, and menstrual irregularities. In addition, women with PCOS exhibit broad range of metabolic consequences, including increased risk of obesity, type 2 diabetes and insulin resistance. Involvement of inflammatory and TGF-β pathways, as well as ovarian rigidity genes in the etiology of PCOS has been postulated, however to date reproducible associations have been identified with only few variants. In order to survey the complete variation in the genome, 58 women with PCOS, and 12 healthy controls were sequenced for the whole genome at 40x coverage by Complete Genomics, Inc. The first phase of this project focuses on 307 candidate genes, chosen for their roles in ovarian rigidity and follicular structure, inflammation, obesity, type 2 diabetes, and TGF-β and NFκB pathways. We found 116,450 unique variant loci mapping to these genes and 5kb around them; with single nucleotide polymorphisms (SNPs) making up 72.9% of the total observed variation, and coding variants constituting less than 0.5%. Using Fisher's exact test for allele counts at coding SNPs, we identified two missense variants in insulin-like 3 (*INSL3*; min$P$ = 0.0013), and one missense variant in each of cathepsin H (*CTSH*; P=0.014), chordin (*CHRD*; P=0.017) and interleukin 1 receptor associated kinase 1 (*IRAK1*; P= 0.020) genes, whose frequencies were increased in PCOS. Small insertions, deletions and substitutions, approximately one to 100 base pairs in length, will be evaluated in a similar way. These results will help us to prioritize and select variants to be genotyped and tested for associations with PCOS in a larger cohort of ~1000 cases and ~1000 controls. In addition to associations with individual variants, the combined effects of mutations in a gene may pose a risk for developing PCOS. Therefore, we will subsequently test for presence of mutational burden in these candidate genes.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2328W

CYP1B1, MYOC and LTBP2 mutations in primary congenital glaucoma patients in the United States. *S. Lim[1], K. Tran-Viet[1], T. Yanovitch[3], S. Freedman[1], T. Klemm[2], W. Call[1], C. Powell[1], A. Ravichandran[1], R. Metlapally[4], E. Nading[1], S. Rozen[2], T. Young[1].* 1) Duke University, Durham, NC; 2) Duke-National University of Singapore Graduate Medical School, Singapore; 3) University of Oklahoma/Dean McGee Eye Institute, Oklahoma, OK; 4) University of California at Berkeley, Berkeley, CA.

Purpose: To screen primary congenital glaucoma (PCG) patient DNA samples for mutations in known PCG genes CYP1B1, LTBP2 and MYOC. Methods: Whole exome sequencing (WES) was performed on affected members of 3 families, and variants were confirmed by Sanger sequencing. Sanger sequencing was performed on additional 44 PCG families. Results: Twelve patients (21.1 %) carried CYP1B1 mutations: 1 nonsense and 6 compound heterozygous mutations. There were no pathogenic LTBP2 and MYOC mutations. Conclusion: This is the largest U.S. PCG cohort study analyzing genetic mutations using WES and Sanger sequencing. Five compound heterozygous CYP1B1 mutations were novel. The low percentage mutation rate in the 3 genes indicates other genes are implicated with PCG. WES may be a powerful tool to identify novel PCG genes.

2329T

Enhancing Sequencing Lab Workflows in the SardiNIA Project. *A. Maschio[1,2], F. Busonero[2], C. Sidore[1,2,3], E.A. Ketterer[4], E.B. Pedersen[4], A.M. Chidester[4], A.M. Gurkan[4], M.R. Coon[4], M. Dei[2], S. Lai[2], A. Mulas[2], M.G. Piras[2], M. Lobina[2], B. Tarrier[4], S. Sanna[2], C.A. Brennan[4], G.R. Abecasis[1], F. Cucca[2,3], R.H. Lyons[4].* 1) Center for Statistical Genetics, University of Michigan, Ann Arbor, MI, USA; 2) Istituto di Ricerca Genetica e Biomedica (IRGB), CNR, Monserrato, Italy; 3) Dipartimento di Scienze Biomediche, Università di Sassari, SS, Italy; 4) DNA Sequencing Core, University of Michigan, Ann Arbor, MI, USA.

Low-pass whole genome sequencing can detect low frequency and non-SNP variants that are not well represented in commercial genotyping arrays. With this purpose we are sequencing 2107 individual genomes from the SardiNIA project each at a depth of 3–4X. Our goal is to use the data to extend our understanding of the genetic factors contributing to many aging related traits. Our sequencing workflow is implemented in the University of Michigan DNA Sequencing Core (UMDSC), which provides expertise and efficient laboratory infrastructure to perform all processing steps (sample QC, library prep, library QC, sequencing, data extraction, data QC and data transfer). A main cause of genotyping errors is related to human factors, that is why to minimize sample tracking errors each sample is genotyped on the Sequenom platform with a set of 36 SNP markers, chosen to be common in the population and previously genotyped with the Affymetrix 6.0 arrays and Illumina MetaboChip. Next, samples proceed to library preparation. With rapidly improving sequencing technology and instrument throughput, we made several improvements to our procedures for preparing paired-end Illumina libraries. We have evaluated multiple automation protocols for library preparation, including the QIAcube (QIAGEN), epMotion 5075 (Eppendorf), and Apollo 324 (IntegenX) protocols, which can facilitate preparation of libraries. Sequencing data generation, which initially used 3 Illumina Genome Analyzers IIx now relies on 7 Illumina HiSeq2000 instruments. Currently runs typically include 24 multiplexed samples per flow cell, resulting in >10GB of sequence data per sample and an average depth of 3–4X after focusing on high-quality mapped bases per each sample. Our results illustrate protocols for lab automation and generation of high quality short read data in a small laboratory.

2330F

Whole-Exome Sequencing Identifies Novel Risk Variants for Thrombotic Storm. *P. Whitehead[1], N.D. Dueker[1], G. Beecham[1], D. Hedges[1], A. Beecham[1], S. Hahn[1], N. Hofmann[1], J.W. Lawson[2], D. Erkan[3], L.R. Brandao[4], A.H. James[5], M.J. Manco-Johnson[6], R. Kulkarni[7], C.S. Kitchens[2], T.L. Ortel[8], J.M. Vance[1], M.A. Pericak-Vance[1].* 1) Hussman Institute for Human Genomics, University of Miami, Miami, FL; 2) University of Florida, Gainesville, FL; 3) Hospital for Special Surgery, Weill Medical Center of Cornell University; 4) Pediatric Hematology/Oncology, The Hospital for Sick Children, Toronto, ON; 5) Department of OB/GYN, Duke University Medical Center, Durham, NC; 6) University of Colorado Denver and The Children's Hospital, Aurora, CO; 7) Pediatrics & Human Development, Michigan State University, East Lansing, MI; 8) Hematology, Duke University Medical Center, Durham, NC.

Thrombotic storm (TS) is a rare disorder defined by ≥ 2 of the following in a short period: 1) more than two acute arterial/venous thromboemboli, or thrombotic microangiopathy, 2) unusual location, 3) progressive/recent unexplained recurrence, or 4) refractory/atypical response to therapy. We hypothesized that prothrombotic genetic risk factor(s) trigger an accelerated form of thrombosis following an initial event. We performed whole-exome sequencing (WES) on two families with TS, each with one case, unaffected parents, and unaffected sibs. Another 25 TS cases were also sequenced. DNA was captured using the Agilent 50mb kit and sequenced using Illumina HiSeq2000. Alignment and genotype calling were performed with Burrows-Wheeler Aligner and Genome Analysis ToolKit. Variants were filtered by function (nonsense, missense, or splice), conservation, damage potential, sequence quality, and rare frequency. We identified variants that met our filtering criteria and were inherited in a de novo or autosomal recessive manner in the two families. Using this approach we identified 2 de novo (SLC26A2; PRMT7) and 7 autosomal recessive (GRIK1, TUBGCP6, ADAMTS9, DPP6, HLA-E, EGFL8, TNFAIP3) variants. We then looked to see if any variants within the 9 identified genes were found in an additional 25 cases. Six of the genes had functional, conserved, potentially damaging, and rare-moderately rare (MAF ≤ 0.05) variants in at least 1 of the other 25 cases; GRIK1, TUBGCP6, ADAMTS9, DPP6, SLC26A2, and TNFAIP3. As a second approach, we identified genes having functional, conserved, damaging, good quality, and rare variants (MAF≤0.01) in ≥ 2 of the 27 cases. A total of 1770 genes met these criteria. Two gene families (ADAMTS/ADAM and DNAH) showed multiple genes within the family meeting these criteria (n=8 ADAMTS/ADAM; n=7 DNAH) and most of these variants have been validated. Further investigation also identified an additional 17 ADAMTS/ADAM genes (all validated) and 5 DNAH genes having a variant in 1 case. Several genes showed variation in more than 10 cases includingZFHX3 and VPS13B. Using whole-exome sequencing we were able to identify several loci that may play a role in the development of thrombotic storm. We are following up functional studies on several TS candidate loci including EGFL8 which has two EGF domains, a common motif in hemostatic and fibrinolytic proteins, and is, thus, potentially involved in coagulation. Further analysis is warranted.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2331W

Identification of Genetic Modifiers Associated with Risk of Stroke in Children with Sickle Cell Anemia. *J.M. Flanagan[1], H. Linder[1], V. Sheehan[1], T.A Howard[1], B. Aygun[2], J.S Hankins[2], M. Wang[3], G.A. Neale[3], R.E. Ware[1].* 1) Texas Children's Hematology Center, Baylor College of Medicine, Houston, TX; 2) Department of Hematology, St. Jude Children's Research Hospital, Memphis, TN; 3) Hartwell Center for Bioinformatics and Biotechnology, St. Jude Children's Research Hospital, Memphis, TN.

Introduction: Stroke is one of the most catastrophic acute complications of sickle cell anemia (SCA), occurring in 11% of patients before 20 years of age. A further 20 to 30% of children with SCA will develop less clinically overt cerebrovascular disease events such as transient ischemic attacks (TIA) and silent infarcts. There is a definite need for biomarkers that could determine the cause of these irreversible cerebrovascular events and which might predict which children are at greatest risk. Previous studies of sibling pairs have shown that there is a genetic component to cerebrovascular disease development but few genetic modifiers have been validated as having a substantial effect on risk of stroke. The aim of this study was to perform an unbiased whole genome search for genetic modifiers of stroke risk in SCA. Methods: Pediatric patients with SCA and documented primary stroke (n=177) were compared to a pediatric control non-stroke group with SCA (n=340). All control patients were over 5 years old and without previous clinical stroke prior to beginning any clinical treatment. Genome wide association studies (GWAS) were performed using genotype data obtained from Affymetrix SNP6.0 arrays. A pooled DNA approach was used to perform whole exome sequencing (WES) by Illumina next generation sequencing of pooled control (n=104), and pooled stroke (n=120) groups. Results: From the Affymetrix SNP6.0 GWAS data, 139 single nucleotide polymorphisms (SNPs) were identified with stroke association. From the WES, 294 non-synonymous mutations were found to be significantly associated with stroke. In combination, 22 mutations identified by WES were located within 500kb of a SNP identified by GWAS. These 22 mutations represent key areas of the genome that are targets for further in depth study. To validate the association of these mutations with risk of stroke, 20 mutations were genotyped in an independent cohort of control (n=236) and stroke (n=57) patients with SCA. One mutation in GOLGB1 (Y1212C) was corroborated as having significant association with risk of stroke (p=0.02). This mutation in GOLGB1 is predicted to effect the golgi associated function of the encoded protein and future studies will focus on how this functional mutation may lead to development of cerebrovascular disease in the context of SCA.

2332T

Implication of rare variants from European-derived adiposity loci in African Americans. *J. Hester[1,2,3], P. Mudgal[1,2], J. Li[1,2], J. Xu[1,2], P. Hicks[1,2,4], B. Freedman[5], D. Bowden[1,2,4,5], M. Ng[1,2].* 1) Center for Diabetes Research, Wake Forest University School of Medicine, Winston-Salem, NC 27157; 2) Center for Genomic and Personalized Medicine Research, Wake Forest University School of Medicine, Winston-Salem, NC 27157; 3) Program in Molecular Genetics and Genomics, Wake Forest University School of Medicine, Winston-Salem, NC 27157; 4) Department of Biochemistry, Wake Forest University School of Medicine, Winston-Salem, NC 27157; 5) Department of Internal Medicine, Wake Forest University School of Medicine, Winston-Salem, NC 27157.

Genome-wide association studies have successfully identified nearly 100 common variants associated with adiposity in multiple populations. However, common variants with relatively small effect can only account for a small proportion of the total variance of BMI. There is potential utility in exploring the impact of rare variants that may have stronger effects in the respective genes. There have been recent efforts focused on sequencing the protein coding regions of the human genome to identify variants that may be potentially functional. 47 exonic variants, predominantly rare ($0.005 < MAF < 0.05$), from seven previously discovered European adiposity loci that replicated association in African Americans ($P < 0.05$) were selected from the Exome Sequence Project (ESP) database and genotyped in 1561 subjects from two African American cohorts. These SNPs were tested for association with BMI after adjustment for age, gender, diabetes status, and admixture in each cohort. Meta-analysis was conducted to combine the results. Meta-analysis of 1561 subjects revealed 3 variants demonstrating significant association at $0.012 < P < 0.045$ with effect sizes between 0.08 and 0.63 SD unit of BMI for each copy of the BMI-increasing allele. The majority of the associations were observed in the *SEC16B* gene. This gene required for organization of transitional endoplasmic reticulum sites and protein export, however, the biological mechanism is largely unknown. Effect sizes for common variants found within the *SEC16B* gene range from 0.03 to 0.05 SD unit of BMI for each copy of the BMI-increasing allele, while the effect sizes observed in rare variants range from 0.08 to 0.43 SD unit of BMI. Our results suggest that rare variants in obesity-associated genes may provide insight into the genetics of obesity in African Americans. A larger sample size will be necessary to confirm or refute the associations.

2333F

Assessing the effect of rare variants on cancer drug-induced cytotoxicity phenotypes using whole genome sequence data. *A. Tikhomirov[1], A. Pluzhnikov[1], C.R. King[2], A. Konkashbaev[1], E. Gamazon[1], M.E. Dolan[3], D.L. Nicolae[1,4,5], N.J. Cox[1].* 1) Section of Genetic Medicine, Department of Medicine, The University of Chicago, Chicago, IL; 2) Department of Health Studies, The University of Chicago, Chicago, IL; 3) Section of Hematology and Oncology, Department of Medicine, The University of Chicago, Chicago, IL; 4) Department of Statistics, The University of Chicago, Chicago, IL; 5) Department of Human Genetics, The University of Chicago, Chicago, IL.

Our goal is to understand the relationship between common and rare variants' contributions to cytotoxicity outcomes, identify genes likely to have rare variants affecting the phenotypes, and characterize network properties of these genes to develop a prediction model. The analyses are performed on 60 unrelated Caucasian (CEU) and 59 unrelated African (YRI) whole genome autosomal sequences obtained from March 2012 release of 1K Genomes data (hg19/b37) comprising over 38M SNPs and structural variants. The lymphoblastoid cell lines (LCL) were phenotyped for cytotoxicity induced by carboplatin, cisplatin, aracin, daunorubicin, and etoposide, as well as for expression of over 14K genes. We use the gene-based method of King et al. (EMMPAT) to obtain a single test statistic per gene combining information from all variants pertinent to the gene, variant allele count and fitness burden, as well as prior information on effect sizes estimated based on MAF. Approximately 16M intragenic SNPs (both rare and common) from over 22K autosomal genes were included in the first round of the analysis supplemented later by over 100K LCL eQTLs (p value < 2e-8). In addition, we use PLINK for single point quantitative trait association tests with cytotoxicity phenotypes. Testing intragenic SNPs alone shows strong effect of rare variants (MAF < 5%) on top association signals for every phenotype regardless of the drug action mechanism and etiology. For the majority of phenotypes, the genes implicated in EMMPAT analyses are congruent with top SNP-based association signals. We extend the gene-based approach to include eQTLs associated with a gene into the model and present the results of comparison between the effects of common vs. rare variants in analyses combining the effects of intragenic SNPs and *cis*-eQTL, as well as both *cis*- and *trans*-eQTLs. We also plan to present results of analyses on at least five additional cytotoxicity phenotypes that have just become available.

2334W

Detecting rare variant effects using extreme phenotype sampling in sequencing association studies. *I. Barnett, S. Lee, X. Lin.* Biostatistics Department, Harvard University, Boston, MA.

In the increasing number of sequencing studies aimed at identifying rare variants associated with complex traits, the test power can be improved by guided sampling procedures. We show both analytically and numerically that sampling individuals with extreme phenotypes can enrich the presence of causal rare variants and leads to an increase in power compared to random sampling. While application of traditional rare variant association tests to these extreme phenotype samples requires dichotomizing the continuous phenotypes before analysis, the dichotomization procedure can decrease the power by reducing the information in the phenotypes. To avoid this, we propose a novel statistical method based on optimal SKAT (SKAT-O) that allows us to test for rare variant effects using continuous phenotypes in the analysis of extreme phenotype samples. The increase in power of this method is demonstrated through simulation of a wide range of scenarios as well as in the triglyceride data of the Dallas Heart Study.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2335T

Whole Genome Sequencing to Identify Rare Genetic Variation Responsible for Resistance to HIV Infection. *P.R. Shea, K. Pelak, K.V. Shianna, D. Ge, D.B. Goldstein, National Institute of Allergy and Infectious Diseases Center for HIV/AIDS Vaccine Immunology (CHAVI).* Center for Human Genome Variation, Duke University, Durham, NC.

Epidemiologic studies of individuals at a high risk of HIV exposure have found that a small proportion of the human population is resistant to HIV infection. Investigations of these high-risk, seronegative (HRSN) cohorts have identified genetic variants in the CCR5 gene that confer high-level resistance to HIV infection. However, CCR5 variation is only able to explain resistance in a minority of HRSN individuals; therefore the remaining mechanism for HIV resistance is unknown. Interrogation of common genetic variation using genome-wide association has thus far been unable to identify additional genetic variants that have consistently been associated with HIV resistance. Increasingly, we have begun to appreciate that rare genetic variation plays an important role in the etiology of complex diseases. To test the hypothesis that rare genetic variation is responsible for resistance to HIV infection, we performed whole-genome sequencing on more than 225 Caucasian patients with severe hemophilia who were potentially exposed to HIV-contaminated blood products from 1979 to 1984, but remain seronegative. Genome data from these patients was compared to over 200 whole-genome sequenced population controls using single variant association testing. A preliminary gene-based collapsing analysis of rare variants performed with 66 HRSNs and 119 controls identified 12 genes that were enriched for rare genetic variants relative to controls at a statistically significant level (empiric $P < 0.05$). Further investigation of these genes revealed that some, but not all, are artifact caused by genes with high levels of background variation. Additional follow up analyses of individual variants and remaining genes showing enrichment will be needed to determine their role in HIV resistance.

2336F

Rare Variants are Strongly Associated with Age of Menarche: First Results from NHLBI-Exome Sequencing Project. *K.L. Bucasas[1], G.T. Wang[1], P.L. Auer[2], M. Kan[1,3], J.M. Murabito[4,5], N. Franceschini[6], E.W. Demerath[7], K.L. Lunetta[8], A. Rodriguez[1], L.A. Lange[6], R.D. Jackson[9], S.M. Leal[1], NHLBI-Exome Sequencing Project.* 1) Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, 77030; 2) Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle WA, 98109; 3) Institute of Nutritional Sciences, Shanghai Institutes for Biological Sciences, Graduate School of Chinese Academy of Science, Chinese Academy of Sciences, Shanghai, 2000032; 4) Sections of General Internal Medicine, Department of Medicine, Boston University School of Medicine, Boston MA, USA, 02118; 5) NHLBI Framingham Heart Study, Framingham, MA, USA, 02118; 6) Department of Genetics, University of North Carolina, Chapel Hill, NC, 27599; 7) Division of Epidemiology and Community Health, University of Minnesota, Minneapolis, MN 55454; 8) Department of Biostatistics, Boston University School of Public Health, Boston, MA, 02118; 9) Center for Clinical and Translational Science, The Ohio State University, Columbus, OH 43210.

Age of menarche is a highly heritable, common complex trait that is recognized as a significant risk factor for many late onset diseases, including cardiovascular disease and cancer. For most complex traits, the contribution of rare variants is largely unknown since previous surveys of genetic architecture have mainly been limited to evaluating common variation. We took a comprehensive approach in studying both common and rare variants by utilizing next-generation technologies that involve massive parallel sequencing of protein coding sequences across the human genome. We analyzed whole-exome sequence data from 2814 African American and European American women from the NHLBI-ESP (Exome Sequence Project) with age of menarche data. We performed race-stratified, gene-based rare variant association testing (minor allele frequency [MAF]< 0.01) by analyzing nonsynonymous and splice site variants using the Combined Multivariate Collapsing (Li & Leal 2008) method. We also applied the Variable Threshold (Price et al. 2010) test which within a gene region maximizes the test statistic over variant allele frequencies and uses permutation to adjust for multiple testing. Additionally, association tests were performed for all individual variants with a MAF > 0.01. Covariates were included in our regression model to adjust for the effects of population substructure (Multi-Dimensional Scaling Components), cohort membership, generational age effects and body mass index. Our results provide new insights into the genetic etiology of age of menarche. We observed that rare variants within multiple genes including LRRC8B $p=9x10^{-7}$ and several individual variants (e.g. rs62260755 $p=3x10^{-6}$) are associated with age of menarche. Interestingly these findings were limited to European Americans and no significant associations were observed for African Americans, despite similar sample sizes (n= 1332 African Americans and n=1482 European Americans). Differences in the allelic architecture between African Americans and European Americans could partly explain this discrepancy. Finally, this study is among the first to demonstrate evidence for the significant influence exerted by rare variants on a complex trait using a large-scale whole exome sequencing approach.

2337W

Detecting Rare Variant Associations with Waist-to-Hip Ratio (WHR) in NHLBI-ESP Female Cohorts. *M. Kan[1,2], K.L. Bucasas[1], P.L. Auer[3], G. Wang[1], A. Rodriguez[1], J. Ellis[4], L.A. Cupples[5,6], Y.D.I. Chen[7], J. Dupuis[5], C. Fox[6], M.D. Gross[8], N. Heard-Costa[6,9], J.B. Meigs[10], J.S. Pankow[11], J.I. Rotter[7], D. Siscovick[12,13], J.G. Wilson[14], J. Shendure[15], R. Jackson[16], U. Peters[3], E. Whitsel[17], J. Zhong[18], D. Lin[19], L. Hsu[3], N. Franceschini[4], C. Carlson[13], K.E. North[17], L.A. Lange[4], A.P. Reiner[13,20], S.M. Leal[1], NHLBI-Exome Sequencing Project.* 1) Department of Molecular and Human Genetics, Baylor College of Medicine, Houston TX, 77030; 2) Institute of Nutritional Sciences, Shanghai Institutes for Biological Sciences, Graduate School of Chinese Academy of Science, Chinese Academy of Sciences, Shanghai, 200032; 3) Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle WA, 98109; 4) Department of Genetics, University of North Carolina, Chapel Hill, NC, 27599; 5) Department of Biostatistics, Boston University School of Public Health, Boston, MA 02118; 6) National Heart, Lung and Blood Institute's Framingham Heart Study, Framingham, MA 01702; 7) Medical Genetics Institute, Cedars-Sinai Medical Center, Los Angeles, CA 90048; 8) Division of Epidemiology and Community Health, University of Minnesota School of Public Health, Minneapolis, MN, 55455; 9) Department of Neurology, Boston University School of Medicine, Boston, MA 02118; 10) Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114; 11) Division of Epidemiology and Community Health, University of Minnesota, MN 55454; 12) Department of Medicine, University of Washington, Seattle, WA 98101; 13) Department of Epidemiology, University of Washington, Seattle, WA 98195; 14) Department of Physiology and Biophysics, University of Mississippi Medical Center, Jackson, MS, 39216; 15) Department of Genome Sciences, University of Washington, Seattle WA 98195–5065; 16) Center for Clinical and Translational Science, The Ohio State University, Columbus, OH 43210; 17) Department of Epidemiology, University of North Carolina, Chapel Hill, NC 27514; 18) Environmental Medicine, Division of Biostatistics, NYU School of Medicine, New York, NY, 10016; 19) Department of Biostatistics, University of North Carolina, Chapel Hill, NC 27599; 20) Carolina Center for Genome Sciences, University of North Carolina, Chapel Hill, NC 27599.

Waist-to-hip ratio (WHR) is a common anthropometric measure in humans that is used to assess abdominal obesity. Recent genome-wide association studies (GWAS) have identified numerous common genetic loci influencing this trait but the contribution of rare variants is not known. To better understand the underlying genetic architecture of WHR, we performed a large-scale whole-exome sequencing study on 1518 African American (AA) and 2337 European American (EA) participants from seven cohorts in the NHLBI-ESP (Exome Sequence Project) with information on WHR. We stratified our association analyses by race and sex because of the observed differences in WHR distributions between EA and AA (p=1.2x10–8) individuals and males and females (p=2.2x10–16). Due to limited sample sizes for males, we concentrated our analysis on the 1209 AA women and 1551 EA women with WHR data. We performed gene-based rare variant association testing (minor allele frequency [MAF] < 0.01) by analyzing nonsynonymous and splice site variants using the Combined Multivariate Collapsing (Li & Leal 2008) method. We also performed aggregate variant association testing by applying the Variable Threshold (Price et al. 2010) test which within a gene region maximizes the test statistic over variant allele frequencies and uses permutation to account for multiple testing. Additionally, association tests were performed for all individual variants with a MAF $\geq$ 0.01. Covariates were included in our regression models to adjust for the effects of population substructure, cohort membership, age, current smoking status and body mass index. Our strongest association was observed for a burden of rare variants within a gene located on chromosome 8 ($p = 3.1x10–7$) which meets the exome-wide significance level ($p < 2.5 \times 10–6$) within the EA subpopulation. No exome-wide significant associations were detected for AA women, possibly because of a smaller sample size and differences in allelic architecture. This is the first study to demonstrate the significant influence exerted by rare variants within a gene on WHR using large-scale whole exome sequencing.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2338T

Identification of Rare Susceptibility Variants for Bipolar Disorder by Exome Sequencing. *C. Cruceanu[1], A. Ambalavanan[2], J.P. Lopez[1], J. Gauthier[2], D. Spiegelman[2], R.G. Lafrenière[2], P. Grof[3], M. Alda[4], G. Turecki[1], G. Rouleau[2,5].* 1) McGill Group for Suicide Studies, Douglas Mental Health University Institute, McGill University, Montreal, Quebec, Canada; 2) Center of Excellence in Neuroscience of the Université de Montréal (CENUM), Montreal, Quebec, Canada; 3) Mood Disorders Centre of Ottawa, Ottawa, Ontario, Canada; 4) Department of Psychiatry, Dalhousie University, Halifax, Nova Scotia, Canada; 5) Centre de Recherche du Centre Hospitalier de l'Université de Montréal (CRCHUM), and Department of Medicine, Université of Montreal, Montréal, Quebec, Canada.

Background: Bipolar disorder (BD) is a complex psychiatric condition characterized by manic and depressive episodes. In spite of the strong support for the role of genetics in BD, molecular studies have by-and-large had minimal success in identifying disease-causing genes. This is likely because of high levels of heterogeneity in the sample sets used, but the problem may be minimized by focusing on a well-defined sub-phenotype of BD. Our group has led efforts to characterize BD patients that respond positively to Lithium (Li) therapy, and shown that Li-response clusters in families. Research in BD genetics to date has consisted of linkage and genome-wide association studies, which presume that common variants in a small subset of genes are the cause for BD. However, the minimal success of this approach suggests that BD is caused by highly penetrant rare variants in a large number of different genes across the population. Methods: Our approach focused on a well-defined clinical subtype of BD (Li-responsive) to minimize clinical heterogeneity, and we used massively-parallel next-generation sequencing to re-sequence the exomes of all affected individuals from multi-generational family units. To identify relevant BD susceptibility genes we prioritized rare variants that segregate with affected status within each family. To validate candidate variants we used Sanger sequencing in all family members. Results: We found that each family shared a only 5–10 potentially highly penetrant (e.g. protein-truncating or missense) or functionally relevant (e.g. 3′UTR, 5′UTR, splicing) variants. After Sanger sequencing validation we were able to narrow in on strong candidates for BD causality or susceptibility within each family, and to further explore the mechanisms by which these variants could lead to pathology using cell line and post-mortem brain samples. Conclusion: By focusing on rare variants, rather than common variants, we hope to have narrowed in on the key genes and biochemical pathways that play an important role in bipolar disorder and can lead directly to clinically relevant diagnostic and therapeutic applications.

2339F

Deep Sequencing in Extended Pedigrees Reveals both Common and Rare Non-Synonymous Variants Influencing Free Triiodothyronine Levels. *J.E. Curran[1], S. Kumar[1], M. Almeida[1], J.M. Peralta[1], T.D. Dyer[1], T.M. Teslovich[2], G. Jun[2], C. Fuchsberger[2], A.R. Wood[3], T.M. Frayling[3], P. Cingolani[4], T.W. Blackwell[2], R. Sladek[5], G. Atzmon[6], S.P. Fowler[7], V.S. Farook[1], S. Puppala[1], J. Laramie[8], S. Lincoln[8], D.M. Lehman[7], G.R. Abecasis[2], R. Duggirala[1], J. Blangero[1].* 1) Texas Biomedical Research Institute, San Antonio, TX, USA; 2) University of Michigan, Ann Arbor, MI, USA; 3) University of Exeter, Exeter, United Kingdom; 4) McGill University, Montreal, Canada; 5) Montreal Diabetes Research Center, Montreal, Canada; 6) Albert Einstein College of Medicine, Bronx, NY, USA; 7) University of Texas Health Science Center at San Antonio, San Antonio, TX, USA; 8) Complete Genomics, Inc., Mountain View, CA, USA.

Thyroid hormones are of paramount importance to human physiology. They are intricately involved in the regulation of many metabolic and physiologic processes, from early development throughout adulthood, including skeletal and neural development, tissue growth and maturation, regulation of metabolism, cardiac function and body temperature, brain maturation, and many more. There are two main circulating thyroid hormones, T3 (triiodothyronine) and T4 (thyroxine), however only the unbound free forms of these hormones are metabolically active, with T3 being the most abundant and also biologically active form. It is well established that there is a correlation between quantitative variation in normal thyroid hormone levels and several disease-related risk factors; however the genetic factors influencing this normal physiological variation remain largely unknown. As part of the San Antonio Family Study (SAFS), we measured free T3 (FT3) in 1,276 Mexican American individuals and found that levels were significantly heritable (h^2= 0.625; p=1.1x10^{-54}). Using a deep whole genome sequencing approach, we concentrated on finding non-synonymous variants influencing FT3 levels in 696 Mexican American individuals from large families, for whom we have both FT3 and sequence variation available (obtained as part of the T2D-GENES Consortium). Initially, we focused on the potential effects of obligately functional protein altering variants that are predicted to be highly deleterious (PolyPhen-2 prediction score > 0.7) of which ~11,000 were observed in our sequencing set. We identified a private variant, Q295H, in the *OR2W5* (olfactory receptor family 2 subfamily W member 5) gene associated with a large increase of 3.4 SDU in FT3 levels (p=8.6×10^{-6}). Although this gene shows some similarities with other olfactory receptors, its function is relatively unknown. Expanding our focus to the remaining non-synonymous variants (~33,000), we identified a second candidate variant influencing FT3 levels. This common variant, V115F, in the *CIDEA* gene is associated with a decrease in FT3 levels of 0.34 SDU (p=3.0×10^{-7}). The V115F variant in *CIDEA* (cell death-inducing DFFA-like effector A) has previously been shown to be associated with metabolic syndrome and obesity, making it an excellent candidate for regulating thyroid hormone levels. Our results highlight the value of focusing on protein-altering variants for the identification of two new QTLs influencing thyroid hormone.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2340W

Exome Sequencing to Identify Causal Variants in a Multiplex Anencephaly Family. *D. Krupp, K. LaRocque-Abramson, K. Soldano, M.E. Garrett, A.E. Ashley-Koch, S.G. Gregory.*Center for Human Genetics, Duke University, Durham, NC.

Neural tube defects (NTDs), including anencephaly, are a common class of human birth defects with an estimated worldwide incidence of 1 in 1000 births. Despite widespread prevalence in all ethnic groups, the majority of human NTDs are sporadic cases, suggesting an underlying oligogenic or polygenic mechanism. To further our understanding of genetic variants contributing to human NTDs, we performed exome sequencing of two affected siblings and their parents from a family with three consecutive anencephalic pregnancies. Exonic regions were enriched for the father and both affecteds using the Illumina TruSeq Exome Enrichment kit. The mother's exome was captured using Agilent's Sure Select Human All Exon 50 Mb kit. All enrichments were performed according to the respective manufacturer's protocol, and libraries were sequenced 2x100nt on an Illumina HiSeq instrument. For the exomes of the father and affected offspring, reads were aligned using the Burrows-Wheeler Aligner, duplicates removed using PICARD, and alignment refinement and variant calling performed with the Genome Analysis Toolkit. Variants were filtered by minimum read depth (3), variant quality score (50), mapping quality (45) and quality by depth (7) to reduce false positives. Enrichment of exonic and non-synonymous variants was evaluated in each individual by chi-squared test for 22 NTD-relevant Gene Ontology terms, including neural tube development, the Wnt receptor signaling pathway, SHH signaling, and cilium morphogenesis. In each affected, we observed suggestive to significant overrepresentation of nonsynonymous variants in the cilium cellular compartment (p=0.0279, p=0.0085) and the cilium morphogenesis biological process (p=0.024, p=0.108). No such enrichments were observed in the father, suggesting that defects in cilium formation and/or cilium-dependent processes underlie these anencephaly cases. Such defects have previously been associated with NTDs in mice and humans. Both affecteds were heterozygous for a novel missense variant in IFT172, not present in the father and rated “probably damaging” by Polyphen-2. The affecteds additionally shared three other low-frequency missense SNPs in intraflagellar transport genes. Further analysis, including that of the mother, is ongoing and will be presented at the meeting.

2341T

Statistical power and effect size overestimation for combined single point and sequence kernel association tests in case-control studies. *A.M. Valdes, F. Lari.* Twin Research, King's College London, London, United Kingdom.

The cost of generating whole genome sequences data implies that the number of cases that can be assayed is low compared to the sample size in a typical GWAS. New methods have been developed which employ gene or region based approaches. Based on published simulation results a regression method using the Kernel association test (SKAT) appears to give good results under most scenarios. *Objective*: to assess the statistical power and determine optimal strategies for signal selection for replication using as discovery whole genome /exome sequence data from 150–200 cases vs the same number of controls. *Methods*: Computer simulations were carried out using publically available phased sequence data from a portion of chromosome 1 from the 1000 genomes project. Variants in the NFBP3 and NFBP1 genes were annotated. Each missense variant was assumed to contribute to disease risk with an odds ratio (OR) of either 2.0 and 5.0, disease prevalence assumed 5%. All risk variants were assessed for single point analysis and both genes and other 10 genes in chr 1 were tested using SKAT. NFBP3 had 16 missense rare variants with a mean minor allele frequency (MAF) of 2.3% (median MAF =2.2%). NFBP1 had 18 missense variants (mean MAF =1.1% ; median MAF =0.7%). We ran 100 replicates for each model. *Results*: A small case control study does not enough power to detect single point association with rare variants even with very high p-values(p<0.01) and variants with MAF <2% cannot be detected even at very relaxed p-values (p<0.05) unless OR >5.0. ORs are consistently overestimated relative to the true effect size and the magnitude of the overestimation is a function of both MAF and p-value. Power for SKAT is lower for low MAF variants (median MAF= 0.7%) than for higher MAF rare variants (median MAF =2%). SKAT has better statistical power than single point association analysis. Statistical power can be improved and type I error reduced if both single point and kernel association results are combined for signal prioritization *Conclusions*: Small case-control studies (total n<= 400) using whole genome/exome data are underpowered for both gene based and single point analysis unless true individual variant effect sizes can be assumed to be OR>5.0. Very low MAFs (<1%) cannot be detected with either type of method even with nominal significance (p<0.05). A combination of single point and gene-based methods can improve power for signal prioritization and candidate region studies.

2342F

Exome sequencing of cluster headache patients with Sardinian ancestry. *L. Jouan[1], H. Daoud[1], B. Oulad Amar Ben Cheikh[1], A. Dionne-Laporte[1], D. Spiegelman[1], O. Diallo[1], E. Henrion[1], A. Szuto[1], J. Gauthier[1], P. A. Dion[1], M. Manchia[2], M. Alda[2], M. Del Zompo[3], G. A. Rouleau[1].* 1) Centre d'Excellence en Neuroscience de l'Université de Montréal, CRCHUM Hôpital Notre-Dame, Montreal, Quebec, Canada; 2) Mood Disorders Program, Department of Psychiatry, Dalhousie University, Halifax, Nova Scotia, Canada; 3) Section of Clinical Pharmacology, Department of Neurosciences, University Hospital Agency of Cagliary, Cagliary, Italy.

Cluster headache (CH) is a devastating disease characterized by recurrent attacks of short-lasting excruciating pain with regular periodicity. Albeit considered for many years a sporadic disease with absence of a heritable component, twin and family studies indicated a role for genetic factors. Complex segregation analyses suggested an autosomal dominant mode of inheritance with low penetrance in some families and a multifactorial inheritance or autosomal recessive in others. To date, only two genes were positively associated with CH, Hypocretin receptor 2 (HCRT2) and Alcohol dehydrogenase (ADH4). To identify novel CH causative genes, we performed exome sequencing of 23 unrelated CH patients from Sardinia, a genetically homogenous population and a valuable resource for fine mapping of susceptibility genes for complex diseases. The exome capture was done using the Agilent SureSelect v4 and the sequencing on Illumina HiSeq2000. Variants were called and annotated using an in-house data pipeline that uses the GATK software. Under the assumption that novel rare variants cause CH, variants were filtered against dbSNP, 1000 Genomes, Complete Genomics data set and >500 in-house exomes to exclude known or common variations. Candidate genes with truncated variants (non-synonymous, splices and Indels) shared by the highest number of CH patients will be prioritized, validated by Sanger sequencing and studied at the functional level. The exome sequencing of CH in patients with Sardinian ancestry has tremendous potential to identify novel CH-causing genes and provide novel insights into our understanding of this disease.

2343W

Whole-genome sequencing analysis of more than 700 members of an extended pedigree. *Y. Huang[1], A. Jasinska[1], V. Ramensky[1], N. Juretic[3], S. Service[1], N. Tran[1], R. Wilson[2], W. Warren[2], G. Weinstock[2], K. Dewar[3], N. Freimer[1].* 1) Center for Neurobehavioral Genetics, UCLA, Los Angeles, CA; 2) The Genome Institute at Washington University in St. Louis, St Louis, MO; 3) McGill University and Genome Quebec Innovation Centre, Montréal, Québec, Canada.

It is hypothesized that much of the genetic risk for complex traits derives from the effects of rare variants. Large-scale population sequencing projects to identify rare trait-related variants have mainly utilized exome-sequencing, and thus have not assayed non-coding variation. Although whole genome sequencing (WGS) is still prohibitively expensive in large population samples, WGS-based studies in extended pedigrees have advantages that make it currently viable to comprehensively capture rare risk alleles: a) There is a good chance in such pedigrees that rare variants are observed in multiple individuals, alleviating the need to conduct sample-wide high depth sequencing; b) The pedigree structure can be leveraged to improve variant calling; c) Due to the limited number of unique haplotypes in a pedigree, a varying-depth sequencing and imputation strategy can be used to obtain nearly comprehensive variants for every pedigree member. We illustrate these points through WGS of the Vervet Research Colony (VRC), which was established in the 1970s from 57 wild monkeys collected from two Caribbean islands, St. Kitts and Nevis. The Caribbean population in turn arose from a small number of vervets brought from West Africa to these islands during the 17–18th centuries. Thanks to the research conducted by a number of research groups during the past four decades, dozens of phenotypes have been measured on a pedigree-wide basis in the VRC. To conduct linkage and association mapping of these traits we sequenced 723 VRC monkeys. We describe a VRC variant discovery pipeline that takes the pedigree structure into account. In particular, we adapted TrioCaller (Chen et al. 2012, submitted), a program that refines genotype-calls from single-locus callers (GATK/SAMtools) and imputes missing genotypes based on haplotype information and trio constraints, to the vervet monkey setting, in which one male monkey could have as many as 25 children with different females. Using this pipeline, we were able to achieve a very low Mendelian inconsistency rate. We also describe a sequencing strategy that prioritizes an individual monkey's sequencing depth based on its position in the pedigree. Results from downsampling showed that we could sequence ~40% of monkeys at 1X with little reduction in the number of variant sites identified (~10 million) and the quality of genotype calls (Mendelian inconsistency rate below 0.5%).

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2344T

Sequencing Initiative Suomi (SISu): identification of loss of function variants enriched in the Finnish population. *K. Rehnström[1,2], E.T. Lim[3,4,5,6], A. Palotie[1,2] on behalf of the SISu Consortium.* 1) Wellcome Trust Sanger Institute, Hinxton, United Kingdom; 2) Institute for Molecular Medicine Finland, FIMM, University of Helsinki, Helsinki, Finland; 3) Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, Massachusetts, USA; 4) Program in Medical and Population Genetics, Broad Institute, Cambridge, Massachusetts, USA; 5) Departments of Genetics and Medicine, Harvard Medical School, Boston, Massachusetts, USA; 6) Program in Genetics and Genomics, Biological and Biomedical Sciences, Harvard Medical School, Boston, Massachusetts, USA.

Because of a founding bottleneck roughly 100 generations ago, the Finnish population offers substantial advantages in the study of rare and low frequency DNA variation. The Sequencing Initiative Suomi (SISu) aims to bring together all Finnish sequencing projects to make data available as a high density reference set for imputation as well as a control set for medical genetic sequencing studies. Currently, the project includes 3800 whole genomes and >5000 exomes from projects including FINRISK, GoT2D (FUSION, Botnia), H2000, METSIM, NFBC66, and UK10K. Of particular interest, the exome data reveals several key features that distinguish the Finns from other European populations that are of key importance to studying rare variation. We see a strong increase of variants with 1–5% frequency, with 40–50% more loss-of-function variants in this frequency range in Finns than other European samples. Since the founding bottleneck was evolutionarily recent, recessive lethal variants may still persist at or near 1% frequency. Unlike most populations, many recessive Mendelian mutations at this frequency have been documented in Finns. We also observe a strong decrease of singletons and very rare variants. We hypothesize that the enrichment of variants which are rare in other parts of Europe could aid in identification of novel recessive diseases similar to the 36 monogenic disorders of the Finnish Disease Heritage. We have identified functional variants which are enriched in Finland, and will genotype these in large population and patient cohorts to identify association with diseases and possible depletion of homozygous variants suggesting lethality. To identify enriched variants, we compared allele frequencies in a set of equally-sized Finnish and European deep coverage exomes and rigorously selected a set of variants that are at least 2–fold enriched in Finns, with a focus on loss of function variants. We are currently genotyping 145 variants in 2000 individuals from the Young Finns Study to validate allele frequencies and identify disease associations. Later, the most interesting 50–60 variants will be genotyped in over 33 000 individuals from the FINRISK studies, a well phenotyped population cohort to increase power to identify association with disease. The samples are linked to national registers, including the hospital discharge register, prescription drug register and cancer registry enabling detailed phenotype mining of rare variant carriers.

2345F

Diabetes related DNA mutations in Americans of Mexican origin with health disparities disclosed by next-gen sequencing. *H. Q. Qu[1], Q. Li[2], Y. Lu[1], S. P. Fisher-Hoch[1], J. B. McCormick[1].* 1) School of Public Health, Univ of Texas Health Science Center at Houston, Brownsville, TX; 2) Endocrine Genetics Lab, The McGill University Health Center (Montreal Children's Hospital), Montreal, Quebec, Canada.

Diabetes related health disparities are serious health issues in a south Texas Mexican American population with a 30% prevalence of diabetes, as shown by our study of our-randomly selected, community-based Cameron County Hispanic Cohort (CCHC:n=2600) (Fisher-Hoch et al., 2010; 2012). With the majority of the diabetes patients being classified as type 2 diabetes, it is unclear what proportion of the diabetes patients may be rare monogenic forms of diabetes, e.g. Maturity Onset Diabetes of the Young (MODY). MODY diagnosis is of clinical importance, both for determining best choices of therapeutic agents and for genetic counseling. However, MODY diagnosis is complicated and expensive, and in this economically disadvantaged community, diagnosis and screening are unavailable. Advanced genomic technologies have provided new opportunities to address this health disparity issue. We performed a NextGen sequencing study on 75 Mexican Americans from our cohort with diabetes and body mass index (BMI)<25, who were identified by younger ages (range from 18–50) from 1551 random participants of CCHC. Serving as the sequencing controls, 75 obese persons without diabetes selected by older age from the CCHC were also studied. To optimize the cost-benefit, we made 6 DNA pools with 25 individuals in each pool (age-ranked), limiting sequencing to 6 DNA samples. The first 3 DNA pools are 75 age-ranked diabetes patients, i.e. Pool 1 with age 18–33 years old; Pool 2 with age 33–43; and Pool 3 with age 43–50. Three other DNA pools are 75 controls. All 11 MODY genes (MODY1–MODY11) as well as 68 other diabetes genes (type 2 diabetes genes and neonatal diabetes genes) were included. Altogether, 873 exons (0.4 million base-pair genomic regions) of these 79 genes were deep-sequenced at 3000x depth, using NimbleGen Sequence Capture Array and Illumina HiSeq2000. Six MODY mutations were identified from Diabetes Pool 1 and Pool 2 (i.e. 12% of diabetes cases with BMI<25 and age≤43). Among these six, 4 novel MODY mutations (predicted by the SIFT algorithm) were identified. In addition to the MODY mutations, we identified 23 damaging mutations from other diabetes genes. Further validation of these mutations by Sanger Sequencing is in progress. The mutation data in this Mexican American population provide critical health information about the rate of MODY in this population, which will help clinical decision of MODY mutation tests and minimize this diabetes health disparity issue.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2346W

Exome Sequencing in Severe COPD Cases and Resistant Smoking Controls from COPDGene. *M.H. Cho[1], M. Emond[2], R.A. Mathias[4], J.E. Hokanson[7], T.H. Beaty[5], N. Laird[6], C. Lange[6], K.C. Barnes[5], M. Bamshad[3], J.D. Crapo[8], E.K. Silverman[1], COPDGene Investigators, NHLBI Exome Sequencing Project.* 1) Channing Laboratory and Pulmonary and Critical Care Medicine, Brigham & Women's Hosp, Boston, MA; 2) Department of Biostatistics, University of Washington; 3) Department of Genome Sciences and Department of Pediatrics, University of Washington; 4) Department of Medicine, Johns Hopkins University; 5) Bloomberg School of Public Health, Johns Hopkins University; 6) Harvard School of Public Health, Harvard University; 7) Colorado School of Public Health, University of Colorado Denver; 8) National Jewish Medical and Research Center.

Background: The burden from chronic obstructive pulmonary disease (COPD) is rising; it now ranks third as a cause of mortality in the United States. Only a minority of cigarette smokers develops disease. Rare protein-coding variants are known to have a role in COPD susceptibility: a small fraction of COPD patients have alpha-1 antitrypsin deficiency (AATD), a well-known Mendelian disorder typically caused by a rare nonsynonymous SNP in SERPINA1 and characterized by severe COPD. COPD is highly heterogeneous; AATD and other monogenic human and smoke-exposed mouse models share the COPD-associated feature of emphysema, or lung destruction. We hypothesize additional rare coding variants, identified by exome sequencing, could also contribute to COPD and emphysema susceptibility. Methods: We selected extremes of COPD susceptibility and emphysema among Non-Hispanic Whites from the COPDGene study, frequency-matched on cigarette smoking. Cases were age < 63.0, severe COPD (GOLD Stage 3–4), and >15% emphysema on CT scan. Resistant smoking controls had age > 65.0, normal lung function, and <5% emphysema. Exome sequencing was performed using the Nimblegen SeqCap EZ V2.0 on Illumina platforms as part of the NHLBI Exome Sequencing Project LungGO. Sequences were aligned using BWA, and variants were called using the Genome Analysis Toolkit. Our primary analysis was a per-gene burden test using a fixed minor allele frequency cutoff (5%) adjusting for exome-based ancestry principal components and P-values calculated by permutation as implemented in Score-Seq. Secondary analyses were performed using SKAT-O. Results: After quality control, a total of 206,546 SNPs, of which 103,809 were nonsynonymous, splice, or stop remained among 205 cases and 196 controls. Quantile-quantile plots demonstrated no evidence of substantial population stratification. Our top genes in our T5 analysis (with P-values from $7.6x10^{-6}$ to $6.0x10^{-4}$) were SHROOM3, SUN2, DMWD, KIF14, PCDHA6, and SLC6A19; however, none of results were exome-wide significant after adjustment for the number of genes tested. Conclusions: Exome sequencing in severe COPD and resistant smoking controls may identify novel genes that contain rare variants that could affect COPD susceptibility. A second phase of association using the Illumina Exome Chip in 2400 subjects is under way.

2347T

Meta-analyses of 6391 Danish individuals identify novel genetic loci associated with circulating levels of Interleukin 18, Insulin-like growth factor-binding protein 1, and Heat shock 70kda protein 1b. *T.S. Ahluwalia[1], N. Grarup[1], A. Albrechtsen[1], T. Sparsø[1], C.H. Sandholt[1], T. Lauritzen[2], D.R. Witte[3], T. Jørgensen[4], S. Hamren[5], M. Rowe[5], J.A. Kolberg[5], Y. Li[1,6], R. Nielsen[7,8,9], J. Wang[1,6,8], T. Hansen[1,10,11], O. Pedersen[1,11,12,13].* 1) Novo Nordisk Foundation Centre for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark; 2) School of Public Health, Department of General Practice, University of Aarhus, Aarhus, Denmark; 3) Steno Diabetes Center, Gentofte, Denmark; 4) Research Centre for Prevention and Health, Glostrup University Hospital, Glostrup, Denmark; 5) Tethys Bioscience, Emeryville, California, USA; 6) BGI-Shenzhen, Shenzhen, China; 7) Department of Integrative Biology, University of California Berkeley, Berkeley, California, USA; 8) Department of Biology, University of Copenhagen, Copenhagen, Denmark; 9) Department of Statistics, University of California Berkeley, Berkeley, California, USA; 10) Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark; 11) Hagedorn Research Institute, Gentofte, Denmark; 12) Faculty of Health Science, University of Aarhus, Aarhus, Denmark; 13) Institute of Biomedical Science, University of Copenhagen, Copenhagen, Denmark.

Circulating levels of specific proteins act as signatures for various pathologies. Interleukin 18 (IL18), heat shock 70kd protein 1B (HSPA1B), Insulin like growth factor binding protein 1 (IGFBP1) and 2 (IGFBP2), and Interleukin 2 receptor alpha chain (IL2RA), are biomarkers with established associations with disorders like type 2 diabetes, cardiovascular diseases, insulin resistance, metabolic stress, inflammation and various forms of cancers. Albeit these circulating proteins thus may act as biomarkers for a plethora of disease traits, studies identifying the underlying genetic components for these biomarker levels are lacking. We therefore sought to identify the genetic determinants responsible for fasting serum levels of these five biomarkers from exome-wide genotype data in 6391 middle-aged Danes. In this study 16,192 coding single nucleotide variants (SNVs), discovered by hybrid-capture and 8X whole exome sequencing in 2,000 Danish individuals, were further genotyped in the population-based Inter 99 cohort (N=5483) and the ADDITION cohort (N=908) with measurements of fasting levels of IL18, IGFBP1, IGFBP2, HSPA1B, and IL2RA. After linear regression analyses in the two study samples, meta-analyses were performed for each biomarker trait in a total of 6391 Danish individuals. Two loci associated with IL18 levels, three loci associated with IGFBP1 levels and one locus associated with HSPA1B levels at $P < 5 \times 10^{-8}$. All these genome wide significant loci were novel. Out of the two loci identified for IL18 levels the stronger signal was for a common missense variant with a minor allele frequency (MAF) of 0.38 on chromosome 2p22 ($P = 1.54 \times 10^{-12}$) and the second novel locus (common and non-synonymous) was situated on chromosome 5q13 ($P = 1.17 \times 10^{-10}$). The three novel loci associated with IGFBP1 levels were located on chromosomes 11q12 ($P = 1.99 \times 10^{-16}$), 4q35 ($P = 2.16 \times 10^{-13}$), and 2p23 ($P = 5.64 \times 10^{-10}$), respectively. All these SNVs were common (MAF > 0.05) and non-synonymous. For circulating HSPA1B levels we identified one novel genome-wide significant SNV situated at chromosome 1q32 ($P = 1.18 \times 10^{-15}$). Exome-sequencing driven mapping for fasting circulating levels of IL18, IGFBP1 and HSPA1B demonstrates two novel loci for IL18, three loci for IGFBP1 and one locus for HSPA1B, respectively. All the identified SNVs are non-synonymous variants.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2348F

Further analysis of a visual migraine aura locus on chromosome 9q21–q22 using exome sequencing. *M.E. Hiekkala[1], M.A. Kaunisto[1,2], B.S. Winsvold[3,4], E. Hämäläinen[2,3], V. Artto[5], S. Vepsäläinen[5], M. Färkkilä[5], M. Kallela[5], A. Palotie[2,3], M. Wessman[1,2].* 1) Folkhälsan Institute of Genetics, Folkhälsan Research Center, Helsinki, Finland; 2) Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Finland; 3) Wellcome Trust Sanger Institute, Hinxton, UK; 4) Oslo University Hospital, Ullevål, Norway; 5) Helsinki University Central Hospital, Helsinki, Finland.

Migraine is a common and disabling neurological disorder characterized by recurrent attacks. Migraine aura is the other main phase of a migraine attack, along with typical headache. Most typically (> 90%) aura is visual. We have identified a visual aura locus in Finnish families on 9q21–q22 (Tikka-Kleemola et al. 2010) residing near familial occipitotemporal lobe epilepsy and combined migraine with aura locus (Deprez et al. 2007). Further evidence for a migraine locus on 9q22 was recently reported in the genetic isolate of Norfolk Island (Cox et al. 2012).

In order to further restrict the visual aura locus and identify susceptibility variants we selected 13 migraine families with visual aura and increased frequency of epilepsy linked to the 9q21–q22. We performed Illumina OmniExpress genotyping in 142 individuals and combined this data with whole exome sequencing data (Agilent's 50 Mb capture kit) from 21 migraine with visual aura patients belonging to eight families. Screening the exome sequencing data on a 40 Mb susceptibility region using Exome Variants Analysis (EVA) -program highlighted 27 candidate genes containing potentially functional, low or rare frequency variants. Many of these candidate genes are involved in functions of membrane transport (five of them being ion channels) or neural development. The frequency of 60 potential variants in the Finnish population was checked against already existing control Finnish exome data. The variants having allele frequency less than 3 % in the Finnish population or that were more common in migraine patients than in Finnish control subjects were included in further analysis. After this filtering step 33 variants remained one of which was nonsense and 24 were missense. So far 18 of these variants (two novel) have been validated using Sanger sequencing. Segregation and replication analyses are in progress.

2349W

Extensions of Human Variation Database in Japanese Integrated Database Project. *A. Koike[1], N. Nishida[2], M. Yoshida[1], M. Kawashima[2], I. Inoue[3], S. Tsuji[4], K. Tokunaga[2].* 1) Hitachi, Ltd. Cent. Res. Lab., Japan; 2) Dept Human Genetics, Univ. Tokyo, Japan; 3) Dept. of Mol. Life Sci. and Mol. Med., Nat. Inst. of Genet; 4) Dept Neurology, Univ. Tokyo.

Recent advances in microarray-based SNP genotyping technology and next-generation sequencing (NGS) technology drastically changed the nature of disease-related genetic study on a cost-effective genome-wide manner. Since these methods produce valuable but large amount of data, these data management has come to an issue. Our organization has developed `Human Variation Database' as a repository database for genome-wide association studies (GWAS) and novel genetic identified variations detected by NGS (https://gwas.biosciencedbc.jp/) to achieve continuous and intensive management of variation data to prevent them from dissipation and for data-sharing among researchers. Up till now, we have widely called for data submission and offered data-repository and data-redistribution services. In this study, we have extended the Human Variation Database to accumulate variations extracted from scientific literatures by manual curation. Disease names are systemized mainly based on disease ontology and more than 3500 variations are currently registered with their variation kinds and their positions on the reference genome/gene/protein-sequences according to the nomenclature defined by HGVS. Collected variations include short/long insertions/deletions and structural variations detected by various measurements, and they are related to disease susceptibility, virus resistance, and drug response. Their statistical genetic results in each population and corresponding clinical conditions are also incorporated to shed light on locus/variation specific clinical characteristics and deepen our understanding of disease mechanisms. The reference genomes in each population are systematized using healthy public data and population difference in disease-related variation are also displayed on the web browser with other information such as conservation scores. Variations and disease are cross-searchable to clarify their relationships. In this presentation, we provide the overview of database structure, data management policies, and population difference of disease related variations in various diseases.

Acknowledgement: This work was supported by the contract research fund “Database Integration program” from the Japan Science and Technology Agency.

2350T

Genetic burden analysis of common associated variants in an isolate population of Southern Ostrobothnia and in extended multiplex families. *V. Leppa[1,2,3], E. Jakkula[1,4], S. Ripatti[1,2,5], P.-A. Gourraud[6], K. Koivisto[7], P. Tienari[8], J. Kaprio[1,9,10], J. Eriksson[2,11,12], T. Lehtimäki[13], S. Koskinen[14], A. Palotie[1,5,15], J. Saarela[1,2].* 1) Institute for Molecular Medicine FIMM, Helsinki, Helsinki, Finland; 2) Department of Chronic Disease Prevention, Institute for Health and Welfare Finland, Helsinki, Finland; 3) Helsinki Biomedical Graduate School, University of Helsinki, Helsinki, Finland; 4) Department of Medical Genetics, Helsinki University Hospital, Helsinki, Finland; 5) Human Genetics, Wellcome Trust Sanger Institute, Genome Campus, Hinxton, Cambridge, United Kingdom; 6) Department of Neurology, University of California San Francisco School of Medicine, San Francisco, California, USA; 7) Department of Neurology, Seinäjoki Central Hospital, 60220 Seinäjoki, Finland; 8) Department of Neurology, Helsinki University Central Hospital, Molecular Neurology Programme, Biomedicum Helsinki, University of Helsinki, Helsinki, 00290, Finland; 9) The Department of Public Health, University of Helsinki, Helsinki, Finland; 10) Department of Mental Health and Substance Abuse Services, National Institute for Health and Welfare, Helsinki, Finland; 11) Unit of General Practice and Hospital for Children and Adolescents, Helsinki University Central Hospital, 00029 HUS, Helsinki, Finland; 12) Folkhälsan Research Centre, Helsinki, Finland; 13) Department of Clinical Chemistry, Fimlab Laboratories, Tampere University Hospital and University of Tampere School of Medicine, Tampere 33101, Finland; 14) Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, and Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku 20521,Finland; 15) Broad The Broad Institute of MIT and Harvard, Cambridge, MA, United States of America.

Multiple sclerosis (MS) is a complex autoimmune disease of the central nervous system. Twin and half-sib studies support the importance of inherited factors in MS predisposition. Close relatives of MS patients have an increased risk of MS, although there are few extended MS families. The prevalence of MS is higher in some population isolates compared to the surrounding areas. The Southern Ostrobothnia (SO) in Finland has a 2–fold prevalence compared to Finland in general and there is familial clustering of MS. Recent genetic burden studies in mixed populations of European descent support the accumulation of common variants in MS families. Our aim was to study the genetic burden caused by common MS associated variants in the Finnish SO isolate and in Finnish multiplex families. We used a weighted log-additive model and 51 SNPs that were reported in a recent international MS GWAS to calculate a genetic burden score. A non-parametric KS-test was used to evaluate the differences between the sample groups in a total of 608 MS cases and 5534 population samples. The genetic burden was significantly increased in the MS patients compared to population samples in all regional and familial comparisons ($p < 0.0001$). However, the analysis of the isolate revealed that the SO cases (n=111, mean=6.64) did not show an increased genetic burden compared to the cases from elsewhere in Finland (n=497, mean=6.77) despite the 2–fold difference in prevalence. Likewise, the SO population samples (n=135, mean=6.31) were not significantly different from the general population (n=5399, mean=6.23). The familial analysis showed similar results: the familial cases (n=63, mean=6.80) were not significantly different from the cases with no reported family background (n=522, mean=6.73). It should be noted, that 44% of the families originate from the isolate. The common variants do not seem to explain the increased prevalence in the isolate or in the Finnish multiplex families. We hypothesize that there are specific variants that contribute to MS predisposition in the isolate region and in the multiplex families.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2351F

Assessing association between known PTPN22 and SUMO4 variants and Type 1 diabetes (T1D) in an Indian population. *S.R. Manyam[1], U. Ratnamala[2], K.R.S.S. Rao[1], S.K. Nath[3], U. Radhakrishna[4].* 1) Department of Biotechnology, Acharya Nagarjuna University, Guntur, AP, India; 2) Department of Pharmacology, Creighton University, Omaha, NE, USA; 3) Arthritis and Clinical Immunology, Oklahoma Medical Research Foundation, Oklahoma City, USA; 4) Green Cross Voluntary Blood Bank Pathology and RIA laboratory, Ahmedabad, India.

Type 1 diabetes (T1D), also known as `insulin-dependent diabetes mellitus (IDDM)' and `juvenile-onset diabetes', results from a cellular-mediated autoimmune destruction of the `gb cells of the pancreas. The incidence of T1D is increasing worldwide, especially in younger children. T1D is a genetically heterogeneous disease, multiple genetic predispositions are believed to be responsible for disease etiology in different populations, and however non-genetic factors such as diet, infection or environmental risk factors play an important role in the expression of the disease. T1D can causes many serious health problems including blindness, hearing impairments, nerve damage, heart attack, kidney failure, osteoporosis, infections, pregnancy complications in women. A number of genetic association and linkage studies of different populations have identified twenty four loci with evidence of linkage for T1D on various chromosomes. In addition, an increased incidence of gene polymorphisms such as PTPN22, SUMO4, CTLA4 and IL2RA associated with T1D was reported, however no common gene mutations or pathogenic causative genomic variations have been identified until. We have recruited 350 sporadic cases with T1D patients, and an equal number of age-matched controls from south India. Recently we completed an initial targeted two known and well-associated their polymorphism in its involvement in the pathogenesis of Type-1 diabetes (i.e. PTPN22 and SUMO4) in 100 selected T1D individuals and equal controls using a single nucleotide polymorphism (SNP) array. The data showed PTPN22 polymorphism (SNP rs2476601, 1858C>T) was not associated with in any of the TD1 subjects of Indian origin as the SNP marker rs2476601 was monomorphic (G/G) in both T1D subjects and controls. The data of SUMO4 gene variation (rs237025, 163A > G, M55V) which has been shown to be a susceptibility marker in type-1 diabetes showed a higher frequency of combined AG and GG genotypes (64%) in the affected individuals than in matched controls (56%) of Indian origin. However, the “G” allele frequency between cases and controls are not statistically significant (40% in cases vs. 35% in controls, OR (C.I.) = 1.23 (0.82–1.85), chi2=1.07, p=0.30), probably due to small sample sizes. To the best of our knowledge, this is the first targeted genetic study on the association of genetic loci in Type 1 diabetic patients in India.

2352W

Duffy Interaction with chromosome 4 determine the number of Malaria episodes of individuals from Western Brazilian Amazon. *J. Pescarini[1,2], L.M. Garrido[1,2], L.C. Pereira[1,2], R.G.M. Ferreira[2,3], L.M.A. Camargo[1,2], H. Krieger[1,2].* 1) University of São Paulo, São Paulo, São Paulo, Brazil; 2) Instituto Nacional de Genética Médica Populacional, Brazil; 3) Fundação Oswaldo Cruz, Rondonia, Brazil.

In the 70′s it was observed that Plasmodium vivax and P. knowlesi use Duffy antigen, which is present in the erythrocyte membrane, as a receptor to entry in the target cell. The absence of Duffy antigen in the erythrocyte, commonly known as FY^-, was recognize just a few years ago as caused by a T -46 C mutation, mainly in the FYB allele. FY^- individuals were proved to be less infected by this two species of Plasmodium than FY^+ individuals (FyA, FyB, FyAFyB), being considered a protection factor to malaria. A Cohort of 924 individuals from Monte Negro and of 178 individuals from Portuchelo, both communities located in the Brazilian Western Amazon, revealed that the variable “number of malarial episodes” has a major genetic component related to it, probably located in chromosome 4 (Feitosa, 2002; Ferreira, 2008). In the same study, 850 indivíduals were tested for different erythrocyte antigens (Systems ABO, Rh, MNSs, Kell and Duffy) showing that FY^- was associated with low “number of malarial episodes” in this population. Based on the hypothesis that a genetic mechanism underlying the “number of malarial episodes” in Western Amazon was located in chromosome 4, Pescarini (2011) genotyped a subsample of 96 individuals (48 cases and their controls, who had more and less malaria episodes, respectively) with Affymetrix Genechip Nsp 250K. Analyzing only this chromosome in PLINK (Purcell., data), it was found an association between polymorphism rs17527389 and the “number of malarial episodes” (p = 0.012, with Bonferroni's corretion). In order to analyze if there is any interaction between rs17527389 and the FY system, Independence analysis was performed using FY^- or FY^+ and A or T alleles for both cases and controls. Significant association was verified (X^2_3=21.69; p<0.001). Odds ratios (OR) of Duffy negative (FY^-/FY^+) for allele A and allele T was also analyzed, showing that isolating “risk” (A) and “protection” (T) alleles, FY^- is still inversely related to individuals who showed higher number of malaria episodes (OR(A)= 0.12; OR(T)>6). In spite of small sample sizes, the presence of genetic interaction between Duffy antigen and rs17527389, was verified and may indicate the existence of undescribed physiological pathways.

2353T

Role of genetic variants of inflammatory genes with inflammation, malnutrition and susceptibility to End Stage Renal Disease in north Indian population. *R. Sharma, R.K. Sharma, A. Saxena, S. Agrawal.* Sanjay Gandhi Post Graduate Institute of Medical Sciences, Lucknow U.P.

Background-TNF α, IL6 and IL10 are potent regulators of inflammation and may be causative factor for various co morbidities associated with renal disease. In present study we have investigated association between inflammation, malnutrition and genetic variants of TNF α, IL6 and IL10 with genetic susceptibility to End stage renal disease(ESRD) in north Indian population. Methods-257 patients and 200 controls were registered in study. All baseline biochemical parameters were measured. Genotyping of TNFα-308G<A,TNFα-238G<A,TNFα-850C<T, IL6–174G<C, IL6–572G<C,IL10–1082A<G,IL10–819C<T and IL10–592C<A was done by PCR RFLP. Serum levels of IL6, IL10 and TNF-α were measured using ELISA. Nutritional status was assessed by SGA scores.Statistical analysis was performed by SPSS(Ver.15) Result and Discussion-Significant differences were observed between ESRD patient and control groups in all the biochemical parameters.TNFα-308G<A(O.R.= 3.0,P-value=0.002)TNFα-238G<A(O.R.=2.4,P-value=0.016),TNFα-850C<T(O.R.=3.1,P-value=0.003, IL6–174G<C(O.R.= 3.7,P-value=0.001),IL6–572G<C (O.R.= 2.1,P-value= 0.007)and IL10–1082A<G(O.R.=2.7,P-value=0.003) were significantly different between patient and control groups. The risk haplotypes were IL6 C/C and TNF-α AAC. On comparing inflamed and non inflamed patients diffrences were seen at IL6–174CC (O.R.=2.0 P-value=0.04),IL10–592AA(O.R. = 2.2,P-value=0.02) IL10–1082 GG,GA(O.R.=2.1 P-value=0.04) and IL10–819TT (O.R. =2.2 P-value=0.01)genotypes. These were associated with 2.0 fold higher risk of inflammation. On grouping patients in 2 categories based on nutritional status-normal and malnourished the risk associated genotypes were-TNF-α-238AA (O.R.=4.1 P-value=0.01)and IL6–572GG(O.R.=2.2 P-value=0.05)among malnourished. Levels of IL6,TNF-α and IL10 were significantly different between patient and control group. Malnourished category showed higher levels of IL6 (P-value=0.03)and TNF-α (P-value=0.04). TNF-α-308AA, TNF-α-850TT, IL6–174CC and IL10–592AA were significantly associated with higher TNF-α, IL6 and lower IL10 levels respectively. Conclusion-Our results suggest that inflammatory markers are not only predictors of disease but also important in prediction of malnutrition associated with chronic diseases like ESRD.

2354F

Genetic predisposition to hypertension associates with increased risk of type 2 diabetes. *Q. Qi[1], J. Forman[2], F. Hu[1], L. Qi[1].* 1) Department of Nutrition, Harvard School of Public, Boston, MA; 2) Channing Laboratory, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA.

Hypertension has been associated with increased risk of type 2 diabetes (T2D). Recent genome-wide association studies have identified multiple loci associated with blood pressure and risk of hypertension; however, it is unknown whether these loci are associated with risk of T2D. Such information could help to understand the potentially causal link between hypertension and T2D. We examined the association between genetic predisposition to hypertension and risk of T2D in two nested case-control studies. The present study included 3082 women (1328 T2D cases, 1754 controls) and 2411 men (1113 T2D cases, 1298 controls) of European ancestry from the Nurses' Health Study and Health Professionals Follow-up Study. A genetic predisposition score was calculated by summing up the number of risk alleles of 29 established blood pressure associated genetic variants. The genetic predisposition score showed consistent associations with risk of hypertension and T2D in men and women (all *P* for heterogeneity >0.46), and therefore the results from the 2 cohorts were combined. Each additional hypertension-predisposing allele in the genetic predisposition score was associated with a 6% increased risk of hypertension (odds ratio [OR] 1.06 [95% CI 1.04–1.08], P = 2.0×10^{-8}), and a 3% increased risk of T2D (OR 1.03 [1.01–1.05], *P* = 0.004). Adjustment for hypertension abolished the association with T2D (*P* = 0.13). We also applied Mendelian randomization analysis, and found that the observed effect size of association between the genetic predisposition to hypertension and T2D was not different from the expected effect size (*P* = 0.13). In conclusion, genetic predisposition to hypertension is associated with an increased risk of T2D. Our results support a potential causal role of hypertension in the development of T2D.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2355W

A systemic monitoring of cleft lip/palate in Hebei province of China through a birth defect surveillance: a 15–year (1996–2010) report. *J-Z. Zhu[1], N. Zhong[1,2,3].* 1) Medical Genetics and Prenatal Diagnosis, Hebei Provincial Hospital, Shijiazhuan, Hebei, China; 2) New York State Institute for Basic Research in Developmental Disabilities, Staten Island, NY; 3) Peking University Center of Medical Genetics, Beijing, China.

Cleft lip with/without palate is the most common visible congenital anomaly, followed by congenital heart diseases, neural tube defects, polydactyly, spina bifida, and hydrocephalus, in Hebei province that is located at northeastern China. Here we present a 15–year systemic monitoring of cleft lip with/without palate in Hebei province through our birth defect surveillance system, which consists of 40 hospitals before year 2000, but expanded to 57 hospitals by year of 2005 and to 60 hospitals by the year of 2006 that is kept to now. Among these monitored hospitals, 16 are in urban and 44 (20 are in mountain area) are in rural areas. The prevalence of total number of people with cleft lip/palate was 664 per 10,000 in 1996, but increased to 1,104 per 10,000 in 2010, which gives an increase of 66%. During this period, the cleft lip/palate was significantly increased by 104.22% in the first 10 years from 1996 to 2005, then decreased from 12.81% in year 2006 to 11.04% in 2010. Looking into the details, we noticed that the decrease was mainly due to a decline of cleft lip without palate from 7.68% to 3.51% but increase of cleft lip with palate from 5.47% in 2006 to 7.35% in 2010. The underlying pathogenic mechanism for this phenomenon is yet unclear. Intensive biological investigation to test if there is any genetic or environmental factors influence the prevalence is recommended.

2356T

De novo and inherited copy number variants are a common cause of short stature. *C.T. Thiel[1], D. Zahnleiter[1], U. Trautmann[1], A.B. Ekici[1], D. Wieczorek[2], E. Kunstmann[3], A. Reis[1], H.-G. Doerr[4], A. Rauch[5].* 1) Institute of Human Genetics, University of Erlangen-Nuremberg, Erlangen, Germany; 2) Institute of Human Genetics, University of Essen, Essen, Germany; 3) Private clinic, Wuerzburg, Germany; 4) Department of Pediatrics and Adolescent Medicine, University of Erlangen-Nuremberg, Erlangen, Germany; 5) Institute of Medical Genetics, University of Zurich, Schwerzenbach-Zurich, Switzerland.

Shortness of stature is one of the most common pediatric concerns and is defined as a body height below -2 SD. The prevalence in the general population is 3%. In the majority of patients with idiopathic growth deficit the etiology remains elusive in the absence of morphological details. Recent GWAS (e.g. GIANT) found significant evidence for both single nucleotide and copy number polymorphisms associated with height variation in the general population. However, these associations explain only a small fraction of the overall variability of human height. To identify novel genetic causes of growth retardation under a “rare variant - frequent disease” hypothesis in 200 indexpatients with idiopathic short stature and their parents we performed molecular karyotyping using Affymetrix SNP 6.0 and CytoScan HD arrays and scored copy number variants (CNVs) with a minimum size of 10 kb. A total of 6,338 aberrations were identified. After exclusion of common polymorphisms using 820 healthy control individuals and comparison with known pathogenic CNVs, we carried out a gene-centric analysis by investigating known gene functions, tissue expression and murine knock-out phenotypes. We considered 20 CNVs (10 %) in 20 patients to be disease causing, including 10 duplications and 10 deletions. All aberrations were larger than 100 kb. 7 were de novo with 1 overlapping the 22q11.1 microdeletion region, and 13 inherited from the also affected parent, including the known microdeletion/duplication loci 1p36, 1q21, 2q33, 22q11.2 und 3q29 in 6 cases. The observed number of de novo CNVs was significantly increased (P value <0.01) compared to the expected level of de novo events. Furthermore, evidence of pathogenicity was strengthened by comparison to a cohort of patients with intellectual deficit and the Decipher database, as well as evaluation of phenotype data from animal models, pathway data and haploinsufficiency scores. 5 CNVs comprise loci of genome wide association (Lango Allen et al. 2010) scored on a level of significance for genome wide CNV association of estimated 2.5E-5. In conclusion, our data indicate that CNVs are one of the main causes of growth retardation in a frequency comparable to other conditions as e.g. intellectual disability. Thus, molecular karyotyping should be considered in idiopathic short stature.

2357F

Involvement of sex in differential expression of miRNA in the circulation of subjects with metabolic syndrome: primarily in women. *Y. Wang[1,2], S. Juo[1,3,4].* 1) Medical Genetics, Kaohsiung Medical University, Kaohsiung, Taiwan; 2) Department of Internal Medicine, National Taiwan University Hospital Yun-Lin Branch, Taiwan; 3) Medical Research, Kaohsiung Medical University Hospital, Kaohsiung, Taiwan; 4) Department of Neurology, Kaohsiung Medical University Hospital, Kaohsiung, Taiwan.

Context: The association between circulating miRNAs and metabolic syndrome (MetS) has not been investigated in detail and it is not known whether the change of circulating miRNAs shows a difference between genders. Objective: The study assessed the expression of circulating miRNAs in subjects with and without MetS, and also explored whether there are gender-specific differences in dysregulations of circulating miRNA levels. Design and Patients: The study analyzed serum miRNA levels by RQ-PCR from 104 human subjects, aged 41 to 73 years. The status of MetS was defined by the Adult Treatment Panel III criteria modified for Asians. Subjects who self-reported a history of cardiovascular diseases or on current medication for hyperglycaemia or hyperlipidemia were excluded. Samples which had contamination due to hemolysis were further abandoned. Results: The percentage of MetS participants was 31.4% in men (n=51) and 32.1% in women (n=53). The levels of circulating miRNAs (let7g and mir-221) were higher in MetS subjects (p=0.002 and p=0.008, respectively). The two candidate miRNAs displayed an increase in blood serum when the number of MetS risk components was increased. However, the alterations of let7g and mir-221 in the circulation were prominent in women (both p <0.05 in women but p >0.1 in men). Among women, the serum level of let7g was elevated in individuals with low levels of high-density lipoprotein cholesterol and high blood pressure (p=0.008 and 0.02, respectively). On contrary, circulating mir-221 was upregulated in female subjects with high fasting glucose level (p= 0.025). Conclusions: The circulating levels of let7g and mir-221 were elevated in individuals with metabolic syndrome and these alterations were primarily in women.

2358W

Evidence of a Causal Relationship between Adiponectin Levels and Insulin Sensitivity: a Mendelian Randomization Study. *H. Gao[1,2,3], T. Fall[1], R. van Dam[2,3], A. Flyvbjerg[4], B. Zethelius[5], E. Ingelsson[1], S. Hägg[1].* 1) Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, SwedenStockholm, Sweden; 2) Saw Swee Hock School of Public Health, National University of Singapore, Singapore; 3) NUS Graduate School for Integrative Sciences and Engineering, National University of Singapore, Singapore; 4) The Medical Research Laboratories, Department of Clinical Medicine, Faculty of Health, Aarhus University, Aarhus, Denmark; 5) Department of Public Health and Caring Sciences/Geriatrics, Uppsala University, Uppsala and Medical Products Agency, Uppsala, Sweden.

The adipocyte-secreted protein adiponectin has been associated with insulin sensitivity in observational studies. We aimed to evaluate whether this relationship is causal using a Mendelian randomization approach. In a sample of 71–year-old Swedish men (n=942) from the Uppsala Longitudinal Study of Adult Men (ULSAM), insulin sensitivity was measured by the euglycemic insulin clamp (M/I ratio). We used genetic variants in the *ADIPOQ* gene as instrumental variables to generate causal estimates for the effect of adiponectin on insulin sensitivity. These estimates were compared with those derived by conventional linear regression of insulin sensitivity on adiponectin. The three *ADIPOQ* variants, rs17300539, rs3774261 and rs6444175 were strongly associated with serum adiponectin levels (P<5.4x10^{-9} for all SNPs). Each of the variants were also significantly associated with M/I ratio in the expected direction (P≤0.022 for all SNPs). Instrumental variable analysis confirmed that genetically determined adiponectin increased insulin sensitivity (β=0.470 for rs17300539, β=0.678 for rs3774261 and β=0.805 for rs6444175; P≤1.4x10^{-3} for all SNPs) at similar effect sizes as observational estimates (β=0.498, $P_{difference}$≥0.136 by Durbin-Wu-Hausman test). The genetically determined and observed adiponectin level associations remained similar after adjusting for body mass index (BMI) to address mediation by adiposity. The observed association between higher adiponectin levels and increased insulin sensitivity is likely to represent a causal relationship not mediated by adiposity.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2359T

Heritability of daily physical activity as assessed by combined heart rate and movement sensing in 771 same-sex twin pairs. *M. den Hoed[1], S. Brage[1], J.H. Zhao[1], A. Nessa[2], K. Westgate[1], U. Ekelund[1], T.D. Spector[2], N.J. Wareham[1], R.J.F. Loos[1,3].* 1) MRC Epidemiology Unit, Institute of Metabolic Science, Cambridge, UK; 2) Department of Twin Research and Genetic Epidemiology Unit, St Thomas' Campus, King's College London, St Thomas' Hospital, London, UK; 3) The genetics of obesity and related metabolic traits program, Mount Sinai School of Medicine, One Gustave L. Levy Place, Box 1003, New York, NY, 10029–6574, USA.

Background: Despite the health-promoting effects of physical activity (PA) and despite efforts to promote PA in the population, the majority of adults do not adhere to the current PA guidelines. The presence of innate, genetic factors that drive PA behavior may explain this contradiction. However, twin and family studies examining the heritability of daily PA have so far been limited by poor PA measurement quality or small sample size. **Objective**: To examine the heritability of daily PA using combined heart rate and movement sensing in a large population of adult twins. **Methods**: PA was assessed in daily life for 6.8 ± 1.2 days (mean ± SD) in 419 monozygotic and 352 dizygotic same-sex twin pairs using Actiheart (1,542 individuals aged 56.3 ± 10.3 yrs, BMI 26.0 ± 4.8 kg/m^2). Actiheart provided information on trunk acceleration along the longitudinal axis of the body, representing bodily movements produced by skeletal muscles. It also provided information on PA energy expenditure (PAEE) and the time spent on sedentary, light intensity and moderate-to-vigorous intensity activities. We estimated the contribution of genetic (dominant and/or additive) and environmental (common and/or unique) factors to inter-individual variance in daily PA after adjusting for sex, age, age^2 and body size (body mass for PAEE; BMI for the remaining measures) using structural equation modeling. **Results**: Dominant and additive genetic factors together explain between 37% (95% CI 29% - 45%) of the variance in acceleration to 48% (95% CI 41% - 55%) of the variance in PAEE. The remaining variance is explained by unique environmental factors. Common environmental factors play at most a marginal role (0.5% - 6.4%). **Conclusion**: Although environmental factors explain the largest proportion of the variance in daily PA, genetic factors also play a considerable role. Innate biological processes may thus explain inter-individual differences in adherence to the PA guidelines in adults. Such processes may relate to physical fitness, perception of ability, sensitivity to internal/external rewarding cues, physical discomfort following PA and/or triggering of food intake when using PA as a weight loss strategy. Future gene discovery studies are anticipated to increase our understanding of PA behavior.

2360F

Relationship between Hysterectomy and Admixture in African American women. *L. Qi[1], R. Nassir[2,3], R. Kosoy[2,3], L. Garcia[1], L.E. Waetjen[4], H.M. Ochs-Balcom[5], M. Gass[6], J. Robbins[3], M.F. Seldin[2,3].* 1) Department of Public Health Sciences, University of California, Davis, Davis, CA, USA; 2) Department of Biochemistry and Molecular Medicine, University of California, Davis, CA, USA; 3) Department of Internal Medicine, University of California, Davis, CA, USA; 4) Department of Obstetrics and Gynecology, UC Davis Medical Center, Sacramento, CA, USA; 5) Department of Social and Preventive Medicine, University at Buffalo, Buffalo, NY, USA; 6) The North American Menopause Society, Cleveland, OH, USA.

Most studies suggest that hysterectomies are more common in African Americans than in other ethnic groups. To assess this ethnic surgical disparity in a novel way, our main goal was to determine whether admixture (the proportion of sub-Saharan African or European origin in individuals) is associated with hysterectomy frequency in African American women in the Women's Health Initiative (WHI). We used ancestry informative single nucleotide polymorphisms (SNPs) to estimate admixture proportions in >10,000 African American women from the WHI. Logistic regression models were used to investigate the association between admixture and hysterectomy with and without controlling for relevant covariates. We also considered other potential risk factors (adiposity, hypertension, and education) for hysterectomy accounting for admixture. African admixture was a strong risk factor even after adjusting for multiple covariates (OR 1.85, P<.0001). The admixture risk for hysterectomy was highest in the 35–39 age group (OR 3.08, P<.0001) and least evident in the oldest age groups (45 or older, ORs range from 1.14 to 1.35, P>.05). Our analyses also indicated that adiposity, hypertension and education were independent risk factors for hysterectomy in this population group. Higher African admixture is associated with higher frequency of hysterectomy in certain age groups. Although further work is needed to determine whether this is due to genetic or environmental factors, the results provide compelling data that genetic studies specifically targeting African American women and diseases associated with hysterectomy maybe especially useful in understanding the pathogenesis and underlying cause of this disparity in health outcome.

2361W

Genetic Loci Associated with C-reactive protein levels and risk of metabolic syndrome, chronic kidney disease, and left ventricular hypertrophy: The Nagahama Study. Y. *Tabara, C. Terao, F. Matsuda.* Kyoto University Graduate School of Medicine, Kyoto, Japan.

Background: Plasma levels of C-reactive protein (CRP) are known to be associated with chronic inflammation and cardiovascular diseases risks. Several genetic variations susceptible for CRP levels have been identified from GWAS in European populations. To clarify unidentified susceptible SNPs, we conducted a GWAS for plasma CRP levels in a Japanese general population. We also investigated association of identified susceptible SNPs with metabolic syndrome (MetS), chronic kidney disease (CKD), and left ventricular hypertrophy (LVH).

Methods: Study subjects were ~3,500 Japanese general population. DNA was extracted from peripheral blood samples. Genome-wide SNP analysis was performed using Illumina 610K quad chips. In the genome-wide association analysis, plasma CRP values were normalized by a rank-based inverse normal translation. Written informed consent was obtained from each participants.

Results: Three known loci, namely rs3091244 in CRP, rs7979473 in HNF1A, and rs769449 in ApoE gene, were reached to genome-wide significance. In addition, a novel suggestive locus (p=6.3*10–7) were identified in the CDKN1A region (rs3176343). Previous reports suggested that a number of genes associated with CRP levels including APOC1, HNF1A, LEPR, GCKR, and HNF4A were directly or indirectly related to metabolic regulatory pathways involved in diabetes mellitus. Suggestive association of the diabetes mellitus associated CDKN1A polymorphism in this study further supports the interpretation. Accumulation of the risk genotype showed step-wise association with CRP levels. However, no direct association was observed between the risk genotypes and MetS, CKD, and LVH, whereas the subjects with MetS (p<0.001), CKD (p<0.001) and LVH (p<0.001) showed significantly higher CRP levels.

Conclusion: Genotypes susceptible for plasma CRP levels were not directly associated with MetS, CKD, and LVH. The lack of association may argue against a causal association of CRP with these diseases.

2362T

Genotype-Phenotype Correlations in Migraine. *A.K. Sullivan, E. Atkinson, F.M. Cutrer.* Neurology, Mayo Clinic, Rochester, MN.

Migraine is a disabling, highly prevalent syndromic disorder which disproportionately afflicts women. Unlike for rare monogenic forms of migraine, approaches to identifying genes for the common forms of migraine have been of limited success. Case control studies have shown modest associations with migraine for several candidate genes, but reported associations have either not been replicated or have only marginal statistical significance. We propose that more clearly defining the clinical phenotype of migraine will help better elucidate the genetic component involved. Approximately 60% of female migraineurs report having migraine attacks around the time of there menstrual period. To investigate genetic factors involved in hormonally modulated migraine, we have selected 24 single nucleotide polymorphisms (SNPs) from a pathway-related group of genes involved in the hypothalamic-pituitary-gonadal-axis (HPGA) that have been implicated in previous studies to confer migraine susceptibility. From our DNA repository of individuals with migraine (the Mayo Migraine Genomic Library which includes age and gender matched controls), we have selected 1750 individuals' ages 18–50, 1132 cases (240 males, 892 females) and 608 controls (258 males, 350 females). We performed logistic regression analyses, and confirmed the association of the SNP rs4680 located within the catechol-O-methyl-transferase gene as well as two other SNPs rs2283265 and rs7131056 that are both located within the Dopamine Receptor D2 gene. When looking at females only, the rs4680 SNP association was still significant. We then compared nonhormonally modulated migraineurs to controls, and we found no associated SNPs. But when hormonally modulated migraineurs were compared to controls, there were 4 SNPs that were found to be significantly associated with migraine: rs4680 SNP, rs22832265 SNP located within the dopamine receptor D2 gene, and 2 SNPs rs2070762 and rs6356 located within the tyrosine hydroxylase gene. These findings could greatly impact clinical management of women with hormonal modulated migraine by suggesting the possible underlying mechanism of their migraine susceptibility leading to development of rational, biologically based therapies.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2363F

Association analysis of *Vitamin D receptor* gene polymorphisms and bone mineral density in postmenopausal Mexican-Mestizo women. *A. González-Mercado[1], J.Y. Sánchez-López[1], J.A. Regla-Nava[1], J.I. Gámez-Nava[2], L. González-López[3], E.A. Aguilar-Chávez[2], J. Durán-González[1], J. Rojo-Contreras[2], A. Celis[2], F.J. Perea-Díaz[1], M. Salazar-Páramo[2], B. Ibarra[1].* 1) División de Genética, Centro de Investigación Biomédica de Occidente, IMSS. Doctorado en Genética Humana, Centro Universitario de Ciencias de la Salud, Universidad de Guadalajara; 2) Unidad de Investigación Médica en Epidemiología Clínica, UMAE, Hospital de Especialidades, CMNO, IMSS; 3) Servicio de Reumatología del HGR110 IMSS.

Introduction Osteoporosis (OP) is a complex multifactorial genetic disorder with interaction of environmental and genetic factors. It is estimated that 46–62% of Bone Mineral Density (BMD) variation can be attributed to genetic factors, including variations in the expressions of different candidate genes encoding proteins such as calciotropic hormones, cytokines, growth factors, bone matrix proteins and growth factors receptors, which participate in bone metabolism. The Vitamin D receptor (*VDR*) gene has been extensively studied in relation to BMD, with controversial results of association in different populations around the world. ***Purpose*** The aim of this study was to analyze the association between the *Fok*I T>C (rs2228570), *Bsm*I G>A (rs1544410), *Apa*I G>T (rs7975232) and *Taq*I T>C (rs731236) polymorphisms and the haplotypes of the *VDR* gene with BMD in postmenopausal Mexican-Mestizo women from Western Mexico. ***Methods*** We included 320 postmenopausal women from Western Mexico and divided them into three groups, according to WHO criteria: 88 non-osteoporotic women (Non-OP; t-score >-1 SD), 144 osteopenic women (Opn; t-score between -2.5 and -1 SD) and 88 osteoporotic women (OP; t-score <-2.5 SD). The BMD measurements at the lumbar (L1–L4) spine, and the left and right femoral neck were performed by dual-energy X-ray absorptiometry. The four SNPs in the *VDR* gene were genotyped by real-time PCR. Statistical analyses were carried out using SPSS v16.0, and Arlequin v3.01. ***Results*** The genotype and allele frequencies of the four *VDR SNPs* were similar in the Non-OP, the Opn and the OP groups, with the following frequencies for the wild-type alleles: *Fok*I (T) 0.472, 0.510, 0.443; *Bsm*I (G) 0.739, 0.780, 0.773; *Apa*I (G) 0.569, 0.610, 0.563; and *Taq*I (T) 0.727, 0.780, 0.773, in the three groups respectively. We found no association of the SNPs with BMD. A haplotype analysis of the four *VDR SNPs* revealed that the TGGT and CGGT haplotypes were the most common combinations across all the groups with statistically significant differences in the frequencies for CGGT between the groups Non-OP and OP (21.8 vs 31.8%), and for TGGT between the groups Opn and OP (34.7 vs 22.9%), this last also showed a statistically significant result in the association analysis with the recessive model (OR=0.1957, IC 95%=0.06–0.67, p= 0.0047).***Conclusions*** Our results suggest a possible protective effect of the TGGT haplotype on the development of OP in postmenopausal Mexican-Mestizo women.

2364W

A novel signal of polygenic inheritance in Genome-Wide Association Studies (GWAS) of diseases and dichotomous traits. *Y. Chan[1,2,3], J.N. Hirschhorn[1,2,3].* 1) Department of Genetics, Harvard Medical School, Boston, MA; 2) Program in Medical and Population Genetics, Broad Institute, Cambridge, MA; 3) Department of Endocrinology, Boston Children's Hospital, Boston, MA.

Genome-wide association studies (GWAS) have been performed for a variety of traits and diseases and their associations with common genetic variants can be detected with sufficiently large sample sizes. The prevailing model for such studies is that variants have modest effects because the genetic contribution to the trait of interest is due to a combination of many loci throughout the genome (polygenic inheritance) as oppose to just one or a few loci. We considered the predicted results from GWAS under a liability threshold model of polygenic inheritance, where many variants contribute additively to an underlying unmeasured quantitative trait and where only individuals with a trait value above a specific threshold are affected with disease. We reasoned that, for a disease with prevalence well under 50%, the statistical power to detect minor alleles that increase disease risk can be substantially higher in a case-control analysis than the power to detect minor alleles that decrease risk, even if the magnitudes of effect on the underlying unmeasured trait are the same. This asymmetry in power raises the possibility that even apparently null GWAS studies may have a signal of polygenic inheritance, reflected in a skew towards nominally significant associations at risk compared with protective alleles. To determine the extent of this power difference, we theoretically evaluated the power to detect variants under varying parameters of threshold, allele frequency, effect size and sample size. We showed that the difference in power, and hence the imbalance between associated risk and protective minor alleles, increases with lower allele frequencies, larger effect sizes and lower disease prevalence until saturation occurs. For example, at a threshold of 1.96 (disease prevalence of 2.5%) with 2000 cases and controls, 1% variants with an effect size of +0.3 and -0.3 standard deviations will have 83.3% and 58.0% power respectively ($\alpha < 0.01$). We also generated sets of simulated data by simulating multiple causal alleles, varying the parameters described above and performed genome-wide case-control association analysis on them. This novel analysis can be performed on GWAS results to gain better insight into the allele frequencies and effect sizes of the causal variants. Furthermore, an imbalance in the associated risk and protective minor alleles can be used to infer the existence of associated variants even if no loci achieve genome-wide significance.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2365T

How allelic variants in INSL3 are associated with not testicular descent? An overview of a group of Mexican patients with idiopathic cryptorchidism. *M. Chávez[1], J. Gutiérrez[1], O. Cuevas[2], J. Rojas[1], O. Gutiérrez[1], D. Landero[1], E. Yokoyama[3], K. Jasso[4], L. Taja[5], R.M. Vigueras[1].* 1) Lab. Biología de la Reproducción, Instituto Nacional de Pediatría, Ciudad de México, Coyoacan, Mexico; 2) Servicio de Urología, Instituto Nacional de Pediatría, Ciudad de México, Coyoacan, Mexico; 3) Departamento de Genética Humana, Instituto Nacional de Pediatría, Ciudad de México, Coyoacan, Mexico; 4) Servicio de Cirugía Colorrectal, Instituto Nacional de Pediatría, Ciudad de México, Coyoacan, Mexico; 5) Lab. de Genética Molecular y Farmacogenética, Instituto Nacional de Cancerología, Ciudad de México, Tlalpan, Mexico.

Cryptorchidism (CO) is a common congenital anormality, affecting at least 2–4% of male births and could present itself like an isolated manifestation (idiopathic). CO is an important factor for male infertility and testicular malignancy on adulthood, its etiology is, in most cases, remain unknown, it is considered a complex disease because in the testicular descent are involved multiple genes like INSL3, RXFP2 and HOXA10. In this study we searched the allelic variants in the gene INSL3 than could contribute as risk factors or susceptibility to idiopathic CO in Mexican patients. 85 patients were included with idiopathic CO and 100 kids on the control group in a first phase, in the INSL3 gene were analyzed four SNPs (p.R102C, p.R102H, p.R105H, c.3334T- C) by allelic discrimination using the Taq-Man method ®. These SNPs weren't detected in our population for this reason we researched for unknown alterations by the method of automatic sequencing of the coding regions of INSL3 gene. The association of SNPs with the phenotype was analyzed with Fisher's exact test, a value of $p \leq 0.05$ was considered significant. The results show us that 100% of patients with CO and the control group were homozygous normal variants and p.R105H p.R102C, while p.R102H variant was detected in just one patient like heterozygous with bilateral CO without family history. The SNP c.3334T-C was the highest allelic heterogeneity in both groups, although there is not significative difference in them was observed the heterozygous genotype predominates in patients with CO. Before, the sequencing of both INSL3′s exons shows two SNP's in exon 1 (p.L42L and p.A60T) and two in exon 2 (p.R105R and p.T86M). There were identified five different alterations or allelic variants and these results represent 5.88% of affected alleles, of which two are for first time reported in the Mexican population. The genotype-phenotype correlation shows us a clear association of the alterations detected and the severe phenotype of our patients with idiopathic cryptorchidism, on the other hand we infer that the genotype TC allelic variant c.3334T-C could be associated with idiopathic cryptorchidism and the CC genotype represents a protective genotype for this disease. Project No. CONACyT 87101/2008, INP-063/2008.

2366F

Evidence of Inbreeding Depression on Human Height. *J.F. Wilson[1], N. Eklund[2,3], N. Pirastu[4], M. Kuningas[5], B.P. McEvoy[6], T. Esko[7], T. Corre[8], G. Davies[9], P. d'Adamo[4], N.D. Hastie[10], U. Gyllensten[11], A.F. Wright[10], C.M. van Duijn[5], M. Dunlop[10], I. Rudan[1], P. Gasparini[4], P.P. Pramstaller[12], I.J. Deary[9], D. Toniolo[8], J.G. Eriksson[3], A. Jula[3], O.T. Raitakari[13], A. Metspalu[7], M. Perola[2,3,7], M.R. Jarvelin[14,15], A. Uitterlinden[5], P.M. Visscher[6], H. Campbell[1], R. McQuillan[1], ROHgen.* 1) Centre for Population Health Sciences, Univ Edinburgh, Edinburgh, United Kingdom; 2) Institute for Molecular Medicine Finland (FIMM), Helsinki, Finland; 3) Department of Chronic Disease Prevention, National Institute for Health and Welfare, Helsinki, Finland; 4) Institute for Maternal and Child Health, IRCCS “Burlo Garofolo”, Trieste, University of Trieste, Italy; 5) Department of Epidemiology, Erasmus University Medical Center, Rotterdam, The Netherlands; 6) Queensland Institute of Medical Research, 300 Herston Road, Brisbane, Queensland 4006, Australia; 7) Estonian Genome Center, University of Tartu, Tartu, Estonia; 8) Division of Genetics and Cell Biology, San Raffaele Research Institute, Milano, Italy; 9) Department of Psychology, The University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, UK; 10) MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, EH4 2XU, Scotland; 11) Department of Immunology, Genetics and Pathology, SciLifeLab Uppsala, Rudbeck Laboratory, Uppsala University, Uppsala, Sweden; 12) Centre for Biomedicine, European Academy Bozen/Bolzano (EURAC), Bolzano, Italy - Affiliated Institute of the University of Lübeck, Lübeck, Germany; 13) Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland; 14) Biocenter Oulu, University of Oulu, Finland; 15) Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, MRC Health Protection Agency (HPA) Centre for Environment and Health, Imperial College London, London, UK.

Stature is a classical and highly heritable complex trait, with 80–90% of variation explained by genetic factors. In recent years, genome-wide association studies (GWAS) have successfully identified many common additive variants influencing human height; however, little attention has been given to the potential role of recessive genetic effects. Here, we investigated genome-wide recessive effects by an analysis of inbreeding depression on adult height in over 35,000 people from 21 different population samples. We found a highly significant inverse association between height and genome-wide homozygosity, equivalent to a height reduction of up to 3 cm in the offspring of first cousins compared with the offspring of unrelated individuals, an effect which remained after controlling for the effects of socioeconomic status, an important confounder. There was, however, a high degree of heterogeneity among populations: whereas the direction of the effect was consistent across most population samples, the effect size differed significantly among populations. It is likely that this reflects true biological heterogeneity: whether or not an effect can be observed will depend on both the variance in homozygosity in the population and the chance inheritance of individual recessive genotypes. These results predict that multiple, rare, recessive variants influence human height. Although this exploratory work focuses on height alone, the methodology developed is generally applicable to heritable quantitative traits (QT), paving the way for an investigation into inbreeding effects, and therefore genetic architecture, on a range of QT of biomedical importance.

2367W

Investigation of mitochondrial D-Loop variations in autism. *F. Piryaei[1], M. Houshmand[1], F. Piryaei[2].* 1) National Institute of Genetic engineering and Biotechnology, Tehran, Iran; 2) Department of Biology, Payame Noor University of Tehran, Tehran, Iran.

Autism as one of three recognized condition in the autism spectrum disorders (ASDs) is a neurodevelopmental, multifactorial disorder. In order to investigate if mutations raise within D-loop with respect to high volume of ROS, regarding this region plays an important role in controlling mtDNA transcription and replication, identified D-loop variation frequencies were compared between two patients and control groups. In this study, a cohort of 24 unrelated idiopathic patients and 100 ethnically-matched Persian control individuals were obtained. PCR sequencing of mtDNA fragments was employed to investigate the involvement of mitochondrial variations in autism. According to our data, C16223T showed a significant difference ($P = 0.043$, $P < 0.05$) between two groups with respect to its higher volume in control group. Hence it suggests a lower risk for expressing autism in individuals who carry this SNP. In order to confirm the role of this variation in reducing the autism expression, follow-up studies should be performed with sizeable samples of patients and controls. Haplogroups harbouring this variation also should be determined.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2368T

Blood-born diagnostic microRNA biomarkers for inflammatory bowel disease. *G. Hemmrich-Stanisak[1], M. Spehlmann[2], A. El-Sharawy[1], M. Hübenthal[1], A. Keller[3], A. Raedler[4], S. Zeissig[2], S. Nikolaus[2], S. Schreiber[1,2], A. Franke[1].* 1) Institute of Clinical Molecular Biology, University of Kiel, Kiel, Germany; 2) Clinic for Internal Medicine I, University Hospital of Schleswig-Holstein, Kiel, Germany; 3) Department of Human Genetics, Saarland University, Homburg, Germany; 4) Department of Internal Medicine, Asklepios Westklinikum Hamburg, Hamburg, Germany.

Invasive methods, such as endoscopy or colonoscopy are still the gold standard for diagnosis and evaluation of Crohn's disease (CD) and ulcerative colitis (UC). Due to their stability in tissues and body fluids, such as peripheral blood, differential signatures of microRNAs (miRNAs) could serve as new noninvasive biomarkers to predict and discriminate these two major subphenotypes of chronic inflammatory bowel disease (IBD). In this study, we determined the expression profiles of 863 miRNAs by array analysis of 114 blood samples (40 CD, 36 UC, 38 controls) to generate phenotype-specific miRNA signatures. Using a threshold of maximum 100 features per signature resulted in miRNA-signatures of 66 (CD) and 95 (UC) deregulated miRNAs that separated cases from controls with a balanced accuracy of 96% and 93%, respectively. To discriminate both subphenotypes a signature of 68 differentially expressed miRNAs is required (balanced accuracy 88%). The minimal miRNA signature consists of 18 miRNAs for CD and only 3 miRNAs for UC to discriminate between cases and controls with balanced accuracies of 80% and 83% respectively. 14 miRNAs are minimally required to distinguish both CD and UC from one another (balanced accuracy 81%). For validation of our results we plan to replicate the approach using an independent IBD-cohort. In addition an observer-blinded study will prove the predictive potential of the resulting phenotype-specific miRNA signatures.

2369F

Influence of emotionally-loaded visual stimuli on experimental jaw-muscle pain and its relation with the tri-allelic polymorphism on the serotonin transporter gene. *E. Horjales-Araujo[1], D. Demontis[2], N.B. Finnerup[1,5], A. Børglum[2], T.S. Jensen[1,3,4], P. Svensson[4,5].* 1) Danish Pain Research Center, Aarhus University Hospital, Aarhus, Norrebrogade 44 Building 1A, 8000; Aarhus C, Denmark; 2) Dept Human Genetics, Institute for biomedicine, Aarhus University, Wilhelm Meters Alle 1240 8000 Århus C; 3) Dept of Neurology, Aarhus University Hospital Norrebrogade 44, 8000 Aarhus C, Denmark; 4) MINDLab, Center of Functionally Integrative Neuroscience, Aarhus University Hospital, Norrebrogade 44, 8000 Aarhus C, Denmark; 5) Department of Clinical Oral Physiology, School of Dentistry; Aarhus University, Vennelyst Boulevard 9 800,0 Aarhus C, Denmark.

Pain is the most common reason patients seek medical attention. However, the symptoms' description and pain perception vary from patient to patient. The perception of pain is determined by a mosaic of genetic, neurobiological, cultural and emotional factors. Recent studies have demonstrated an association between some genotypes and pain perception. In specific, great focus have been given to the tri-allelic polymorphism in the promoter region of the serotonin transporter gene in relation with pain perception. In this study we aim to investigate whether the modulatory effect of the visual stimuli on muscular pain perception observed in our previous study is genotype dependent. This study was approved by the local ethical committee. DNA extracted from saliva (using OC-100; DNA genotek, Ontario, Canada) obtained from 380 Scandinavian healthy volunteers (between 18 and 39 years old). To determine the tri-allelic 5–HTTLPR genotype, PCR- reactions were carried out using the GoTaq® Hot Start Polymerase (Promega) and 80ng of genomic template. Additionally, 10 µl of the PCR product were digested for 2hrs at 37° with 1µl MSP1 (New England Biolabs, Ipswitch, MA, USA). The genotypes were evaluated from running PCR products and digested PCR products on agarose gels. A total of 150 healthy subjects (50 of each genotype for serotonin transporter gene (s/s, s/l and l/l) proceed to the next phase of our study. In this phase, experimental jaw muscle pain was evoked by injection of 0.2 ml hypertonic saline into the masseter muscle. With each injection an emotion (positive, negative or neutral) was generated by visual stimuli (pictures taken from International Affective Picture System). The participants were asked to continuously rate pain intensity (electronic visual analogue scale - VAS) as well as unpleasantness/pleasantness of the conditioning stimuli (pictures). Hypertonic saline evoked moderate levels of pain in all subjects. In participants with high expression of the protein serotonin transporter, conditioning with negatively-loaded pictures increased VAS pain scores compared with neutral and positive emotions. In general, participants with intermediate expression of the protein reported had experienced higher pain intensity than participants with high or low expression of the protein. However there were no significant effects of the pictures on pain perception in participants with neither intermediate nor low expression of the protein.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2370W

Genetic Association Analysis between Glycine Cleavage System and Schizophrenia. *A. Yoshikawa[1], F. Nishimura[1], H. Hibino[2], K. Kohda[3], M. Tochigi[1], Y. Kawamura[4], T. Umekage[5], T. Sasaki[6], K. Kasai[1], C. Kakiuchi[1].* 1) Neuropsychiatry, Medicine, The University of Tokyo, Tokyo, Japan; 2) Department of Psychiatry, Fukui Memorial Hospital. Kanagawa, Japan; 3) Department of Physiology, The Keio University School of Medicine, Tokyo, Japan; 4) Department of Psychiatry, Sakae Seijinkai Hospital. Kanagawa, Japan; 5) Division for Environment, Health and Safety, The University of Tokyo, Japan; 6) Department of Physical and Health Education, Graduate School of Education, The University of Tokyo, Japan.

Background: Schizophrenia is a devastating psychiatric disease with worldwide lifetime risk of approximately 1%. The pathophysiological mechanisms underlying schizophrenia have not been fully elucidated. Hypofunction of NMDA receptors (NMDAR) has been implicated in the pathophysiology of schizophrenia, but precise mechanisms are not known yet. Glycine is the co-agonist of NMDAR and glycine cleavage system (GCS) is the main pathway which catalyzes glycine. Interestingly, the activity of the GCS is co-localized with NMDAR in neurogenic regions especially in the hippocampus. In addition, several studies suggest that the serum concentration of glycine/D-serine is altered in patients with schizophrenia. Therefore, we hypothesized that the dysfunction of GCS is involved in the pathophysiology of schizophrenia. GCS is a mitochondrial multi-enzyme complex and is composed of GLDC (glycine dehydrogenase (decarboxylating)), AMT (aminomethyltransferase), GCSL (glycine cleavage system protein L) and GCSH (glycine cleavage system protein H). Here we conducted the genetic association study between GLDC and schizophrenia. Aim of the study: To investigate the genetic association of GCS genes with schizophrenia. Methods: We performed a case-control association study between GLDC and schizophrenia in 384 patients with schizophrenia diagnosed by DSM-IV criteria and 384 age and sex matched healthy controls. All the subjects were ethnically Japanese. Written informed consent was obtained from all the subjects. The Human Genome, Gene Analysis Research Ethics Committee of the Faculty of Medicine, University of Tokyo approved the study. We selected the 24 tag SNPs by tag SNPs picker of the Hapmap database (MAF>0.1, r2>0.8). Genotyping was performed by the fluorogenic 5′ nuclease method (Taqman) using ABI7900HT. We analyzed the association by using the Haploview programs. Results: We observed no significant association with any tested SNPs or haplotypes of GLDC. Discussion: No evidence for the association was found between GLDC and schizophrenia in the Japanese population. Assessing the association with larger samples will be needed.

2371T

A gene search for a new spinocerebellar ataxia. *K. Ishikawa, N. Sato, Y. Hashimoto, M. Obayashi, H. Matsuo, H. Mizusawa.* Dept Neurology, Tokyo Med & Dental Univ, Tokyo, Japan.

Spinocerebellar ataxia (SCA) is a group of neurodegenerative conditions showing autosomal dominant cerebellar ataxia as a cardinal clinical feature. SCA contains clinically and genetically heterogeneous conditions. Mutations in 22 SCAs (SCA1, 2, 3/Machado-Joseph disease (MJD), SCA5, 6, 7, 8, 10, 11, 12, 13, 14, 15, 17, 18, 23, 27, 28, 31, 35, 36 and dentato-rubral pallidoluysian atrophy (DRPLA)) have been identified, and the gene loci are known in the remaining 9 SCAs (SCA4, 19/22, 20, 21, 25, 26, 29, 30, 32). We have collected 332 index patients with SCAs. After screening common SCA mutations which are SCA1, 2, MJD, SCA6, DRPLA, and SCA31, we found 66 patients (22%) did not harbor any of these 6 mutations. We embarked a positional cloning combined with neurological investigations to identify a new mutation for SCA. Although there was a clinical heterogeneity among these 66 patients, we found that a majority of patients show relatively early age-of-onset (i.e. onset before 35 years of age) and predominantly cerebellar dysfunction. Genotyping was performed on 30 individuals from 6 families using the Affymetrix 6.0 SNP Array. The obtained data were subjected for linakge analysis using a computer software MERLIN. We found supporting data that these families show linkage to a few chromosomal segments not previously assigned for SCAs. We are currently in a process to search mutations in these segments using exome sequencing, copy number variation assay and candidate gene approach. Further study to screen mutations among candidate genes and to collect more families would be needed to identify a new mutation for SCA.

2372F

Combining genotype and gene expression data to unravel the genetics of Major Depressive Disorder. *S. Mostafavi[1], A. Battle[1], X. Zhu[2], S. Montgomery[1,3], JB. Potash[4], MM. Weissman[5], C. Haudenschild[6], C. McCormick[6], KB. Beckman[7], R. Mei[8], A. Urban[2], D. Koller[1,3], DF. Levinson[2].* 1) Department of Computer Science, Stanford University, CA; 2) Department of Psychiatry and Behavioral Sciences, Stanford University, CA; 3) Genetics, Stanford University, Stanford, CA; 4) Department of Psychiatry, University of Iowa, Iowa City, IA; 5) Department of Psychiatry, Columbia University, New York, NY; 6) Illumina, Inc., La Jolla, CA; 7) University of Minnesota, Minneapolis, MN; 8) Centrillion Biosciences, Mountain View, CA.

The goal of this project is to identify and characterize the variability in the transcriptome that is associated with major depressive disorder (MDD), and to link such associated variability with environmental or causal genotypic variations. To this end, we have collected RNA-sequencing data from whole blood (over 60 million reads per subject), SNP genotype data (737,187 common SNPs), and a large number of phenotypic measurements for 949 subjects, including 474 fully eligible MDD cases and 475 controls. All subjects were of self-reported European ancestry, and recruited from a survey panel that was representative of the US population. Panel members were screened with an online CIDI-SF for (a) absence of current substance dependence, and (b) presence of either recurrent MDD or the absence of any two-week period with more than 3 of 9 MDD criteria. These individuals were then interviewed with the SCID (DSM-IV) after giving a blood sample for DNA and RNA (Paxgene collection system). Beyond quantifying gene expression levels, RNA-sequencing allows us to quantify more detailed expression traits such as isoform variability and allele-specific expression (ASE). Using these data, we have developed a statistical model for identifying groups of co-expressed expression traits (modules), including ASE, gene level and isoform variability, that are predictive of case control status, independently of other phenotypic covariates that can confound such associations. In particular, a major challenge in understanding the expression variations that underlie disease status is the abundance of spurious correlations caused by a number of confounding factors, such as age, gender, BMI, and drug intake. To address this issue, our method (1) models these confounding factors and their potential effect on expression variability; and (2) identifies gene modules associated with depression in our dataset, with varying degrees of additional support from genotypic variations. Our preliminary results identify several modules that predict disease status and severity. We are in the process of characterizing these modules for biological meaning, and their link to disease severity. These results provide new insights into genetic factors underlying MDD risk.

2373W

Transcript profile analysis of interleukin-1 beta gene in immature and adult zebrafish brain after seizure. *P.G. Barbalho, F.R. Pinto, R. Mangolin, C.V. Maurer-Morelli.*Department Medical Genetics, University of Campinas, Campinas, Campinas-SP, Brazil.

Introduction: Epilepsy is a common neurological disorder characterized by recurrent spontaneous seizures that afflicts nearly 50 million people worldwide. Proinflammatory cytokines, such as interleukin-1beta (il1b), has shown to be up-regulated in surgical specimens of pharmacoresistent patients and in experimental rodent models of seizure; however the profile of this cytokine in zebrafish model for epilepsy is unknown. In this study we aimed to investigate the il1b transcript profile in adult and immature zebrafish brain after seizure induced by Pentylenetratazole (PTZ). Material and Methods: Zebrafish were maintained according to standard procedures and all experiments were approved by animal ethical committee/UNICAMP. Adult and larvae (7days post fertilization) zebrafish were separated in Seizure (SG) and Control (CG) groups. Animals from SG group were individually exposed to 15mM PTZ (2–3min for adults and 20min for larvae). Animals from CG were exposed to same handling condition, but in normal bath water. A total of n=6 samples were used for each group and age. A single adult sample was composed by pooling two brains and 20 heads were used to compose each larvae sample. At the time 12h, 24h and 48h after seizure insult, animals were anesthetized, the brains were immediately isolated, frozen in N2(L) and total RNA extracted by Trizol (Invitrogen). Reverse transcriptase quantitative-PCR amplifications were carried out in triplicates with efla as endogenous control (TaqMan®, Applied Biosystems). The relative quantification (RQ) was calculated by the equation $RQ=2^{-\Delta\Delta CT}$. Statistical analysis was performed by Kruskall-Wallis test. Results: The mean ± Standard Deviation results obtained were: (i) adult: CG12h 0.95±0.31; SG12h 0.84±0.19; CG24h: 0.75±0.23; SG24h: 0.76±0.27; CG48h 0.85±0.16; SG48h 0.99±0.41; (ii) immature: CG12h 1.84±0.46; SG12h 2.00±0.61; CG24h: 2.18±1.07; SG24h: 1.15±0.51; CG48h 2.5±0.58; SG48h 2.02±0.63. No statistical significance was found between CG and SG ($p \geq 0.05$). Conclusion: This is the first study investigating il1b transcript profile in zebrafish brain after seizure insult. The results indicated that il1b transcript levels have a different pattern in zebrafish as compared to those described in rodent models of epilepsy. In order to better characterize the model for epilepsy studies, further studies are under way to investigate the il1b expression pattern at 0h, 6h and 72h after PTZ induced-seizure.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2374T

An integrative analysis pinpoints the pathogenesis of autism in the cerebellar vermis. *D.-A. Clevert, S. Hochreiter.* Institute of Bioinformatics, Johannes Kepler University Linz, Austria.

Motivation: We investigated neurodevelopmental dysfunctions in autism spectrum disorders (ASD) by an integrative analysis including the three largest genome-wide studies on associations between copy number aberrations (CNA) and ASD, the BioGPS tissue atlas, the Allen brain atlas, and in situ hybridization histochemistry data from the developing mouse brain. In contrast to the original association studies, we considered "ASD candidate genes" each of which is the only CNA-impaired gene in an ASD case, therefore, presumably causing ASD. For extracting ASD candidate genes, we developed an analysis pipeline for rare and small CNAs. Rare CNAs are supposed to be more disease-specific, because CNAs that cause ASD with high probability are assumed to be de novo and quickly vanish in the population due to their low reproductive fitness. Small CNAs affect only few genes and, therefore, are very specific concerning the genes they are impairing. **Results:** ASD candidate genes that are indentified independently in all CNA studies include the neurexins *CNTNAP2* and *NRXN1*, the catenin *CTNNA3*, the cadherin *CDH13*, and the contactins *CNTN5* and *CNTN6*. Gene set enrichment analysis of ASD candidate genes showed that significant biological processes are all related to cell and synaptic adhesion the postsynaptic density, membrane and synapse. At data from the BioGPS, the Cancer Genome Anatomy Project, and the Allen brain atlas, ASD candidate genes have significantly different variations in their expression values in cerebellum compared to other genes, where at the Allen brain atlas cerebellar vermis lobes I-II, IV, VI, VII, and VIII where most significant. In situ hybridization histochemistry data indicate that ASD candidate genes are primarily expressed in the developing mouse cerebellum. **Discussion:** Our results, that locate the pathogenesis of ASD in the cerebellar vermis, are consistent with pathological studies of ASD cases, where, in over 90% of the examined brains, well-defined cerebellar abnormalities were found. Also studies on children with vermal lesions showed phenotypes like speech disorders and behavioral disturbances similar to autism. The high percentage, 60–80%, of ASD cases showing motoric deficits again hints at the cerebellum. We explain 4:1 male to female ratio in ASD by the regulatory influence of estrogen on the development of the cerebellum. The human estrogen 17β-estradiol enhances in the cerebellum synaptic connectivity.

2375F

Common genetic variants in the GNB3, ARNTL, and NPAS2 loci are associated with objectively measured sleep traits in the elderly. *D.S. Evans[1], N. Parimi[1], C.M. Nievergelt[2], K.L. Stone[1], T. Blackwell[1], S. Redline[3], S. Ancoli-Israel[2], E.S. Orwoll[4], S.R. Cummings[1], G.J. Tranah[1], the Study of Osteoporotic Fractures (SOF) and the Osteoporotic Fractures in Men (MrOS) Study Groups.* 1) Research Institute, California Pacific Medical Center, San Francisco, CA; 2) Department of Psychiatry, University of California, San Diego, La Jolla, CA; 3) Department of Medicine, Brigham and Women's Hospital and Beth Deaconess Medical Center, Harvard Medical School, Boston, MA; 4) School of Medicine, Oregon Health and Science University, Portland, OR.

Aging is associated with changes in sleep quality and an increase in sleep disturbances, which in turn are associated with multiple adverse health outcomes. The internal biological clock controls many aspects of sleep and is regulated by a conserved set of clock genes. We tested the association between common genetic variation in clock genes and objectively measured actigraphic sleep and activity rhythm traits in two large population-based cohorts of elderly participants, the Study of Osteoporotic Fractures (SOF) (n=1407, 100% female, mean age 84 years) and the Osteoporotic Fractures in Men (MrOS) study (n=2527, 100% male, mean age 77 years). Common genetic variation in 30 candidate genes was captured using 529 single nucleotide polymorphisms (SNPs). Meta-analysis revealed that two SNPs in non-coding DNA (rs1047776 and rs2238114) in the GNB3 gene region were significantly associated with wake after sleep onset (rs1047776 $P_{ADD} = 2x10^{-5}$, rs2238114 $P_{ADD} = 5x10^{-5}$). SNP associations in the gene-rich GNB3 region could reflect the effect of variants in at least nine genes due to long-range linkage disequilibrium (LD). Publically available eQTL data from CEU lymphoblastoid cell lines were used to identify genes whose RNA expression might be regulated by rs1047776 and rs2238114. While GNB3 was the candidate gene in this region, neither SNP was associated with GNB3 RNA levels ($P > 0.05$). However, rs2238114 and rs1047776 were significantly associated with LRRC23 RNA levels ($P = 5x10^{-8}$, $P = 0.02$, respectively). LRRC23 is expressed in ocular tissue, and its leucine-rich repeat domain could be involved in protein interactions. In MrOS participants, SNPs within the binding partners ARNTL and NPAS2 were associated with a shift toward a later sleep and wake onset time (sleep onset time: ARNTL rs3816358 $P_{2DF} = 1x10^{-4}$, NPAS2 rs3768984 $P_{2DF} = 5x10^{-5}$; wake onset time: rs3816358 $P_{2DF} = 3x10^{-3}$, rs3768984 $P_{2DF} = 2x10^{-4}$) and the interaction between the two SNPs was significant (sleep onset time $P_{INT} = 0.003$, wake onset time $P_{INT} = 0.001$). In addition, results from a genome-wide association study of objective sleep traits will be discussed. These findings support the importance of clock genes in influencing the regulation of sleep traits in the elderly. The integration of gene expression data with SNP association results led to the serendipitous identification of LRRC23 as a potential candidate gene to regulate sleep fragmentation.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2376W

GWAS of alcohol abuse and dependence in the CATIE schizophrenia sample. *V.F. Gonçalves[1], C. Zai[1], A. Paterson[2], L. Sun[3], J.L. Kennedy[1], J. Knight[1].* 1) Neuroscience section, CAMH, Toronto, Canada; 2) Hospital for Sick Children, Toronto, Canada; 3) Statistics Division, Dalla Lana School for Public Health, Toronto, Canada.

Alcohol abuse or dependence (AB/AD) is one of the most frequent comorbidities diagnosed among schizophrenia (SCZ) patients. There is evidence that genetic factors influence AB/AD and the identification of genetic variants related to alcoholism in schizophrenia patients will provide us with better understanding concerning the biological mechanisms of addiction in psychiatric disorders. In this work, it was our objective to identify genetic variants associated with AB/AD in SCZ subjects. Lifetime alcohol abuse or dependence was diagnosed using the SCID research interview and the analyses were conducted considering the presence or absence of symptoms. After initial quality control of both SNPs and samples, data analysis was performed on 421,474 SNPs (Affy500K/Perlegen164K chip) in 391 case-control European-Americans (EA) and 213 case-control African-Americans (AA) from the NIH Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) trial. The association analyses were first performed separately for the EA and AA samples adjusting for the effects of sex, age, site, and the first two principal components. No SNPs achieved genome-wide significance level ($p<5x10(-8)$) in either sample. In the EA sample, the most significant SNP was rs12681165 (chr8) with p=2.6x10(-7). In the AA sample, the most significant SNP was rs7242277 (chr18) with p=2.6x 10(-6). This SNP is located close to the galanin receptor 1 (GALR1) gene, which has been associated with nicotine dependence in humans. Meta-analysis of these two genome-wide association studies was performed. The most interesting SNP was rs1248032 with p-value= 6.2x10(-6). This SNP is located close to the T-box (TBX5) gene. The genomic prioritization method (Sun et al., 2006) was used in both samples considering 1157 SNPs in 57 alcohol-related candidate genes. These SNPs were assigned to a high priority group and all remaining genome-wide SNPs of other genes assigned to a low priority group, and the stratified false discovery rate (SFDR) was applied to the two groups. However, no genome-wide significant association evidence was observed for the AB/AD phenotype in either sample. Genotyping of another schizophrenia sample (N=600) is currently being conducted, and this sample will be used for replication of the SNPs identified in this study. Alternative strategies under consideration include functional SNP analysis to maximize detection of true associations based on biological mechanisms.

2377T

Familial Study of Essential Tremor (FASET) at Columbia University. *N. Hernandez[1], D. Tirri[1], L. Clark[2,3], E. Louis[1,3,4,5].* 1) Gertrude H. Sergievsky Center, Columbia University, New York, NY; 2) Department of Patholgoy and Cell Biology, Columbia University, New York, NY; 3) Taub Institute for Research on Alzheimer's Disease and the Aging Brain, Columbia University, New York, NY; 4) Department of Epidemiology, Mailman Public School of Health, Columbia University, New York, NY; 5) Department of Neurology, Columbia University, New York, NY.

Essential tremor (ET) is one of the most common movement disorders, with prevalence estimated at 4.6% for individuals age 65 years and older. Family studies in addition to twin studies have provided strong evidence for a genetic contribution to ET. To date, linkage studies have identified three susceptibility loci at chromosomes 3q13 (ETM1; OMIM:190300), 2p22–25 (ETM2; OMIM:602134) and 6p23 (ETM3; OMIM:611456) however the genes and causal mutations have yet to be identified. The FASET (Familial study of Essential Tremor) study (NS073872) at Columbia University was initiated to identify and recruit ET families into a genetic study of ET to determine modes of inheritance and identify susceptibility genes. In year 1 of this project, between the period of 10/10/2011 and 4/2/2012 we have conducted a total of 98 in person assessments (home visits and videotape), of which 25 subjects are probands and and 73 are relatives. Total of 90 blood samples and 8 cheek swabs have been sent to the Human Genetics Resource Core at Columbia University for DNA isolation. All subjects recruited into the study were evaluated using the following assessments: 1) a structured clinical questionnaire that collects demographic data and medical data 2) a standardized 35–minute videotaped examination that includes a 20 minute tremor examination (which allows for the detailed assessment of postural, kinetic, intention and rest tremors, the motor portion of the Unified Parkinson's Disease Rating Scale and a detailed dystonia examination 3) Phlebotomy (3 tubes of blood, 8–10ml/tube per subject). All videotaped examinations were reviewed by a neurologist (Dr Louis) while blind to the pedigree structure. Overall, a total of 10 families and 45 family members have completed interviews and assessments. In year 1 of NS073872 we are proceeding with whole genome sequencing in six of the completed families. We have selected the two most distantly related affected subjects to prioritize for sequencing. We will present the pedigree structures of the 10 completed families and preliminary sequencing data. As of May 2012, causal genes for ET have not been identified. The identification of causal genes will lead to an understanding of ET pathogenesis that will allow the development of treatments. Funding: NIH, NINDS #R01 NS073872 (Louis, Clark), the Parkinson's Disease Foundation (Louis, Clark), the Arlene Bronstein Essential Tremor Research Fund, and the Claire O'Neil Essential Tremor Research Fund.

2378F

Significant Linkage of Familial Essential Tremor to a novel locus on Chromosome 5. *J.E. Hicks, I. Konidari, J.R. Gilbert, W.K. Scott.* Hussman Institute for Human Genomics, University of Miami, USA.

Essential tremor is a neurological condition characterized by tremor during voluntary movement. To date, three loci causative of familial essential tremor have been identified. We examined the five largest ET families in our 49–family dataset for genetic linkage using a 440,512 SNP Affymetrix Axiom array. 122 individuals (48 with ET; average 9.6 per family) from these five families were studied. Multipoint parametric linkage analysis was performed with a subset of 6,384 SNPs using an affecteds-only dominant model in SIMWALK2. Individuals were considered to be affected with definite ET if they met the consensus criteria developed by the Movement Disorder Society. For one family, linkage analysis identified two regions providing significant evidence of linkage to ET. The strongest of these was on chromosome 5 (MLOD score 3.6, 1–LOD down support interval: 168,502,329–174,594,583, hg19 coordinates). The second was on chromosome 18 (max LOD score 3.2, 1–LOD down support interval: 2,122,401–3,475,401). In order to fully utilize genotype information from all SNPs, identity by descent (IBD) sharing analysis was employed. The BEAGLE and GERMLINE software packages were used to assess IBD sharing in all affected relative pairs in each family. Regions were considered candidate regions if all affected individual pairs were at least IBD=1 for SNPs in the region. This analysis excluded the linkage peak on chromosome 18, as there was not sharing between all individual pairs. Segment sharing under the linkage signal on chromosome 5 reduced the candidate region by 1.3Mb to a 1024 SNP haplotype in the region 169,267,941–174,050,867. This region contains 21 genes, several of which are excellent biological candidate genes. While other families did not achieve significant evidence for linkage, maximum LOD scores corresponded to haplotypes shared IBD by all affected individuals in two more families, one family at 5q14.3, and the other family at 1p34.3 and 20p12.1. This indicates that while linkage may not be able to reach traditional significance thresholds, IBD haplotype sharing may be able to localize disease loci in these families.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2379W

Autism and Obesity: Is There Any Mechanistic Connection? *Z. Talebizadeh[1,2], M. Miralles[1].* 1) Pediatrics, Children's Mercy Hosp, Kansas City, MO; 2) University of Missouri-Kansas City School of Medicine, Kansas City, MO.

Autism might be accompanied with other symptoms. These associated symptoms have been used in subject stratification for linkage and association studies resulting in valuable discoveries. Herein, the central point of our model is based on the concept that obesity might be a co-occurring condition in autism. Recently, there have been suggestive reports for a higher rate of obesity in autism. It is not clear if the suggested higher rate of obesity is related to psychotropic medication side effects or to the disorder itself. However, in-depth literatures search suggests that indeed some scientific merit can be found that backs up the potential mechanistic link between autism and obesity. Implications in some metabolic pathways have been independently associated with autism and obesity. To assess the possible relationship between autism and obesity, we evaluated the expression level of genes in a given metabolic pathway in autistic subjects with and without obesity. Subjects were ascertained from the Autism Genetics Resource Exchange (AGRE). Obesity status was determined using Body Mass Index percentiles. PCR arrays were conducted on lymphoblastoid cell line-derived RNA from 16 AGRE subjects. Despite the small sample size in this pilot study, differential expression was detected for several genes in autism-obese group compared with autism-non obese group. Interestingly, two genes with previously reported connections to both autism and fat metabolism were seen among the differentially expressed genes in the obese group. Our pilot study demonstrates that autistic subjects with obesity may represent a more homogeneous subset of this highly heterogeneous population. Applying such a stratification method may provide a way to better understand underlying genetic mechanism in a subset of autistic subjects with a co-occurring obesity. Our study introduces a novel, practical, and effective method to connect different lines of phenotypic and genomic data in unraveling the etiology of this highly heterogeneous neurodevelopmental disorder.

2380T

Haplotypes in the expression quantitative trait locus of interleukin-1β gene are associated with schizophrenia. *M. Yoshida[1], K. Shiroiwa[1], K. Mouri[1], T. Sasada[1], H. Ishiguro[2,3], T. Inada[4], T. Arinami[3], O. Shirakawa[5], A. Hishimoto[1].* 1) Department of Psychiatry, Kobe University Graduate School of Medicine, kobe, Hyogo-ken, Japan; 2) Department of Neuropsychiatry and Clinical Ethics, University of Yamanashi School of Medicine, Kofu, Japan; 3) Department of Medical Genetics, Doctoral Program in Social and Environmental Medicine, Graduate School of Comprehensive Human Sciences, University of Tsukuba, Tsukuba, Japan; 4) Department of Psychiatry, Seiwa Hospital, Institute of Neuropsychiatry, Tokyo, Japan; 5) Department of Neuropsychiatry, Kinki University School of Medicine, Osaka, Japan.

Recent genome-wide association studies (GWAS) and gene expression analyses have revealed that single nucleotide polymorphisms (SNPs) associated with complex diseases such as schizophrenia are significantly more likely to be associated with expression quantitative trait loci (eQTL). The interleukin-1β (IL1B) gene has been strongly implicated in the susceptibility to schizophrenia. In order to test this association, we selected five tag SNPs in the eQTL of the IL1B gene and conducted a case-control study using two independent samples. The first sample was comprised of 528 schizophrenic patients and 709 controls and the second sample was comprised of 576 schizophrenic patients and 768 controls. We identified two SNPs and several haplotypes as being significantly associated with schizophrenia. Previous reports indicated that one major haplotype that was protective against schizophrenia reduced IL1B transcription, while two risk haplotypes for schizophrenia enhanced IL1B transcription. Therefore, we measured IL1B mRNA expression in PAXgene-stabilized whole blood from 40 schizophrenic patients and 40 controls to explore the possibility of using five tag SNPs as schizophrenic trait markers. A multiple regression analysis taking confounding factors into account revealed that the T allele of rs4848306 SNP, which is a protective allele for schizophrenia, predicted reduced change in IL1B mRNA expression, regardless of phenotype. Our results appear to support the previous hypothesis that IL1B contributes to the genetic risk of schizophrenia and warrant further research on the association of eQTL SNPs with schizophrenia.

2381F

Selection of Candidate Genes from Linkage Regions Identified by the Individual Behavioral Domains of Autism Spectrum Disorders. *A. Hare[1], J. Flax[1], Z. Fermano[1], M. Azaro[1], A. Seto[1], S.Y. Cheong[2], S. Buyske[3], B. Zimmerman-Bier[4], C.W. Bartlett[2], L.M. Brzustowicz[1].* 1) Rutgers University, Department of Genetics, Piscataway, NJ; 2) The Batelle Center for Mathematical Medicine, The Research Institute at Nationwide Children's Hospital and Department of Pediatrics, The Ohio State University, Columbus, OH; 3) Rutgers University, Department of Statistics and Biostatistics, Piscataway, NJ; 4) Saint Peter's University Hospital, New Brunswick, NJ.

Autism is a complex disorder defined by symptoms in three domains: 1) Communication Deficits, 2) Social Deficits, and 3) Restricted and Repetitive Behaviors. The New Jersey Language and Autism Genetics Study (NJLAGS) has studied families who have at least one individual with a diagnosis of Autism and at least one individual with Specific Language Impairment (SLI) with the goal of identifying behavioral biomarkers related to language as well as to other conditions associated with autism. Behavioral data assessing language impairment (LI), reading impairment (RI), social impairment (SI), and obsessive-compulsive behaviors (OCB) has been collected for these families, allowing the identification of behavioral phenotypes for use in linkage analyses. Linkage regions have been identified on 13q21.2 (OCB), 14q32.31 (SI), 15q23 (LI), 15q26.2 (SI), and 16p12.3 (RI). With no compelling evidence for association in these regions, we used Ingenuity Pathway Analysis (IPA) to aid in the selection of candidate genes for fine mapping analysis. A list of genes for each linkage region was identified (RefSeq, NCBI Build 37) and biological functions unique to each list were identified using IPA's core analysis. Linkage regions were grouped by domain and were analyzed in IPA separately. Genes playing a role in nervous system development and function were identified and given a score determined by the number of relevant functions. All domains were analyzed and scored together in order to mimic an overall autism model. Scores from the autism model were added to scores obtained in the individual domains analysis. Then candidate genes were prioritized by overall score. The LI domain identified 12 candidate genes of interest, the SI domain identified 3 genes of interest and the OCB domain identified 1 gene of interest. The highest overall scoring candidate gene, AKT1, was identified in the SI domain. AKT1 was identified in IPA's “PTEN Signaling”, “Reeling Signaling in Neurons”, and “Axonal Guidance Signaling” canonical pathways. Other high-scoring candidates in these pathways include NTRK3 (LI), IGF1R (SI), and PPP2R5C (SI). Inclusion of these genes in the overall autism model strengthened their overall score and increased support for roles in “Axonal Guidance Signaling” and “mTOR Signaling.” This method of ranking, in conjunction with other databases such as DAVID and KEGG, has identified strong candidate genes for future fine mapping analyses.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2382W

Association of Neuregulin 1 polymorphisms with a north Indian schizophrenia cohort. *P. Kukshal[1,2], T. Bhatia[3], A.M. Bhagwat[2], R.E. Gur[4], R.C. Gur[4], S.N. Despande[3], V.L. Nimgaonkar[5], B.K. Thelma[1].* 1) Department of Genetics, University of Delhi, South Campus, New Delhi, India; 2) C. B. Patel Research Centre, Vile Parle (West), Mumbai, India; 3) Department of Psychiatry, Dr. RML Hospital, New Delhi -110 001, India; 4) The Schizophrenia Research Center, Department of Psychiatry, University of Pennsylvania School of Medicine, Philadelphia, PA, USA; 5) Departments of Psychiatry and Human Genetics, Western Psychiatric Institute and Clinic, University of Pittsburgh School of Medicine and Graduate School of Public Health, 3811 O'Hara Street, Pittsburgh, PA 15213, USA.

Background: Neuregulin 1 (*NRG1*) plays important roles in the development and plasticity of the brain. *NRG1* polymorphisms have been proposed as risk factors for several common disorders and most strikingly with Schizophrenia. Genetic variation in *NRG1* has been associated with schizophrenia in several studies across ethnic groups and its association with cognition is also well documented. Methods: We tested for association of *NRG1* variants with schizophrenia in a moderately large north Indian cohort (n= 1020 cases and controls each). Cases diagnosed based on DSM IV criteria were age and gender matched to controls from the same geographical region. Informed consent was taken from each participant. On a subset of the cohort (n=116 cases and n=170 controls) a computerized neuropsychological battery was administered and their correlation with neurocognitive measures was evaluated. In the first step, a total of 35 SNPs using iPLEX and three microsatellite markers by fragment analysis were tested for association with Schizophrenia. In the second step, regression analysis for eight neurocognitive parameters was carried out with only SNPs showing allelic association. Results: Five variants rs35753505 (p=0.04; OR=1.147(1.01–1.31)), rs4733263 (p=0.04; OR=1.14(1.01–1.31)), rs4281084 (p=0.05; OR= 0.86(0.75–1.0)), rs6994992 (p=0.026; OR=1.15(1.02–1.3)) and rs10093107 (p=0.05; OR=0.89(0.78–1.0)) and one microsatellite marker 420 M9–1395 (p=0.016) from the 5′ region showed association with the disease but none withstood bonferroni corrections. 5 SNPs namely rs35753505 (p=0.05); rs1354336 (p=0.01); rs10093107 (p=0.03); rs3924999 (p=0.02) and rs11780123 (p=0.006) showed genotypic association. A sliding window haplotype revealed associations in and around rs10093107 in the 3–10 marker haplotypes from the 5′ UTR. On regression analysis of neurocognitive measures associations were observed with emotion and attention. Conclusions: Considering the isoforms and the multiple roles and functions that *NRG1* is involved in, it is believed that regulatory variants in this gene may be responsible in the disease etiology. SNPs from the *NRG1* promoter region are associated with schizophrenia in this large north Indian cohort. Association reported in other populations for rs6994992 and rs35753505 was also replicated. Association with cognitive measures was also observed. These results warrant fine mapping of this gene to identify critical variants for schizophrenia.

2383T

Single nucleotide polymorphism (SNP) fine mapping of the *AVPR1A* and *OXTR* loci reveals an *OXTR* association with a social memory endophenotype in probands with autism and their family members. *A. Lori[1], I. Lee[2], J.F. Cubells[1,3], E.B. Binder[1,4], T. Lehtimäki[5], K. Puura[6], K.N. Conneely[1], L.J. Young[3], D. Skuse[2].* 1) Department of Human Genetics, Emory University School of Medicine; Atlanta Georgia, GA, USA; 2) Behavioral and Brain Sciences Unit, UCL Institute of Child Health; London, UK; 3) Departments of Psychiatry and Behavioral Sciences, Emory University School of Medicine; Atlanta GA, USA; 4) Max-Planck Institute of Psychiatry; Munich Germany; 5) Department of Clinical Chemistry, Tampere University and University Hospital; Tampere, Finland; 6) Department of Child Psychiatry, Tampere University and University Hospital; Tampere, Finland.

Autism spectrum disorders (ASDs) are neurodevelopmental syndromes characterized by deficits in social behaviors and communication, and by idiosyncratic interests and repetitive behaviors (DSM-IVTR). Since the ASD diagnosis encompasses a very heterogeneous set of behavioral phenotypes, endophenotypes related to specific aspects of human social cognition could be potentially superior targets for ASD-related genetic analysis. Vasopressin (AVP), oxytocin (OXT) and their receptors play important roles in evolutionarily conserved social behaviors such as social memory and social attachment. Variants in the genes *AVPR1A* and *OXTR*, encoding respectively the AVP-1a receptor and the OXT receptor, have been reported to associate with ASD. In this study, we evaluated whether single nucleotide polymorphisms (SNPs) at *AVPR1A* and *OXTR* loci were associated with variation in human social cognition by evaluating 4 specific cognitive tasks (face recognition memory, accuracy in detecting direction of eye gaze, recognition of facial emotion of fear, Autism Diagnostic Observation Schedule subscale scores) in 190 families with a single offspring affected by ASD. Criteria for SNP selection included a minor allele frequency of >0.01 and pairwise r2≤ 0.85 among validated SNPs in HAPMAP and SNP Browser. A total of 30 SNPs for *AVPR1A* and 59 SNPs for *OXTR* were successfully genotyped using a Sequenom platform. Genetic associations were tested using QFAM-PLINK for quantitative familial traits (BC-Gene). Correction for multiple testing was performed using p-ACT (Conneely and Boehnke, 2007, AJHG 81:1158–68). One SNP (rs237887) in *OXTR* showed an experiment-wide significant association with face recognition memory (uncorrected p=0.00015, p-ACT= 0.023). This intronic SNP is located in a region predicted to interact with several transcription factors, suggesting a possible functional effect of the SNP on OXTR gene expression. We detected an experiment-wide significant association between a potentially functional SNP in *OXTR* and face recognition memory ability in children with ASD and their family members. The results add to evidence for a role of the OXT system in influencing key components of human social behavior, and suggest for the first time that variation in *OXTR* associates with variation in specific components of social cognition in humans. Support: NIH grant MH056897 to LJY; Wellcome Trust, Nancy Lurie Marks Family Foundation National Alliance for Autism Research to DS.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2384F

The Effects of *ABCB1* Polymorphisms and Expression in Perinatal Depression. *S.E. Parets[1], A. Lori[2], M.N. To[3], D.J. Newport[3], Z.N. Stowe[4], J.F. Cubells[1,2,3], A.K. Smith[1,2].* 1) Genetics and Molecular Biology Program, Emory University, Atlanta, GA; 2) Department of Human Genetics, Emory University, Atlanta, GA; 3) Department of Psychiatry& Behavioral Sciences, Emory University School of Medicine, Atlanta, GA; 4) Psychiatric Research Institute, University of Arkansas for Medical Sciences, Little Rock, AR.

High levels of maternal depressive symptoms have been widely associated with poor infant outcome and provide compelling evidence to continue antidepressant treatment throughout pregnancy. *ABCB1* encodes P-glycoprotein (P-gp), which is integral in controlling the passage of substances, including cortisol and antidepressants, across the blood-brain and placental barriers. Polymorphisms in *ABCB1* account for differences in the rates of substrate transport and associate with depressive symptoms and treatment response. For this study, we recruited 186 women in their 3rd trimester undergoing pharmacotherapy with antidepressants. We then examined the association between maternal polymorphisms in *ABCB1* and the maximum 3rd trimester score on the Hamilton Rating Scale for Depression (HRSD) using linear regression. Associations were observed with 3 exonic SNPs: rs1128503, rs2032582 and rs1045642 ($p<.0001$). Since they are in LD, we examined common haplotypes (≥5%) and observed that the G-G-C haplotype (53.9%) was associated with higher maternal depression ($p=6x10^{-5}$). We next examined the association between neonatal (umbilical cord blood; N=78) genotypes and birth weight using a linear regression model that accounted for the level of antidepressant in the cord blood and its relative affinity for the P-gp transporter. Associations between the neonatal genotypes in 3 SNPs (rs2032582, rs2032583 and rs2235040) and high birth weight were observed ($p<.05$). The G allele of rs2032582 was associated with both higher maternal depression and higher neonatal birth weight, which has been associated with poor infant outcome. This polymorphism alters protein activity/substrate specificity. We next evaluated placental *ABCB1* expression (N=30) and observed association with maximum maternal HRSD score during pregnancy ($p=.02$) and maternal perceived stress during pregnancy ($p=.012$). Interestingly, *ABCB1* expression also increased proportional to gestational age ($t=2.58$; $p=.016$) and APGAR scores ($t=2.48$; $p=.02$) at delivery. Our results suggest that maternal depressive symptoms and stress may be mediated, in part, by genetic variation in *ABCB1*. Expression changes in this gene correlate with developmental outcomes in the offspring and support further investigation of *ABCB1* in perinatal depression.

2385W

HMNC1 and COMT gene polymorphism associated with postpartum depression. *M. Romano-Silva, D.M. Miranda, P. Figueira, H. Correa, A. Alvim-Soares.* INCT de Medicina Molecular, Faculdade de Medicina, Universidade Federal de Minas Gerais, Av Alfredo Balena, 190, Belo Horizonte-MG, CEP 30130–100, Brazil.

Postpartum depression is a common disorder affecting women, with consequences for both mother and child. A genetic determinant for PPD is suggested by many genetic epidemiologic studies. One hundred and seventeen women were randomly selected among those who delivered in a maternity and filled in and filled in the Edinburgh Postpartum Depression Scale (EPDS) questionnaire about 8 weeks after delivery. POMC, MC4–R, HMCN-1 and COMT polymorphisms were analysed by PCR. Differences in genotype frequency were calculated by X2 test and the difference between groups was tested with Student's t test. Tests were two-tailed and results significant when $p \leq 0.05$. Those women that scored 13 or greater on EPDS were included in the PPD group [25 (18,97%)], and a total of 94 woman (81,03%), who scored 12 or less, became the Control group. No differences in POMC and MC4R genotype distribution were observed between these groups. We found a significant interaction between the development of depressive symptoms in postpartum and polymorphisms in HMCN-1 ($p=0.03$) and COMT ($p=0,01$). Hemicentin-1 (HMNC-1) gene encodes an extracellular protein that contains four estrogen receptor-binding sites and is involved mainly in the cell migration and protein anchorage. In conclusion, this study supports the notion that PPD is a result of complex interactions between multiple variables, including genetic factors. It is encouraged independent replications in larger samples and further investigation on these gene effects in PPD.

2386T

Disrupted in Schizophrenia 1 gene in autist patients. *P. Sarda[1], J. Puechberty[1], M. Girard[1], A. Schneider[1], S. Taviaux[1], G. Iefort[1], D. Genevieve[1], A. Baghdadli[2].* 1) Department of medical genetics, Arnaud de Villeneuve Hospital, Montpelier, France; 2) Child psychiatry department, Center of resource autism, Montpellier, France.

Disrupted in Schizophrenia 1 (DISC1) is a key susceptibility gene for major mental disorders particularly for psychiatric disorders (schizophrenia, depression, bipolar disorder and schizoaffective disorder) In two previous Finnish and Chinese studies, DISC1 polymorphisms were associated with autism and Asperger syndrome. Two reports have revealed abnormalities of DISC1 in three individuals with autism spectrum disorder. The first report concerns two brothers with autism and mild mental retardation with a large 2.07 Mb duplication on 1q42.2 including the DISC1 gene and inherited through the normal father. The second report concerns a 3–year-old male with developmental delay and autistic behavior who has a large 2Mb interstitial deletion of chromosome 1q42 involving DISC1, DISC2, and TSNAX inherited from his non-affected mother. We report the case of a 4–year-old boy with mild developmental delay and autistic behavior. Molecular karyotyping using SNP 6.0 Affymetrix arrays revealed a 100.7 kb duplication in the first exon of the DISC1 gene. This 100.7 kb duplication was inherited through the father. Anomalies were confirmed by qPCR analysis. The father presented some difficulties in social interaction and communication during childhood. He obtained a secondary degree but had initial difficulties and repeated 2 years during grade school and high school. These observations confirm evidence for a potential role of the DISC1 gene in autism spectrum disorders. This work was funded by the “Programme Hospitalier de Recherche Clinique régional Languedoc-Roussillon”.

2387F

Phenotypic description of a child with a 10q23.1 deletion involving only *NRG3*. *N. Sobreira[1], D. Avramopoulos[2], A. Pulver[2], B. Menten[4], B. Loeys[3], D. Valle[1].* 1) Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD; 2) Department of Psychiatry, Johns Hopkins University School of Medicine, Baltimore, MD; 3) Center for Medical Genetics, Antwerp University Hospital and University of Antwerp, B-2650 Antwerp, Belgium; 4) Center for Medical Genetics, Ghent University Hospital, Ghent, Belgium.

Our group and others have reported deletions in the 10q22q23 region typically ranging from a 4.7 Mb to 7.6 Mb region between 2 low copy repeats. The phenotype of the patients with the deletions is variable and includes developmental delay, macrocephaly, broad forehead, deep-set eyes, upslanting palpebral fissures, smooth philtrum, thin upper lip, cerebellar anomalies, cardiac defects and congenital breast aplasia. Balciuniene et al. (2007) suggested that the 10q22q23 region harbors one or more genes important for cognitive and behavioral development and Chen et al. (2009) showed that *NRG3*, deleted in most of the 10q22q23 deletions, is involved in schizophrenia. Other frequently involved 10q22q23 genes include *BMPR1* associated with juvenile polyposis and *PTEN* associated with Cowden syndrome. We describe a 5 year-old male with short stature, weight < 3rd percentile, microcephaly, brachycephaly, thin upper lip, smooth philtrum, low-set ears, hypospadius, inguinal hernia, sagittal clefts of the thoracic vertebrae, and developmental delay. The patient had a minimum 436 Kb deletion (chr10:83,652,547–84,088,659) by SNP genotyping array limited to the 5′ region of *NRG3*. To determine the precise deletion size and any related chromosomal aberrations we included a 851 Kb region of chr10 (chr10:83,375,977–84,227,714) in a 4.6 Mb targeted capture kit that included other unrelated chromosomal rearrangements and followed with next-generation sequencing. We analyzed the sequencing results using the Integrative Genomics Viewer and by comparing the coverage data between our patient and 5 controls. We found that the breakpoints of the approximately 570 Kb deletion in the telomeric end of the deletion is in a 10 Kb region of LINE repeats belonging to the L1 family (L1PA2) and that the centromeric end of the deletion is in a 6 Kb region of LINE repeats belonging to the family L1 (L1PA2 and L1M5). The detailed characterization of this smaller deletion and of the phenotype of our patient suggests the association of *NRG3* disruption with common features described in patients with larger 10q22q23 deletions including developmental delay, smooth philtrum, and thin upper lip, and with less frequent features like short stature, low weight, microcephaly, brachycephaly, hypospadius, inguinal hernia, and sagittal clefts of the thoracic vertebrae. Our results make *NRG3* a strong candidate gene for the developmental delay and dysmorphic features associated with the 10q22q23 deletions.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2388W

Determination of a rare variant and common variants in COMT for their involvement in the etiology of smoking dependence using Mid-South Tobacco case-control study samples. *J. Yang[1], S. Wang[1], J.Z. Ma[2], T.J. Payne[3], M.D. Li[1].* 1) Psychiatry and Neurobehavioral Sciences, University of Virginia, Charlottesville, VA; 2) Public Health Science, University of Virginia, Charlottesville, VA; 3) Otolaryngology and Communicative Sciences, University of Mississippi Medical Center, Jackson, MS.

Catechol-O-methyltransferase (COMT) is responsible for degrading catecholamines, including the neurotransmitters dopamine, epinephrine, and norepinephrine. Given the role of the dopamine reward pathway in nicotine dependence (ND), in a previous study, we showed that several common variants and haplotypes within the COMT gene are significantly associated with smoking quantity, the Heaviness of Smoking Index, and the Fagerström Test for ND score (FTND) in both European Americans (EA) and African Americans (AA) of the Mid-South Tobacco Family (MSTF) sample. To confirm the involvement of COMT in ND, we performed a case-control-based association analysis in an independent sample, called the Mid-South Tobacco Case-Control study, which consists of 1387 EA subjects and 3161 AA subjects recruited primarily from the state of Mississippi since 2005. The SNPs selected were rs6269, rs4633, rs5031015, rs4818, rs4680, and rs174699, where rs4680 represents a Val/Met missense polymorphism and rs503105 is a rare variant located in exon 4 of the membrane-bound COMT transcript or exon 2 of the soluble COMT transcript that leads a non-synonymous change of G to A (Ala/Thr). Our analysis revealed a significant association of rs4680 (Val/Met) with FTND in the EA (P = 0.00057) and pooled (P = 0.00223) samples and of rs503105 (Ala/Thr) with FTND in all samples (P = 0.00072 for the AA, P = 0.00385 for the EA, and P = 5.2E-08 for the pooled samples). We further found that the A allele of rare variant rs5031015 represents a significant risk effect in the AA (frequency 0.19%, beta = -1.583), EA (frequency 0.4%, beta = -1.945), and pooled (frequency 0.25%, beta = -1.835) samples. In addition, we identified a haplotype, A-T-G-C-A-T, formed by SNPs rs6269, rs4633, rs5031015, rs4818, rs4680, and rs174699, with a significant protective effect against ND in the EA sample (frequency 49.1%, P = 0.00037). In sum, our results not only provide evidence for a significant association of COMT variants with ND in an independent sample but also demonstrate that the rare variant rs5031015 likely plays a significant role in susceptibility to ND. (Supported by NIH grant DA-012844).

2389T

Copy number variations (CNV) and agenesis of the corpus callosum, narrowing the recurrent 8p duplication interval. *S. Esmaeeli Nieh[1], L. Fernandez[1], S. Sajan[2], E. Rider[1], P. Bukshpun[1], M. Wakahiro[1], J.T. Glessner[3], H. Hakonarson[3], S.L. Christian[2], W.B. Dobyns[2], E.H. Sherr[1].* 1) Department of Neurology, University of California, San Francisco, San Francisco, CA; 2) Center for Integrative Brain Research, University of Washington, Seattle, WA;; 3) Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, PA.

Background: Because of low copy repeats (LCRs) and common inversion polymorphisms, the human chromosome 8p is prone to a recurrent cytogenetically visible deletion/duplication rearrangement which is often seen with the brain malformation agenesis of the corpus callosum (AgCC). AgCC is often associated with intellectual disability and autistic features. The genetics of AgCC is complex, however, prior reports from our lab and others have suggested that de novo duplications on 8p may play an important causative role. **Hypotheses:** We hypothesize that individuals with brain malformations have a high rate of de novo copy number variations (CNV) including the 8p22–p11.21 recurrent event and that screening a large cohort might identify smaller CNVs that narrow the critical region. **Methods:** We identified patients with AgCC by a comprehensive radiological and clinical review of our IRB approved AgCC cohort (n>500). Blood samples were obtained from the proband and both biological parents, when available. DNA was extracted and run on an Illumina 610 Quad Chip array. Data were analyzed using PennCNV. CNV's that were selected for analysis contained greater than 9 SNPs, were longer than 30kb, had a PennCNV confidence level above 10 and contained one or more genes. Manual curtain, merging of contiguous large CNVs and PCR confirmation were conducted. **Results:** From a total of 275 AgCC proband arrays, we identified 5 individuals with inverted duplication on 8p22–p11.21, all with similar genomic boundaries. We also identified two individuals who had isolated duplications in this interval. Interestingly, these two patients had complete callosal agenesis and microcephaly, but did not have cardiac defects, a feature common to the larger genomic rearrangement. This finding defines a narrowed critical interval ~10.5Mb on 8p21.2–p12. This region encompasses ~80 genes including several strong candidate genes associated with neurological disorders. Functional interpretation of non-coding regions using Great (version 2.0.1) revealed the highest Binom Rank with a raw P-value=3.1e-6 for central nervous system disease. **Conclusions:** Duplications on 8p constituted the most recurrent genomic regions identified in AgCC patients in our cohort and there is a specific enrichment of neurodevelopmental disorders' genes. These data support a strong etiologic link between AgCC and neurodevelopmental disorders, likely shared molecular and developmental pathways.

2390F

Serotonin Transporter and Receptor Genes Significantly Impact Nicotine Dependence through Genetic Interactions in both European American and African American Smokers. *Z. Yang[1,2], C. Seneviratne[1], S. Wang[1], J.Z. Ma[3], T.J. Payne[4], J. Wang[2], M.D. Li[1].* 1) Psychiatry and Neurobehavioral Sciences, University of Virginia, Charlottesville, VA; 2) Shanxi Key Laboratory of Environmental Veterinary Science, Shanxi Agricultural University, Shanxi, China; 3) Public Health Sciences, University of Virginia, Charlottesville, VA; 4) Otolaryngology and Communicative Sciences, University of Mississippi Medical Center, Jackson, MS.

Pharmacologic studies have demonstrated a modulatory effect of nicotine on serotonergic signaling, primarily through genes encoding the serotonin transporter (SLC6A4) and the 5–HT3AB subunits HTR3A and HTR3B. Thus, the functional state of these genes may alter the susceptibility to nicotine dependence (ND). In this study, we examined the impact of variations in the three genes on ND in 1366 individuals from 402 African American (AA) and 671 individuals from 200 European American (EA) nuclear families. The ND of each smoker was assessed with three commonly used measures: smoking quantity (SQ), Heaviness of Smoking Index (HSI), and Fagerström Test for Nicotine dependence (FTND). Association analysis at the individual SNP level revealed marginal association of rs10160548 in HTR3A with SQ and HSI in AA, 5–HTTLPR in SLC6A4 with FTND in EA, and rs11606194 in HTR3B with FTND in the pooled sample. Additional haplotype-based association analysis revealed a few major haplotypes in HTR3A that were significantly associated with ND in the AA, EA, and pooled samples. However, none of these associations remained significant after correcting for multiple testing except for a haplotype G-C-C-T-A-T formed by SNPs rs1150226, rs1062613, rs33940208, rs1985242, rs2276302, and rs10160548 in HTR3A for the AA sample Finally, we demonstrated significant interactions among variants rs1062613 and rs10160548 in HTR3A, rs1176744 in HTR3B, and 5–HTTLPR and rs1042173 in SLC6A4, affecting all three ND measures in the AA, EA, and pooled samples. Of these significantly interacting variants, rs1062613 in HTR3A and 5–HTTLPR and rs1042173 in SLC6A4 are known to modulate expression of the corresponding gene and rs1176744 in HTR3B to alter serotonin conductance by 5–HT3AB. Serotonergic signaling clearly plays a significant role in ND through interaction among variants in the genes encoding the serotonin transporter and receptors.

2391W

Genetic Association of *FKBP5* and *CRHR1* with Cortisol Response to Acute Psychosocial Stress in Healthy Adults. *P.B. Mahon[1,2], P.P. Zandi[1,2,3], J.B. Potash[4], G. Nestadt[2], G.S. Wand[1,5].* 1) Johns Hopkins Mood Disorder Center, Baltimore, MD; 2) Department of Psychiatry, Johns Hopkins School of Medicine, Baltimore, MD; 3) Department of Mental Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD; 4) Department of Psychiatry, University of Iowa, Iowa City, IA; 5) Department of Medicine, Johns Hopkins School of Medicine, Baltimore, MD.

Chronic dysregulation of the hypothalamus-pituitary-adrenal (HPA) axis activity is related to the onset and course of several neuropsychiatric and stress-related disorders. Thus, it is important to identify factors that account for the observed inter-individual variability in HPA axis activity. Studies suggest that cortisol response to stress has a moderate to strong genetic etiology and that FKBP5 and CRHR1 are key proteins regulating the HPA axis response to stress. Polymorphisms in the *FKBP5* and *CRHR1* genes have been associated with mood disorders. The goal of the present study was to examine the relationship between polymorphisms in these genes with HPA axis response to psychological stress in one of the largest Trier Social Stress Test (TSST) cohorts yet examined. Our sample included 368 healthy, young adults who underwent the TSST. Participants were carefully and extensively screened to include only healthy individuals in order to minimize state effects (smoking, alcohol use, medical conditions, etc.) in favor of examining genetic effects on cortisol stress response. Salivary cortisol was measured at multiple time points both before and after the stressor. After quality control, nine SNPs in *FKBP5* and four in *CRHR1* were available for analysis. Single marker and haplotypic analyses were carried out in PLINK and STATA 11.0. Exploratory analyses were also conducted using six stress-related variables as effect modifiers or mediators of the association. The strongest evidence was for an association of rs4713902 and rs7757037 in *FKBP5* with baseline cortisol response to the TSST under a recessive model (p=0.0003 and 0.004, respectively). In addition, we identified a sex-specific effect of three SNPs in *FKBP5* on peak response to the TSST (interaction p=0.02–0.009). The four SNPs in *CRHR1* (rs7209436, rs110402, rs242924, rs242938) were associated with peak cortisol response to the TSST under a recessive model (p=0.04–0.007). We have provided additional evidence for association of variants in *FKBP5* and*CRHR1* with HPA axis response to psychosocial stress. These findings have implications for interindividual variation in HPA axis activity and potentially for the etiology of neuropsychiatric and stress-related disorders.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2392T

Associations of Genetic Polymorphisms in *GRIN2B* and *GABRG2* with Epilepsy Susceptibility. *C. Chen[1], C. Hung[1], Y. Kao[1], H. Liou[2].* 1) Department of Pharmacy, College of Pharmacy, China Medical University, Taichung, Taiwan; 2) Pharmacology and Neurology, National Taiwan University Hospital and College of Medicine, National Taiwan University, Taipei, Taiwan, Republic of China.

Purpose: Epilepsy is one of the most common chronic neurological disorders. Alterations of neurotransmitter receptors may be associated with epilepsy susceptibility. GABA and glutamate are the two most widely discussed neurotransmitters which may play a major role in epilepsy. The aim of this study was to investigate the possible role of polymorphisms in GABA and glutamate receptor genes (*GRIN2B*, *GABRG2* and *GABRA1*) in epilepsy susceptibility in Taiwan population. **Method:** Epileptic patients and age, gender matched normal controls were recruited at the department of Neurology of National Taiwan University Hospital (NTUH). 10 ml blood was drawn to extract genomic DNA after signed inform consent. Single nucleotide polymorphisms on *GRIN2B*, *GABRG2* and *GABRA1* (-200G>T and -421C>A for *GRIN2B*; c.588C>T and c.643A>G for *GABRG2*; IVS11+15A>G, c.-142A>G and c.-31C>T for *GABRA1*) were genotyped by real-time PCR SNP analysis. To analyze the associations between SNPs and epilepsy susceptibility, Pearson's $\chi 2$ tests were performed and Bonferroni corrections were used to adjust for multiple comparisons. To further identify the possible gene-gene interaction, logistic regression was performed under adjustment of age, gender and epilepsy syndromes.**Result:** There were 500 epilepsy patients and 200 healthy controls enrolled in this study. Results of statistical analysis showed that the T allele carriers of -200G>T in *GRIN2B* were more likely to be normal subjects, while A allele of c.643A>G in *GABRG2* was over presented in patients with epilepsy (for *GRIN2B* -200G>T, OR:0.464; 95%CI=0.300–0.719; p=0.0006; for*GABRG2* c.643A>G, OR=6.480; 95%CI=3.319–12.652; p<0.0001). There is no significant association between other polymorphisms and epilepsy susceptibility. In the most fitted regression model, under adjusting for confounding factors, *GRIN2B* -200G>T and *GABRG2* c.643A>G influence the occurrence of epilepsy concurrently. **Conclusion:** These results suggested that some polymorphisms in genes encoding for neurotransmitter receptors may be associated with epilepsy susceptibility in Taiwan population. Further studies in different populations are warranted to clarify whether these genetic effects can be observed in different ethnics.

2393F

Interpreting the impact of mutations in *SCN1A* in the context of the most severe phenotypes amongst the generalized epilepsies with febrile seizures plus. *M. Gonsales[1], P. Preto[2], M.A. Montenegro[2], M.M. Guerreiro[2], I. Lopes-Cendes[1].* 1) Department of Medical Genetics, University of Campinas, Campinas, Brazil; 2) Department of Neurology, University of Campinas, Campinas, Brazil.

Dravet and Doose syndromes are severe phenotypes within the clinical spectrum of generalized epilepsy with febrile seizures plus. Several mutations in the neuronal voltage-gated sodium channel a1–subunit gene (*SCN1A*) have been identified in patients within this spectrum. However, the prognostic value of these mutations and a possible correlation with the different clinical subtypes remain unclear. Therefore, the aim of this study was to search for *SCN1A* mutations in patients with Dravet and Doose syndromes, thus establishing genotype-phenotype correlations using prediction algorithms to determine potentially deleterious mutations and comparative analysis with mutations previously described. We performed *SCN1A* mutation screening in 15 patients with Dravet and 13 with Doose syndrome. Eight algorithms were used to analyze the possible impact of the mutations in protein function. In addition, MLPA was used to detect copy number variations (CNVs) within *SCN1A*. Mutations previously described in the literature were compiled and analyzed regarding their location in the protein and their potential deleterious effect. Twelve mutations were identified in patients with Dravet: six missense, three splice-site, two frameshift and one in-frame deletion. Patients with Doose syndrome showed no mutations. Moreover, no copy number variants were identified. The compilation of previously described *SCN1A* mutations revealed a total of 518 nucleotide variants and 32 CNVs in Dravet syndrome while only two mutations were found in patients with typical Doose syndrome. Missense mutations are the most frequent type in Dravet (44%), and amino-acid changes are mostly located in the pore-forming and C-and N-terminal regions of the protein, as also observed in our cohort. The high frequency of *SCN1A* mutations in patients with Dravet suggests that molecular testing is useful for individuals with this phenotype, but not clinically relevant in Doose syndrome. Our results, consistent with the analysis of the 518 mutations previously described, show a predominance of missense mutations, initially considered exclusive of mild phenotypes. These substitutions involve conserved amino-acid located in functionally important regions, while other types of mutation are rather distributed along the protein. The analysis using multiple prediction algorithms was effective, especially for predicting the deleterious effect of transmembrane substitutions. Supported by FAPESP/CNPq.

2394W

Next-generation sequencing and mutation screening in patients with familial mesial temporal lobe epilepsy: investigating the chromosome 18p11.31 candidate locus. R.*O. Santos[1], M.G. Borges[1], W.T.S.T. Ide[1], C.S. Rocha[1], R. Secolin[1], F.M. Artiguenave[1], C.L. Yasuda[2], A.C. Coan[2], M.E. Morita[2], F. Cendes[2], C.V. Maurer-Morelli[1], I. Lopes-Cendes[1].* 1) Medical Genetics Dept, University of Campinas, Campinas, São Paulo, Brazil; 2) Neurology Dept, University of Campinas, Campinas, São Paulo, Brazil.

Background: The epilepsies form a group of chronic neurological diseases characterized by recurrent seizures. Familial mesial temporal lobe epilepsy (FMTLE) is a clinically well characterized syndrome with an autosomal dominant inheritance. Our previous results pointed to a region on chromosome (ch) 18p11.31 as genetically linked to FMTLE. The identified candidate region spans approximately 6Mb and contains 18 annotated genes. **Objectives:** The present work aims to sequence the entire candidate region previously identified on ch 18p11.31 in patients from a large family segregating FMTLE. **Methodology:** DNA samples were amplified from three patients and one control by long range PCR and sequenced using the SOLiD System® (Applied Biosystems, USA). Sequence data was aligned and analyzed by a logic algorithm developed in our laboratory using Perl language. Results from alignment were compared with databanks in order to verify possible functional relationship with the disease physiopathology. **Results:** We identified 5019 single nucleotide polymorphisms (SNPs) in the sequenced region. Among them, 2474 SNPs were not present in the databases of polymorphisms previously described. After performing a series of pipeline adjustments using the bioinformatics algorithm we found three SNPs present only in the three patients and absent in the control. **Conclusions:** Next-generation sequencing allowed a faster and reliable analysis of the candidate region previously identified by linkage studies in FMTLE. Using an interactive bioinformatics pipeline, we have identified putative functional variants in relevant candidate genes for FMTLE on ch 18p11.31. Among them are those related to cortex development, development of neuronal projections, axons guidance and signal transduction. The next step is to confirm these changes in the others family members with FMTLE and verification of the segregation of these variants in this and other families with similar phenotype. FAPESP, São Paulo Brazil.

2395T

Characterization of a GEFS+ locus on 6q16.3q22.31. *P. Szafranski[1], G. Von Allmen[2], B. Graham[1], S. Fogarty[2], A.A. Wilfong[3], S.-H.L. Kang[1], J.A. Fereira[4], W. Bi[1], J.A. Rosenfeld[5], L.G. Shaffer[5], S.W. Cheung[1], P. Stankiewicz[1], S.R. Lalani[1].* 1) Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX; 2) Dept of Neurology, University of Texas Health Science Center at Houston, TX; 3) Pediatric Neurology, Texas Children's Hospital, Baylor College of Medicine, Houston, TX; 4) Dept of Pediatrics, University of South Florida, Tampa, FL; 5) Signature Genomic Laboratories, PerkinElmer, Inc., Spokane, WA.

Generalized epilepsy with febrile seizures plus (GEFS+) is an autosomal dominant syndrome with variable expressivity and incomplete penetrance. To date mutations causing GEFS+ have been identified in four genes encoding sodium channel subunits, *SCN1A, SCN1B, SCN2A*, and *SCN9A* and two genes encoding GABA receptor subunits, *GABRD* and *GABRG2*. In addition, linkage studies have implicated seven other genomic regions. Using linkage analysis in 15 subjects from a large family with GEFS+, Poduri et al. (2009) obtained a maximum lod score of 4.68 at *D6S1706* (between *NUS1* and *SLC35F1*) and identified an ~ 18.1 Mb shared haplotype on 6q16.3q22.31 in affected family members, defining a novel GEFS+ locus. Of note, this interval does not contain either ion channel or GABA receptor genes; no mutations were found in 16 candidate genes in this interval, including *NUS1* and *SLC35F1*. Rosenfeld et al. (2012) reported five individuals with seizures and deletions involving 6q22 and narrowed the critical epilepsy region to ~5 Mb on 6q22.1q22.31. We present six additional patients with deletions overlapping this region, two of whom manifested seizures, patients 3 and 4. Patient 1 is a 7–year-old boy with autistic spectrum disorder in whom we identified an ~ 1.6 Mb deletion in 6q22.2q22.31. Patient 2 is a 6–year-old boy with developmental delay (DD) with an ~ 5.5 Mb *de novo* deletion in 6q21q22.2. Patient 3 is a 12–year-old girl with moderate DD and epilepsy and an ~ 318 kb *de novo* deletion. Patient 4 is an 11–year-old girl with epilepsy and an ~ 247 kb *de novo* deletion in the overlapping region. Patient 5 is an 8–year-old boy with DD and an ~2.9 Mb *de novo* deletion in 6q22.1q22.31. Patient 6 is a 4–month-old boy with a congenital heart defect and DD who had an ~168 kb deletion entirely within *SLC35F1*; he carried additional, clinically significant copy number changes. We sequenced deletion breakpoints in patients 1–4 and narrowed the critical region to ~ 180 kb between *NUS1* and *SLC35F1* on 6q22.1q22.31. This region includes two ESTs, bearing no similarity to any known gene, and a Piwi-interacting RNA (piRNA) gene, *piR-56963*. piRNAs form complexes with Piwi proteins and have been linked to epigenetic and post-transcriptional gene silencing. Recently, neuronal piRNAs have been implicated in the control of memory-related synaptic plasticity. We hypothesize that loss of piR-56963 or dosage alteration of *SLC35F1* or *NUS1* may contribute to GEFS+ with incomplete penetrance.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2396F

Identifying disease genes in a large highly inbred consanguineous kindred from Turkey with idiopathic generalized epilepsy. *F.N. Tuncer[1], S.A. Ugur Iseri[1], M. Calik[2], A. Iscan[3], G. Altiokka Uzun[4], U. Ozbek[1].* 1) Department of Genetics., Istanbul University Institute for Experimental Medicine, Istanbul, Turkey; 2) Department of Pediatric Neurology, Harran University, Faculty of Medicine, Sanliurfa, Turkey; 3) Department of Pediatric Neurology, Bezmialem Vakif University, Faculty of Medicine, Istanbul, Turkey; 4) Neurology Department, Istanbul University, Istanbul Faculty of Medicine, Istanbul, Turkey.

Epilepsy is a complex neurological disorder that affects 1% of the world's population. Among different forms of epilepsies, idiopathic generalized epilepsies (IGEs) are characterized by bilateral and synchronous generalized seizures in the absence of detectable brain lesions or metabolic abnormalities. Thus, the primary etiology for this disorder is believed to be genetic. The proposed study includes a large highly inbred consanguineous kindred with multiple IGE affected individuals. The ultimate aim is identifying a new epilepsy gene to help in delineating the molecular mechanisms of this disease. Physical, neurological and electroencephalography (EEG) examinations were performed on the subjects recruited with information on family history. These examinations revealed 6 affected individuals being 4 siblings and 2 cousins, who exhibit varying types of seizures including generalized tonic-clonic (GTCS), absence and simple febrile seizures. Affected members of this family along with 4 unaffected individuals were genotyped using Illumina Human HumanCytoSNP-12 BeadChip (300K). The genotype data were delineated using easyLinkage Plus software platform, where Mendelian genotyping errors were determined using PedCheck, multipoint lod scores were calculated under the assumption of autosomal recessive inheritance and haplotypes were constructed through GeneHunter. Linkage analysis was performed utilizing 4 affected sibs, their parents and healthy brother. Candidate regions were found on chromosomes 8, 9 and 16 with lod scores of 0.68, 1.20 and 1.08, respectively. Homozygosity mapping was performed on these candidate regions including the genotype of the affected cousin, which revealed partitioned homozygous regions on these chromosomes. Additionally, a 2.79 cM of homozygous region on chromosome 12 was depicted in all affected individuals along with the father and the healthy brother of the affected siblings. These preliminary results support a genetic aspect of epilepsy in this family, but needs to be further delineated using next generation sequencing to depict the gene(s) responsible for the disease. Hence, prospective work includes exome sequencing on the selected family members, as well as consultation with the family members to re-confirm the disease onset and progress.

2397W

Haplotype sharing and linkage analyses of multigenerational families with multiple sclerosis identified two chromosomal regions at 12p12.3–q12 and 16q21–q22.3. *N. Matsunami[1], J. Stevens[1], L. Baird[1], B. Otterud[1], T. Varvil[1], T. Leppert[1], J.W. Rose[2], M.F. Leppert[1].* 1) Department of Human Genetics, University of Utah, Salt Lake City, UT, USA; 2) Department of Neurology, University of Utah, Salt Lake City, UT, USA.

Multiple Sclerosis (MS) is a complex neuro-inflammatory and autoimmune disorder of the central nervous system. Various family studies have shown that MS has a strong genetic component, particularly at the HLA locus on chromosome 6p21.32. We have carried out Affymetrix Genome-Wide Human SNP array 6.0 experiments on affected and unaffected members of 19 multigenerational Utah families with MS. The number of affected individuals per family ranges from 2 to 7. Phased haplotype sharing analysis and linkage studies identified two regions of interest; 12p12.3–q12 (hLOD= 6.4) and 16q21–q22.3 (hLOD=4.5). The chromosome 12 region overlaps extensively with the previously reported region described by Vitale et al. (Human Molecular Genetics, 2002, Vol.11, No. 3, 295–300) in another multiplex MS family. The size of our shared phased haplotype regions on chr12 and 16 are 21 Mb, containing 95 genes, and 9 Mb, containing 152 genes, respectively. In order to test the hypothesis that rare functional variants in these regions may contribute to MS in these families, we used a custom targeted-enrichment approach to selectively capture exon and promoter regions (Agilent SureSelect) and subsequently sequenced 25 individuals from 11 families that supported chr12 and 16 linkages on an Illumina HiSeq instrument. Reference sequence alignment was carried out using BWA (Li H & Durbin R, 2010) and Novoalign (www.novocraft.com). Variant annotation, classification, and prioritization were carried out using ANNOVAR (Wang K, Li M, & Hakonarson H, 2010), VAAST (Yandell M et al., 2011) and the DNA-Seq Analysis Package of Golden Helix SNP & Variation Suite (SVS, Golden Helix Inc.). We are now in the process of molecular validation of putative functional variants. Following molecular validation, we intend to compare findings with gene functions relevant to the MS pathophysiology.

2398T

Homozygosity Mapping and Exome sequencing Identifies a Gene on 11q24.2 for non-syndromic autosomal recessive intellectual disability (NS-ARID) in a Pakistani family. *MA. Rafiq[1], IA. Balouch[1,2], K. Mittal[1], A. Mikhailov[1], M. Ayub[3], P. John[2], JB. Vincent[1,4].* 1) Molecular Neuropsychiatry, Centre for Addiction and Mental Health, Toronto, Ontario, Canada; 2) Atta-ur-Rehman School of Applied Biosciences (ASAB), National University of Sciences and Technology (NUST), H-12 Sector, Islamabad-Pakistan; 3) Tees, Esk and Wear Valleys NHS Foundation Trust United Kingdom; 4) Dept. of Psychiatry, University of Toronto, Toronto, Ontario, Canada.

Intellectual disability (ID), or mental retardation (MR), is a neurodevelopmental disorder that has a huge impact on the health care system. Here we report a study on non-syndromic autosomal recessive intellectual disability (NS-ARID) in a consanguineous Pakistani family. Two affected and one unaffected members of a family were genotyped using Affymetrix 500K single-nucleotide polymorphism (SNP) microarrays. Analysis of microarray data identified three genomic regions (11q24.1–q25, 14q11.2 and 17q24.2–q24.3) that were homozygous-by-descent (HBD) in both affected individuals. All family members (2 affected and 6 unaffected) were genotyped by using highly polymorphic microsatellite markers present in three regions of HBD discovered by microarray analysis. By using this approach we were able to exclude HBD for the 14q11.2 and 17q24.2–q24.3 regions. Using whole exome sequence capture followed by next generation sequencing on the SOLiD 4, we identified a homozygous splice donor site G>A mutation in both affected individuals in a gene on 11q24.2. Sequencing of cDNA from both affected individuals revealed a new splice site 44 nucleotide downstream in the intronic region. Consequently, 15 extra amino acids are added to the protein, and are predicted to disrupt crucial protein function. Studies to measure the effect of mutation on protein activity are being conducted.

2399F

Investigation of inherited copy number variations in extended Asperger syndrome families. *M. Rossi[1,2], K. Rehnström[3], O. Pietiläinen[1,3], V. Leppä[1], T. Hiekkalinna[2], T. Nieminen-von Wendt[4], J. Barrett[3], I. Hovatta[1,5,6], S. Ripatti[1], A. Palotie[1,3].* 1) Institute for Molecular Medicine Finland (FIMM), Helsinki, Finland; 2) Department of Molecular Medicine, National Public Health Institute, Helsinki, Finland; 3) Wellcome Trust Sanger Institute, Hinxton, Cambridge, United Kingdom; 4) Unit of Child Neurology, Hospital for Children and Adolescents, Helsinki, Finland; 5) Research Program of Molecular Neurology, University of Helsinki, Helsinki, Finland; 6) Department of Medical Genetics, University of Helsinki, Helsinki, Finland.

Background: Asperger syndrome (AS) belongs to the Autism spectrum disorders (ASDs). AS is characterized by impairments in social interaction and communication, and restricted interests and activities but lacks major cognitive deficiencies. Prevalence estimates for AS vary between 2.5–2.9/1000. Rare copy number variants (CNVs) have emerged as a potential genetic factor for several neurodevelopmental disorders. Low frequency and family specific CNVs have been shown to be involved in ASDs. Our aim was to analyze large multigenerational AS families for inherited CNVs that may have incomplete penetrance and moderate to strong influence on AS predisposition. Materials and methods: We genotyped 177 individuals (naffected= 125) from 24 AS families using Illumina HumanOmni1–Quad BeadChip and Illumina HumanHap550 BeadChip. As controls, we used a set of 6125 population samples genotyped with Illumina 670K BeadChip. First, we analyzed the families for linked loci using a non-parametric multipoint linkage (NPL) analysis (MERLIN). We searched these family specific identical-by-descent (IBD) shared genomic regions for CNVs segregating with AS phenotype. CNVs were called using QuantiSNP and PennCNV. Results: We assumed incomplete penetrance and included CNVs segregating to all except one affected individuals within each family. We identified three intronic deletions in genes XPR1 (4.7kb), MRPL44 (5.5kb) and MIPOL1, and one intergenic 5.5kb duplication on chromosome 6q27. All CNVs were found in separate families except MIPOL1 deletion and intergenic duplication in the same family. Conclusions: We identified four inherited structural variants that may be involved in AS etiology. These CNVs overlap with three genes that have no previously known function related to AS. XPR1 is a retroviral virus receptor and MIPOL1 has been associated with mirror-image polydactyly. MRPL44 encodes for mitochondrial ribosomal proteins and has been linked to mitochondrial cardiomyopathy. Moreover, there are family specific CNVs at several previously identified ASD candidate loci, such as 15q11–13. Unlike previous reports for example in infantile autism, we did not observe a burden of large CNVs in AS. Our findings imply that, together with other genetic factors, rare, heritable CNVs may have a role in AS etiology.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2400W

Gene analysis of familial and sporadic amyotrophic lateral sclerosis (ALS) in Kyushu island of Japan. *H. Furuya[1], A. Watanabe[1], H. Arahata[1], Y. Kawano[1], N. Fujii[1], H. Kikuchi[2], T. Fujiwara[3], T. Hokonohara[3], A. Iwaki[3], H. Shibata[3], Y. Fukumaki[3].* 1) Department of Neurology, National Omuta Hospital, Omuta, Fukuoka, Japan; 2) Murakami Karindo Hospital, Fukuoka 812–2161, Japan; 3) Division of Human Molecular Genetics, Research Center for Genetic Information, Medical Institute of Bioregulation, Kyushu University, Fukuoka 812–8582, Japan.

Objective: We analyzed reported causative genes for unrelated familial and sporadic amyotrophic lateral sclerosis (ALS) patients in Kyushu island, a western part of Japan. Background: ALS is a fatal neurodegenerative disease characterized clinically by progressive paralysis leading to death from respiratory failure. The morbidity prevalence rate of ALS is 5–6/100,000 and there are currently no effective therapies. Approximately 5% of cases are familial in nature, whereas the bulk of patients diagnosed with the disease are classified as sporadic. To date, a number of genes have been discovered as causative for classical familial ALS, such as *SOD1, TDP-43, FUS, CHMP2B*, and *OPTN*, and these genes cumulatively account for less than 25% of familial cases. To reveal the hidden familial cases and causative genes of ALS patients, we conducted gene analysis of clinically diagnosed ALS patients. Methods: We screened mutations of four causative genes (*SOD1, TDP-43, FUS, CHMP2B*) in probands of two ALS families and 85 sporadic cases, by the direct sequencing method. Results: We found a known missense mutation (R521C) in *FUS* of one ALS family and a novel *FUS* mutation (H517P) in another ALS family. In sporadic cases, one missense mutation (R521H) in *FUS*, one know missense *SOD1* mutation (C146R) were found. Conclusions: We found two *FUS* missense mutations in familial ALS. Out of 85 sporadic cases, two missense mutations (2.4%) in known ALS related genes were found. These results suggest that there is more heterogeneity of gene mutations in sporadic ALS cases in Kyushu island, than we have expected previously.

2401T

Replication and generalization of genetic risk factors for depression, anxiety and panic attack in a population-based cohort with African, European and Mexican ancestries. *J.N. Vega, K. Brown-Gentry, D.C. Crawford, T.A. Thornton-Wells.* Center for Human Genetics Research, Vanderbilt University, Nashville, TN.

Anxiety and mood disorders are two of the most common psychiatric disorders, with lifetime prevalence estimates above 20% and heritability estimates of 35–50%, despite clear environmental stressors and risk factors. Many functional candidate genes have been implicated in anxiety, stress and mood; however, investigation within non-European populations is incomplete. As part of the Epidemiologic Architecture for Genes Linked to Environment (EAGLE), we conducted an association study of previously-reported genetic risk variants for anxiety or depression in the ethnically diverse population-based cohort, the National Health and Nutrition Examination Survey (NHANES) ascertained and maintained by the Centers for Disease Control and Prevention (CDC). For this study, we accessed NHANES 1999–2002, which includes psychological survey data for 596 non-Hispanic whites (NHWs), 231 non-Hispanic blacks (NHBs), and 354 Mexican Americans (MAs). We selected 35 SNPs in 23 candidate genes previously implicated in anxiety, mood or personality disorders or the hypothalamic-pituitary-adrenal (HPA) axis. We performed genotyping using the Sequenom and Taqman OpenArray platforms and tested for association using logistic regression with an additive allelic model. The G allele of rs6269 in COMT was associated with decreased anxiety in MAs (OR=0.447, CI=[0.262–0.763], p= 0.0031), consistent with previous reports that the G allele is associated with higher COMT activity and lower pain sensitivity. The T allele of rs6999100 in TRIM55, downstream of CRH, was associated with decreased rates of anxiety in non-Hispanic whites (OR=0.633, CI=[0.426–0.941], p=0.0236), consistent with a previous report that the T allele is under-transmitted to individuals with behavioral inhibition to the unfamiliar, with is an intermediate phenotype and risk factor for panic disorder and social phobia. The T allele of rs4570625 in the upstream regulatory region of TPH was associated with decreased rates of depression in non-Hispanic blacks (OR = 0.453, CI = [0.234–0.876], p = 0.0187), in contrast to previous reports in Caucasians that the T allele is associated with amygdala responsivity, panic disorder, anxiety-related disorders and affective disorders. Despite being underpowered, we were able to replicate and characterize genetic associations for anxiety, depression and panic attack. Larger, diverse studies are required to characterize other associations with small genetic effect sizes.

2402F

GENOME-WIDE GENETIC STUDIES OF TWO CONSANGUINEOUS PAKISTANI MULTIPLEX PEDIGREES WITH SCHIZOPHRENIA AND BIPOLAR DISORDER. *L. Xiong[1,2,3], S. Zhou[3], M. Christian[3,4], S. Provost[5], G. Asselin[5], M.Q. Brohi[6], M. Denton[7], R. Joober[8], L.E. DeLisi[9], G.A. Rouleau[10].* 1) Dept. of Psychiatry, University of Montreal, Montreal, Quebec, Canada; 2) Centre de recherche Fernand-Seguin, Hôpital Louis H. Lafontaine, Montreal, Quebec, Canada; 3) University of Montreal Hospital Research Centre, Montreal, Quebec, Canada; 4) Department of Zoology, University of Sindh Jamshoro, Sindh, Pakistan; 5) Montreal Heart Institute, University of Montreal, Montreal, Quebec, Canada; 6) Sir Cowasjee Jehangir Institute of Psychiatry, Sindh, Pakistan; 7) Department of Biochemistry, University of Otago, Australia; 8) Montreal Douglas Hospital, Department of Psychiatry, McGill University, Montreal, Quebec, Canada; 9) Department of Psychiatry, Harvard Medical School, Boston, MA, USA; 10) Ste-Justine Hospital Research Center, University of Montreal, Montreal, Quebec, Canada.

Background: Pakistan has the highest rate of endogamous/consanguineous marriage in the world. The potential influence of such a long history of consanguineous unions on overall levels of homozygosity genome-wide and on genetic disease profiles remains largely under-investigated, especially for more common and complex diseases. In the past few years, with the cooperation from local clinicians, we have identified and characterized some large multiplex pedigrees aggregated with schizophrenia (SCZ), schizoaffective disorder (SAF) and bipolar disorder (BPD) from Sindh, Pakistan. **Methods:** Full family studies were carried out in the field and at local clinics by the local geneticist, psychiatrist, and psychologist; and further validated by two independent psychiatrists. Standard clinical instruments, i.e. DIGS/FIGS, were used in each participating individual. A high density SNP array genome scan was performed in two selected SCZ and BPD pedigree, respectively. Parametric and nonparametric linkage analyses were carried out on these two pedigrees separately and jointly using selected SNPs genome-wide. Whole exome sequencing was also performed in all affected individuals from these two selected pedigrees. **Results:** Data analysis has first excluded potential disease-associated common cytogenetic abnormality and copy number variant (CNV) in each pedigree. Joint SIMWALK multipoint linkage analysis under an autosomal dominant model in these two pedigrees identified three regions with LOD scores above 3.0, one region on chromosome 2 (LOD-1 interval 17 cM), and two on chromosome 11 (LOD-1 intervals: 2.6 cM and 5.6 cM); one of the two pedigrees produced maxLOD score of 3.35 at the chromosome 2 locus. Both loci were further supported by NPL statistics. Further haplotype sharing analyses with more informative microsatellite markers from these two candidate regions in these two pedigrees have identified major shared haplotype(s) within each pedigree. These results strongly support high risk genetic variant(s) in these pedigrees. Combined exome sequencing data analyses are underway. **Discussion:** Preliminary linkage analyses strongly suggest major gene effect in these two large consanguineous pedigrees aggregated with SCZ and BPD, with potential candidate regions identified, harbouring high-penetrant risk variants.

2403W

A comprehensive genotype-phenotype study for *CHRNA7* and *CHRFAM7A* in autism and intellectual disability. *J. Ge[1], C.P. Schaaf[1], B. Zhang[1], A.L. Hall[1], M. Ali[1], S. Tinschert[2], U. Moog[3], A. Patel[1], P. Stankiewicz[1], A.L. Beaudet[1].* 1) Molecular and Human Genetics, Baylor College of Medicine, Houston, TX; 2) University of Dresden, Germany; 3) University of Heidelberg, Germany.

Deletions and duplications of chromosome 15q13.3 have been associated with autism, schizophrenia, epilepsy, ADHD and intellectual disability. The most likely candidate gene within this region is *CHRNA7*, encoding the α-7 neuronal nicotinic acetylcholine receptor which functions as a ligand gated ion channel at neuronal synapses. Studies of *CHRNA7* are complicated by a chimeric gene *CHRFAM7A*, a fusion of the last 6 exons of *CHRNA7* and exons D through A of *FAM7A*, which is expressed in human brain and elsewhere in the body. A small duplication of*CHRNA7* with uncertain phenotypic effect occurs in about 1 in 185 individuals. *CHRFAM7A* is present as 0, 1, or 2 copies per chromosome and commonly has a 2–bp deletion in exon 6. CHRFAM7A is reported to have a dominant negative effect on CHRNA7. Complete genotyping for the *CHRNA7* cluster requires detailed copy number analysis for all segments, determination of the 2–bp deletion, and sequencing of all coding regions distinguishing *CHRNA7* and *CHRFAM7A* by analysis of cDNA or genomic DNA with long-range PCR. BP4 BP5 inversions occur and may affect CHRNA7 function. We have performed extensive genotyping and are expanding the analysis for 50 Bp4 Bp5 deletion, 25 duplication, 8 *CHRNA7* deletion and 177 *CHRNA7* duplication families with penetrant and nonpenetrant members detected in the CLIA lab, for three families with homozygous deletion of *CHRNA7*, one family with *CHRNA7*triplication, and for ~500 Simons Simplex autism trios, ~200 NIMH schizophrenia plus epilepsy cases, and ~600 NIMH normal controls.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2404T

Genetic variant of AKT1 and AKTIP associated with risk for late-onset depression in a Brazilian population. *P. Pereira[1], M. Romano-Silva[1], M. Bicalho[1,2], E. Moraes[2], L. Malloy-Diniz[1,2], R. Nicolato[1], G. Pimenta[1], M. Mello[1], I. Bozzi[1], D. Miranda[1].* 1) UFMG, Belo Horizonte, Minas Gerais, Brazil; 2) Centro Jenny Faria de Assistência à Saúde do Idoso e da Mulher, Hospital das Clínicas, Universidade Federal de Minas Gerais (UFMG), Brazil.

Major depressive disorder (MDD) is one of the most prevalent and costly neuropsychiatric diseases. DSM-IV defines late-onset depression (LOD) as a disorder that starts after or at the age of 65 years old. Several lines of evidence have suggested that gene-gene interactions may confer susceptibility to MDD. The serine/threonine protein kinase protein B (PKB), also known as Akt1, is a downstream enzyme that has been implicated in the pathogenesis of neurotransmitter-related disorders, such as depression. The PI3K/AKT/GSK3b pathway is involved in neuronal survival, plasticity and also neurogenesis. The identification of a novel AKT1– binding protein (AKTIP) was pointed to as an important new target for investigation. AKTIP binds directly to AKT1, enhancing the phosphorylation of regulatory sites and this modulation are affected by Akt1 activation. The association of AKT1 and AKTIP polymorphisms with depressive symptoms was not investigated in late-onset depression (LOD). The aim of the present study was to examine the association between polymorphisms in the AKT1 and AKTIP genes and late-onset depression (LOD) in the Brazilian population. We genotyped nine tagSNPs in the AKT1/AKTIP genes in 188 outpatients with LOD and 79 healthy individuals from the control group. The participants were evaluated using DSM-IV criteria, MINI-PLUS and the Geriatric Depression Scale (GDS). Our findings suggested an association between the tagSNP rs3730358 homozygous A/A (p= 0.006) and risk for LOD. A strong association of allele A and increased risk for LOD was demonstrated for the tagSNP rs3730358 (p-value= 0.003). Our study is the first one to establish the association of the AKT1/AKTIP gene and late onset depression and further studies are necessary to clarify the functional role of these proteins.

2405F

Whole genome gene expression analysis of patients with 22q11 Deletion Syndrome and controls. *K.K Mantripragada, S. Chawner, M.B van den Bree, R.G Thomas, L. Bates, M.J Owen, N.M Wiliams.* Institute of Psychological Medicine and Clinical Neurosciences, School of Medicine, Cardiff University, Cardiff, United Kingdom CF14 4XN.

The 22q11 Deletion Syndrome (22q11DS), also known as Velo-Cardio-Facial Syndrome(VCFS), has a variable constellation of phenotypes including cardiac malformations, dysmorphology and high prevalence of psychiatric diseases, including schizophrenia. Genes located within the 22q11 deleted region have been suggested to alter the expression of several genes involved in maturation and development of neurons and neurotransmission. Using Illumina HumanHT-12 v4 Expression BeadChip system, we investigated genome-wide gene expression of lymphoblastoid cell lines (LCLs) from 17 individuals including 7 VCFS patients with schizophrenia, 7 VCFS patients without schizophrenia and 3 non-affected parents of VCFS individuals. We analysed gene expression across the genome with special attention to 22q11 genes to examine whether haploinsufficiency of 22q11 genes influenced functional genetic networks relevant to schizophrenia. Our preliminary data from VCFS individuals shows that 17% of the genes display altered expression (p value <0.05) between the schizophrenia and non-schizophrenia patients. Pathway analysis suggests that these genes are significantly enriched in 20 pathways. These pathways fall into three themes: embryonic brain development, immune response signalling and protection and recovery to brain damage. This pilot study substantiates previous reports which support 22q11 deletion syndrome as a research model for schizophrenia. This indicates that deletion of 22q11 region in conjunction with other genetic factors is relevant in schizophrenia. These are preliminary findings and are limited by small sample numbers. Further analysis using extensive sample numbers is required to implicate specific pathways in schizophrenia.

2406W

Fine-mapping of a region of interest in an extended family with multiple cases of autism. *N.H. Chapman[1], K. Ankeman[1], R. Bernier[1], A. Estes[1], J. Munson[1], Z. Brkanac[1], W. Raskind[1], H. Coon[2], E.M. Wijsman[1].* 1) University of Washington, Seattle, WA; 2) University of Utah, Salt Lake City, UT.

We describe a genome scan and follow-up of autism in a single large family, made up of 47 people in 4 generations, with 11 affected people and 10 unaffected siblings of affected people. STR genotypes are available on 45 individuals, and dense SNP genotypes are available on 42.

An initial single-marker genome scan using STR-based parametric linkage analysis identified several regions with suggestive lod scores(>1.0). One of these was consistently larger than the others, and was robust to details of the parametric model. We focused on this region to explore the use of dense SNP data to refine the region of interest. Multipoint linkage analyses were done in three stages: Stage 1: using STR markers alone; Stage 2: using STRs and evenly spaced supplemental SNPs with maximum heterozygosity; and Stage 3: using STRs and carefully selected SNPs.

Due to the size of the pedigree, a Markov chain Monte Carlo (MCMC) approach was used to calculate likelihoods and lod scores. The program gl_auto, available in MORGANv3.0, was used to sample patterns of inheritance consistent with the observed marker genotypes, conditional on the pedigree structure, marker map, and marker allele frequencies. At each marker of interest, we used the program IBDgraph to sort samples into classes with equivalent inheritance patterns. Finally, at each marker we calculated likelihoods for each class, and used a weighted sum over classes as the estimated likelihood to obtain lod scores for each marker.

The results of Stage 1 identified a region of 7.2 megabases(Mb), covered by 8 STR markers spaced an average of 1.9cM apart. In Stage 2, 20 SNPs were added (avg. spacing 0.5cM), narrowing the region by more than 50% to 3.15 Mb. Finally, in Stage 3, 6 of the supplemental SNPs were replaced with others that were informative in meioses where recombination events occurred. This reduced the region of interest a further 34% to 2.1Mb. We are now evaluating variants identified in this region for plausibility as autism susceptibility alleles.

2407T

Investigation of GWAS findings in schizophrenia in a case control-and a family sample from Indonesia. *D.B. Wildenauer[1], W. Qin[1], M.D.B. Wildenauer[1], A.A.A.A. Kusumawhardani[2], B. Benyamin[3], S.G. Schwab[4], Indonesia Schizophrenia Genetics Consortium.* 1) CCRN School of Psychiatry, UWA, Claremont, Australia; 2) Dept of Psychiatry, University of Indonesia, Jakarta, Indonesia; 3) Queensland Brain Institute, University of Queensland, Brisbane, Australia; 4) University of Erlangen-Nuremberg, Erlangen, Germany.

Schizophrenia is a devastating mental disorder with onset in early adulthood, affecting not only the individual but also having severe impact on their families and on the socio-economical framework of a society. The relatively high heritability (up to 80%) suggests the presence of genetic susceptibility factors, which may be identified in large case control samples. Population specific susceptibility alleles may exist. Identification of these alleles may aid in gene identification as well as in characterization of different course and outcome in developing and developed countries. We have consecutively ascertained 1081 individuals with schizophrenia and 1111 non-psychiatric controls in five Mental State Hospitals in the area of Jakarta, Indonesia. Diagnosis was made according to the DSMIV criteria by psychiatrists, trained in the use of a structured interview (DIP) in Bahasa Indonesia translation. We used a DNA test panel with 374 SNPs from non-coding regions (Illumina) and characterized all samples in respect to DNA quality, duplications, sampling error. For assessment of ancestral background and removal of ethnical out-layers the programs STRUCTURE and EIGENSTRAT were used. In addition, we have an independent sample of 138 families from Indonesia with 2 or 3 affected siblings(318 affected, 686 individuals in total) available for family based association analysis. We investigated DNA variants located in the CUB Sushi and multiple domains 1 gene (CSMD1), chr 8p23.2, and in the HLA-region on chromosome 6p (Ripke et al; Nat Genet. 2011;43:969–76, Yue et al; Nat Genet. 2011;43:1228–31). In the case control sample, rs10503253, rs10503256 (CSMD1 gene) and rs1635, rs2142731 (MHC region) were associated with schizophrenia (p-values = 0.0097, 0.0066, 0.047 and 0.016, respectively). In addition, rs10503253 produced a p=0.027 in the family sample. Even though limited in sample size, these findings may contribute to evaluation of susceptibility loci detected in recent GWA studies and demonstrate the usefulness of relatively homogeneous samples with different ancestral background.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2408F

Association analysis of mitochondrial DNA with suicidal behavior in schizophrenia. *V. De Luca, A. Hassan, N. Hettige, C. Borlido, G. Polsinelli.* Dept Neurogenetics, Univ Toronto, Toronto, ON, Canada.

Mitochondrial dysfunction is caused by DNA mutations in the mitochondrial chromosome or the nuclear chromosomes. Patients with psychiatric illness may carry specific mitochondrial haplotypes. To date, this issue has received little attention in the literature, and mitochondrial DNA has been under-investigated in psychiatric patients. This work describes the association analysis of 42 SNPs in the mitochondrial DNA in 81 patients with diagnosis of schizophrenia who received a comprehensive assessment for suicidal behavior. Lifetime suicide attempt was ascertained with the following standardized instruments: Beck Suicide Ideation Scale and the Columbia-Suicide Severity Scale. The presence of early of trauma was assessed using the CTQ. In our sample there were 40 subjects who attempted suicide at least once in their life. The analysis of the mitochondrial SNPs showed no association with suicide attempt but the SNP T12706C showed a slight trend (p= 0.063). Mitochondrial DNA is not associated with suicide attempt in schizophrenia in this pilot sample however psychiatric disorders, including mood disorders and psychosis should be investigated for association with mitochondrial DNA in future studies.

2409W

Application of whole genome resequencing to identify the causative gene for an autosomal recessive oculopharyngodistal myopathy. *M. Taira[1], H. Ishiura[1], Y. Ichikawa[1], B. Ahsan[1], Y. Fukuda[1], J. Mitsui[1], Y. Takahashi[1], K. Higasa[2], J. Yoshimura[2], I. Nishino[3], J. Shimizu[1], S. Murayama[4], E. Uyama[5], S. Morishita[2], J. Goto[1], S. Tsuji[1].* 1) Department of Neurology, The University of Tokyo Graduate School of Medicine, Tokyo, Japan; 2) Department of Computational Biology, The University of Tokyo Graduate School of Frontier Sciences,Tokyo,Japan; 3) Department of Neuromuscular Research, National Institute of Neuroscience National Center of Neurology and Psychiatry, Tokyo, Japan; 4) Department of Geriatric Neuroscience, Tokyo Metropolitan Institute of Gerontology, Tokyo, Japan; 5) Department of Neurology, Kumamoto Takumadai Hospital, Kumamoto, Japan.

[Purpose] Oculopharyngodistal myopathy (OPDM) is characterized by adult-onset of eye and facial muscle weakness, distal muscle weakness and atrophy, and pharyngeal involvement. Since its original description on four Japanese families, patients presenting with similar clinical phenotypes have been described in 23 families by 2011. The modes of inheritance in these families are variable and the causative genes have not been identified. We applied whole genome resequencing to identify the causative gene for an OPDM family with autosomal recessive mode of inheritance. [Methods] A patient and three unaffected siblings born to consanguineous parents were enrolled in the study. We first conducted SNP (single nucleotide polymorphism) genotyping (Affymetrix 100K chip), and parametric linkage analysis using SNP HiTLink (Y Fukuda et al, 2009) and Allegro. Whole genome resequence analysis was carried out by Illumina HiSeq2000. For exons in the candidate regions with low coverages (<= 3X), direct nucleotide sequence analysis was further conducted to identify variants in these exons. [Results] Patient is a 66–year-old male. His initial symptom was dysarthria at the age of 17. When he was 58 years old, restricted eye movement, severe paralysis of facial muscles, difficulty in swallowing, distal upper extremity and proximal-dominant lower extremity muscle weakness, and respiratory failure were noted. Expansion of the GCG repeat sequence in PABPN1 was excluded. Linkage analysis showed approximately 75.9Mb with a maximum LOD score of 1.57. Whole genome resequence analysis showed 3,661,037 single nucleotide variations and 580,832 short insertion/deletions, of which 66,466 and 54,145 were novel, respectively. Among these variants, we identified only one novel nonsynonymous variant in a gene on the candidate regions, substituting threonine for evolutionally conserved alanine. The substitution was not found in 276 Japanese healthy subjects. Mutational analysis of this gene in 6 other patients presenting with similar clinical symptoms of OPDM did not reveal any mutations. [Conclusion] We have identified a putative gene for an autosomal recessive form of OPDM. Further analysis on additional OPDM families will be needed to validate this gene as the causative gene for OPDM.

2410T

Glucocerebrosidase Mutations and Dementia with Lewy Bodies. *L. Clark[1], K. Mirzozoda[1], J.H. Lee[1,4,5], R. Mayeux[1,2,3,4,5,6], L.S. Honig[1,4,6].* 1) Taub Institute for Research on Alzheimer's Disease and the Aging Brain, Columbia University, New York, NY; 2) Department of Pathology and Cell Biology, Columbia University, new York, NY; 3) Center for Human Genetics, Columbia University, New York, NY; 4) Gertrude H. Sergievsky Center, Columbia University, New York, NY; 5) Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY; 6) Department of Neurology, Columbia University, New York, NY.

Background: Mutations in the glucocerebrosidase gene (GBA) are a risk factor for clinically-diagnosed Parkinson Disease and for neuropathologically confirmed Dementia with Lewy Bodies (DLB). Our prior work showed that in 187 autopsied brains from controls and patients with DLB and Alzheimer disease (AD), carriers of GBA mutations were more likely than non-carriers to have cortical Lewy bodies regardless of AD pathology. In this study we test whether in a clinic population, GBA mutations relate to DLB. Methods: All subjects enrolled in the Columbia University Alzheimer's Disease Research Center with available DNA underwent full gene sequencing of GBA. This included 432 subjects, of whom 191 (44%) were male and 195 (45%) had dementia, 13 (3%) with clinical DLB. Results: Thirty subjects (6.9%) had 9 different GBA mutations including 8 missense and 1 silent (synonymous); 2 mutations were novel, not reported previously. The most common variants were N370S (8 subjects) and K-27R (10 subjects), and these were strongly non-randomly distributed by ethnicity: 100% (10/10) of K-27R were in non-Caucasians (6 non-Hispanic blacks, 2 Hispanic, 2 other), while 85% (17/20) of the other mutations were in Caucasians, consistent with likely African origins for K-27R, and likely Ashkenazi origin for N370S. Classifying subjects by their described mutation phenotypic effect in terms of Gaucher's disease revealed: 1 silent (I119I), 8 mild (all N370S), 3 severe (2 L444P and 1 R463C), and 18 with unknown phenotypic effect (10 K(-27)R, 5 E326K, 1 W184R, 1 V460L, 1D443N). Overall, 11/432 (2.5%) of tested subjects, or 11/30 (37%) of subjects with a GBA mutation, had mild or severe mutations. Comparison of the 30 mutation-carriers with non-carriers revealed no differences in demographics or diagnoses. However, comparing mild/severe carriers to non-carriers showed an increased risk of DLB (18.2% vs 2.7%; p= 0.04), and trends towards higher proportion of males (64% vs 43%), and higher UPDRS score (10.4±15.5 vs 7.59±12.0). Conclusions: GBA mutations occurred in 7% of subjects; the K-27R signal sequence variant is likely associated with African origin. We did not observe a significant association of GBA mutation carrier status with clinical DLB. But confining the analysis to carriers of mutations known to have a phenotypic effect in Gaucher's disease, did show an association with clinical DLB in this referral population, like that previously demonstrated for autopsy-proven DLB.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2411F

Both ISRE and GAS Interferon regulatory sequences found in chromosome 21 genes that are important in brain development and function. *A. Jagadeesh, J. Nagappa, K. Sampson, H. North, L. Gutman, P. Lenferna de la Motte, M. Sellers, I. Dhanoya, D. Elyaman, J. Husni, L.E. Maroun.* Microbiology and Immunology, St.Matthew's University School of Medicine, Grand Cayman, Cayman islands.

The interferons can cause neuropathology.The cells from a person with trisomy 21 (T21)(Down syndrome(DS))would be expected to show a 50%increase in sensitivity to interferon (IFN) based on the presence on Ch21 of 3 of the 4 main interferon receptor component genes.The expected presence of a 50%increase in IFN receptors on T21 cells has been demonstrated.However,T21 cells are excessively sensitive to IFN action in the range of 5 to 10 fold greater than their euploid counterparts.We test our hypothesis that this greater-than-expected sensitivity could be due to the presence on Ch21 of genes that are a target for IFN up regulation by using online search engines to search for IFN regulatory sequences in or close to Ch21 genes.The control site for IFN-alpha regulated genes is termed the ISRE (Interferon Stimulated Response Element). The comparable site for IFN-gamma regulated genes is termed the GAS(Gamma Activated Sequence).Our strict criteria for labeling a gene as potentially interferon regulated required contiguous homology to at least 11 of 13 or 12 of 15 bases of a known ISRE,plus complete homology to at least one known GAS sequence, all located within the gene or, just upstream of the gene transcription start site. In addition,we consider here only genes potentially involved in brain development, function or dysfunction.To date,the following four genes have been found that meet these stringent criteria (based on homology with documented ISRE/GAS sequences):**DOPEY2:** The distribution of this protein during brain development has led to the suggestion that it plays a role in DS.Sites found:ISRE:FactorB,HLA-A,GAS:FCGR1,ICAM-1.**SIM2:**SIM2is a transcription factor found in the brain regions altered in DS.Sites found:ISRE:MX1,GAS:c-fos67[in tandem with ICAM-1].**PCBP3:**PCBP3 is a splicing factor that can alter the mRNA for Tau,a suspect gene in DS. Sites found:ISRE:BF,HLA-A(present twice),Factor B, GAS:ICAM-1.**RCAN1:**RCAN1 is a regulator of calcineurin.Its aberrant expression is associated with altered neuronal vesicular trafficking.Sites found:ISRE:ISG20,GIP3,IFITM3,GAS:IRF-1.There has been no previous indication in the literature that these genes are IFN regulated. They join a group of genes on Ch21 already shown to be IFN regulated (NCAM2, mir155, MX, and CD18).Together,these genes support the possibility that the supersensitivity of Ch21 cells to the IFNs may be due to the presence on Ch21 of both IFN receptor genes and genes that are a target of IFN action.

2412W

Genome-wide SNP analysis in a large consanguineous Turkish family reveals diagnosis as Unverricht-Lundborg Disease. *U. Ozbek[1], S.A. Ugur Iseri[1], F.N. Tuncer[1], I. Karacan[1], O. Ozdemir[1], G. Altiokka Uzun[2], M. Calik[3], A. Iscan[4].* 1) Genetics Department, Istanbul University, Institute for Experimental Medicine (DETAE), Istanbul, Turkey; 2) Neurology Department, Istanbul University, Istanbul Faculty of Medicine, Istanbul, Turkey; 3) Department of Pediatric Neurology, Harran University, Faculty of Medicine,Sanliurfa, Turkey; 4) Department of Pediatric Neurology, Bezmialem Vakif University, Faculty of Medicine, Istanbul, Turkey.

Progressive myoclonic epilepsies (PMEs) are composed of a rare group of inherited neurodegenerative diseases comprising 1% of all epileptic syndromes that are characterized by myclonic, tonic-clonic seizures, progressive neurological dysfunction, dementia and ataxia. This study highlights a large inbred consanguineous family from Turkey with 6 affected sibs out of 10 having a progressive form of myoclonic epilepsy. Affected individuals have been subjected to physical, neurological and electroencephalography (EEG) examinations with consultation on family history. Examinations revealed homogenous clinical findings on all affected individuals with tonic seizures starting at age 9. Among the affected sibs, 2 daughters have been confined to wheelchair since the age of 13. The eldest son is wheelchair free, but suffers from recurrent atonic (dropping) seizures. All affected sibs exhibit mental retardation, except for the two youngest sons. EEG recordings display generalized epileptiform activity for affected individuals, which are accompanied by normal cranial magnetic resonance imaging available for 3 affected sibs. All affected and 3 unaffected members from this family were genotyped using Illumina Human HumanCytoSNP-12 BeadChip (300K). The genotype data were delineated using easyLinkage Plus software platform, where Pedcheck was used to determine Mendelian genotyping errors and GeneHunter was used both to calculate multipoint lod scores under the assumption of autosomal recessive inheritance and to construct haplotypes. Analyses revealed a locus on chromosome 21q22 with a lod score of 3.86. The linkage peak encompassed a region of 16.64 cM between rs4816712 and rs2183593. Cystatin B (CSTB) gene was found to reside under this peak and has been screened for mutations in all members of this family, since dodecamer repeat expansions in the promoter region of this gene have been ascribed to the onset of a subtype of PME; Unverricht-Lundborg Disease (ULD). PCR amplification of the first exon and 5′ UTR of this gene has marked a lower threshold copy number for the onset of ULD. Our findings enabled the accurate diagnosis of this family as ULD patients, which were initially thought to be a PME family. This finding will be used as a preventative approach for future offsprings from the mutation carriers within this family.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2413T

***CRHR2* differentially modulates anxious temperament and plasma cortisol in a rhesus macaque model of childhood anxiety.** *G.L. Fawcett[1,2], M. Raveendran[2], J.A. Oler[4,5], A.S. Fox[6,7], S.E. Shelton[4,5], N.H. Kalin[4,5,6,7], J.M. Cheverud[8], D.M. Muzny[2], R.A. Gibbs[1,2], J. Rogers[1,2,3].* 1) Molecular and Human Genetics, Baylor College of Medicine, Houston, TX; 2) Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX; 3) Southwest National Primate Research Center, San Antonio, TX; 4) Department of Psychiatry, University of Wisconsin, Madison, WI; 5) HealthEmotions Research Institute, University of Wisconsin, Madison, WI; 6) Department of Psychology, University of Wisconsin, Madison, WI; 7) Waisman Laboratory for Brain Imaging and Behavior, University of Wisconsin, Madison, WI; 8) Department of Anatomy and Neurobiology, Washington University in St. Louis, St. Louis, MO.

Development of anxiety and depression are influenced by the corticotrophin releasing hormone (CRH) pathway. Genetic associations have been identified for clinical anxiety, depression, and suicide for *CRHR1*. Work in animal models has indicated a role for *CRHR2* in anxious temperament, although human studies provide conflicting evidence (Villafuerte, et al. 2002, Ishitobi, et al., 2012). We used a well validated rhesus macaque model to assess anxiety (n=441) by examining anxious temperament (AT), a composite phenotype including blood cortisol measurements, freezing behavior, and vocalization behavior among juveniles in response to a brief mildly stressful human intruder challenge. AT shows strong heritability, as do freezing and vocalization: freezing (h2=0.56, p=1x10–7), vocalization (h2=0.46, p=5x10–6), and AT (h2=0.50, p=1.8x10–6). Cortisol heritability was low but nearly significant (h2=0.11, p=0.06). We analyzed each of these traits using multiple regression analyses performed in SOLAR and using 48 LD-adjusted SNPs identified by exon and putative promoter targeted re-sequencing. For AT, we identified strong association of one non-synonymous SNP (p=0.015 (additive, dominance-minor)) in the first exon for *CRHR2* and of a non-synonymous SNP (p= 0.005 (additive)) in the last exon for *CRHR2*. We observed a highly complex interaction of *CRHR2* SNPs influencing AT and individually with each of the three component traits that comprise AT. We identified a complicated genetic architecture comprised of multiple contributing polymorphisms, with different SNPs in *CRHR2* influencing different aspects of the composite phenotype, but the primary contributions coming from effects on vocalization and/or freezing. Two SNPs associated with AT also contributed to variation in cortisol (*SNP44874* p=0.008 (additive) and *SNP49411* p=0.034 (dominance-major)). Further work will investigate other aspects of the genetic architecture of AT and its components, as well as examine the influence of *CRHR2* on brain metabolism in the hippocampus and amygdala.

2414F

Deregulation of neuronal cytoskeleton dynamics and androgen receptor gene pathways converge to aberrant synapse modulation in autism. *K. Griesi Oliveira[1], D. Y. Sunaga[1], L. A. Cruz[1], E. Vadasz[2], M. R. Passos-Bueno[1].* 1) Department of Genetics, Biosciences Inst,Univ Sao Paulo, Sao Paulo, SP, Brazil; 2) Psychiatry Institute, School of Medicine, University of Sao Paulo, Sao Paulo, Brazil.

The identification of autism spectrum disorders (ASD) causes is hampered by their genetic heterogeneity. However, it is evident that, despite genetic heterogeneity, the different genetic alterations leading to ASD are implicated in common molecular pathways or biological processes. In this scenario, the search for differentially expressed genes (DEGs) between ASD patients and controls is a good alternative to identify the molecular etiology of such disorders. Here, we employed genome-wide expression analysis to compare the transcriptome of dental pulp stem cell (DPSC) of idiopathic autistic patients (n=7) and control samples (n=6). A significant proportion of the identified DEGs are expressed in brain (p=0.001) and are involved in mechanisms previously associated to ASD such as protein synthesis, cytoskeleton regulation, cellular adhesion and alternative splicing. Androgen receptor (AR), several of its regulated genes, and CHD8, a gene that participates in AR-mediated transcriptional regulation, have their expression increased in the ASD patients. All these mechanisms ultimately converge to regulation of synapse formation and plasticity, which then could be considered the core alteration in ASD. Based on our results we suggest the use of ASD patients' stem cells transcriptome to better understand the pathways that lead to this disorder.

2415W

Calcium Channel Genes are Associated with Autism Spectrum Disorders. *A. Lu[1], X. Dai[1], R.M. Cantor[1,2].* 1) Department of Human Genetics and Center for Neurobehavioral Genetics, UCLA, Los Angeles, CA; 2) Department of Psychiatry, David Geffen School of Medicine, UCLA, Los Angeles, CA.

Alternation of synaptic homeostasis is a plausible biological process for developing Autism Spectrum Disorder (ASD). Calcium channel genes (CCG) are fundamental to this process, and evidence implicating these genes in ASD has been accumulating. Based on this, we conducted a targeted association study using existing GWAS data and imputation methods to assess potential roles for these genes. A panel of 2,176 SNPs (703 genotyped and 1473 imputed) in 10 CCG that each encodes alpha1 subunit proteins of distinct calcium channels, was tested for association with ASD in 543 multiplex Caucasian families (1,103 trios) from the Autism Genetics Resource Exchange (AGRE) and 1,651 multiplex and simplex Caucasian families (1,678 trios) from the Autism Genome Project (AGP). IMPUTE2 and a combined reference panel from the HapMap3 and the 1,000 Genomes Project were used to increase the coverage density. Family based association was tested using the FBAT software, which controls for population stratification and accounts for the non-independence of siblings within multiplex families. The level of significance was set to provide a Bonferroni correction for the size of this panel. Associated SNPs that exhibited an interaction with sex were tested using a case/pseudocontrol approach. It is striking that two significantly associated SNPs in CACNA1C, a gene with a known mutation that results in the Timothy Syndrome a Mendelian disorder which features of ASD, were identified among others in this panel of common SNPs. These associations provide additional support for the role of common CCG variants in ASD. They also indicate that sequencing studies for rare variants in these genes in ASD families are warranted.

2416T

X-chromosomal dystonia parkinsonism syndrome: Disease-specific sequence change DSC3 of *TAF1/DYT3* multiple transcript system affects dopamine metabolism, vesicular transport, and synapse function. *U. Mueller[1], D. Nolte[1], A. Hofmann[2], J. Schultze[2], T. Herzfeld[1].* 1) Inst Human Genetics, Justus-Liebig Univ, Giessen, Germany; 2) Life and Medical Sciences (LIMES)-Center, University of Bonn, Germany.

The X-chromosomal dystonia parkinsonism syndrome (XDP) is an adult-onset movement disorder. XDP is caused by changes within the *TAF1/DYT3* multiple transcript system. This system is composed of 43 exons, including the 38 exons of *TAF1* and 5 down-stream exons (d1–d5). The 43 exons are transcribed into multiple splice variants that are composed of combinations of*TAF1* and d exons. d exons 2–4 can also be transcribed independently under the control of a specific promoter. Several disease-specific changes were found within *TAF1/DYT3*. Most are located in introns including a SVA transposon in intron 32 of *TAF1*. However, one sequence change, DSC3, is located within an exon (d4). This raises the possibility of a role of DSC3 in the pathogenesis of XDP. We tested this hypothesis in a cell system. Neuroblastoma cells were transfected with expression constructs containing DSC3 (construct d2–d4/DSC3) and wild-type exons d2–d4 (construct d2–d4/wt). Expression profiling was performed. The effect of DSC3 on overall gene expression was dramatic. Expression of 362 genes differs between cells containing d2–d4/wt and cells containing d2–d4/DSC3. Annotation clustering was performed. The three annotation categories with the highest enrichment scores were identified. Annotation cluster 1 (enrichment score 3.32) included genes related to dopamine metabolism, vesicular transport, and synapse function. Twenty-six genes related to Ca^{++} metabolism and function were assigned to annotation cluster 2 (enrichment score 2.53). Genes responsive to stress, mainly mediated by reactive oxygen species (ROS) are grouped in annotation cluster 3. DSC3 suppresses physiologic regulation of these genes by wild-type d2–d4, possibly acting as a regulatory RNA. The data shows a major role of exons d2–d4 in the regulation of genes required for normal brain function. Interference with this role by DSC3 is a likely pathological mechanism in XDP. Disturbance of dopamine function and of Ca^{++} metabolism can explain the movement disorder phenotype; loss of protection against reactive oxygen species may account for the neurodegenerative changes found in XDP.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2417F

Identification of microduplications of the *ATRX* gene in Autism Spectrum Disorders. *B.A. Oliveira[1,2,3], I.C. Conceição[1,2,3], C. Correia[1,2,3], P. Jorge[4], R. Santos[4], C. Café[5], J. Almeida[5], S. Mouga[5,8], F. Duque[5,7], G. Oliveira[5,6,7,8], A.M. Vicente[1,2,3].* 1) Instituto Nacional de Saúde Dr. Ricardo Jorge, Lisbon, Portugal; 2) Center for Biodiversity, Functional & Integrative Genomics, Lisboa, Portugal; 3) Instituto Gulbenkian de Ciência, Oeiras, Portugal; 4) Centro de Genética Médica Doutor Jacinto de Magalhães, Porto, Portugal; 5) Unidade Neurodesenvolvimento e Autismo, Centro de Desenvolvimento, Hospital Pediátrico (HP), Centro Hospitalar e Universitário de Coimbra (CHUC), Coimbra, Portugal; 6) Centro de Investigação e Formação Clinica do HP-CHUC, Portugal; 7) Faculdade de Medicina da Universidade de Coimbra, Portugal; 8) Instituto Biomédico de Investigação em Luz e Imagem, Faculdade de Medicina da Universidade de Coimbra, Portugal.

X-linked syndromes and ASD are two very heterogeneous neurodevelopmental conditions that often present intellectual disability (ID) and autistic behaviours in common. Mutations and duplications of the *ATRX* gene, located on Xq21.1, are associated with the alpha-thalassemia/mental retardation syndrome, X-linked (ATR-X syndrome). Most *ATRX* mutations fall in the zinc finger or helicase functional domains, suggesting a link between chromatin remodeling, DNA methylation and gene expression in neurodevelopmental processes. The role of *ATRX* in normal brain may be related to cell survival during early neuronal differentiation, and it is a regulator of alpha-globin gene expression. In 342 Portuguese ASD probands, genotyped as part of a large genomic screening carried out by the Autism Genome Project (AGP) using the Illumina Infinium1M microarray, we noted a considerable number of microduplications including the *ATRX* gene. Such Copy Number Variants (CNVs) were not present in 4964 control subjects of European ancestry with no obvious psychiatric disease history, from available databases. After CNV validation by quantitative PCR, we confirmed a large duplication including *ATRX* in 10 ASD individuals, 2 females and 8 males. The two female probands did not present ID or major dysmorphic features and the CNV was *de novo* in one patient, while the other showed maternal inheritance. In this patient we observed a skewed X-inactivation (8:92) assessed by *HUMARA*assay, not present in the mother or the patient with a *de novo* duplication; origin of the inactivated *ATRX* allele is under study. The duplication in these female patients involved additional genes (*COX7B, PGAM4, MAGT1, TAF9B, ATP7A and PGK1*). In the eight male patients the duplication was much smaller, including only part of the *ATRX* gene, and was *de novo* in five subjects. The clinical presentation of the male patients was heterogeneous, with two individuals with moderate ID, and one showing dysmorphologies not characteristic of the classical ATR-X syndrome. A haematological phenotype is not obvious in any of the 10 patients, but manifestations of alpha-thalassemia may be subtle in individuals with *ATRX* mutations or structural alterations, and the presence of inclusion bodies was not assessed. Our data evidence microduplications of the *ATRX* gene as a likely etiology for ASD in approximately 3% of cases, thus broadening the clinical spectrum of known genetic alterations associated with the disease.

2418W

Alanyl-tRNA synthetase mutation in dominant distal hereditary motor neuropathy and Charcot-Marie-Tooth disease. *H. Takashima[1], A. Hashiguchi[1], J. Hu[2], Z. Zhao[1, 2], Y. Higuchi[1], Y. Sakiyama[1], S. Tokunaga[1], Y. Okamoto[1,3], A. Yoshimura[1].* 1) Dept Neurology and Geriatrics, Kagoshima Univ, Kagoshima, Japan; 2) Dept Neuromuscular disease, Third Hospital of Hebei Med Univ Shijiazhuang City, Hebei, China; 3) Dept Molecular Human Genetics Baylor college of Medicine, Houston, TX.

<Objective> Distal hereditary motor neuropathy (dHMN) is also known as distal spinal muscular atrophy or a variant of Charcot-Marie-Tooth disease (CMT).Four aminoacyl-tRNA synthetases (AARSs) have been implicated in CMT/dHMN. Although mutations in AARS cause axonal CMT, no published reports linking AARS mutations to the dHMN phenotype exist. To identify a new genetic cause of dHMN, which is also known as a variant of CMT.In addition, to screen the AARS mutations in dHMN and CMT patients. **Methods** We screened for the mutations of AARS gene in 450 inherited neuropathy patients. We also screened for the mutations of 28 CMT or related pathogenic genes using an originally designed microarray resequencing DNA Chip. The purpose-built GeneChip CustomSeq Resequencing Array (Affymetrix, Inc., Santa Clara, CA) was designed to screen for CMT and related diseases such as spinocerebellar ataxia with axonal neuropathy, and dHMN. To confirm the mutation revealed by our DNA chip, the proband and family members underwent genetic analysis by direct sequencing. **Results and Discussion** We found a mutation c.2677G>A (p.D893N) in a Chinese family with HMN phenotype. All 4 family members considered to be clinically affected proved to have the heterozygous AARS p.D893N mutation, while none of the 4 unaffected relatives harbored this mutation. We found 2 additional mutations (2 novel mutations)in Japanese CMT patients. **Conclusion** An AARS mutation caused dHMN in a Chinese family. AARS mutations result in not only a CMT phenotype but also a dHMN phenotype. AARS mutations were rare cause in Japanese CMT patients.

2419T

Association of interleukin 6 gene polymorphisms with genetic susceptibilities to spastic tetraplegia in males: a case-control study. *Q. Xing[1,5], M. Chen[1,5], T. Li[2], S. Lin[1,5], D. Bi[2], D. Zhu[2], Q. Shang[3], C. Ma[3], Q. Li[1,5], L. He[1,5], C. Zhu[2,4].* 1) Institute of Biomedical Science, Fudan University, Shanghai 200032, China; 2) Department of Pediatrics, the Third Affiliated Hospital of Zhengzhou University, Zhengzhou 450052, China; 3) Department of Pediatrics, Zhengzhou Children's Hospital, Zhengzhou 450053, China; 4) Center for Brain Repair and Rehabilitation, Institute of Neuroscience and Physiology, University of Gothenburg, Sweden; 5) Bio-X Center, Key Laboratory for the Genetics of Developmental and Neuropsychiatric Disorders, Shanghai Jiaotong University, Shanghai 200030, China.

OBJECTIVE: To explore the genetic association between the polymorphisms of the interleukin 6 gene (IL6) and cerebral palsy (CP) in Chinese population. METHODS: A total of 542 CP patients and 483 healthy control children were recruited in this study to detect 5 single nucleotide polymorphisms (rs1800796, rs2069837, rs2066992, rs2069840 and rs10242595) in the IL6 locus. Genotyping of SNPs was performed by the MassArray platform based genotyping approach. The SHEsis program was applied to analyze the genotyping data. RESULTS: Of the 5 selected SNPs, no significant allelic and genotypic association was found between CP patients and controls. However, subgroup analysis found significantly difference in allele frequencies between spastic tetraplegia in males compared with controls at rs1800796 (OR=1.39, P=0.033, P=0.099 after SNPSpD correction) and rs2069837 (OR=1.58, P=0.012, P=0.035 after SNPSpD correction). The frequencies of the C allele of rs1800796 and the A allele of rs2069837 were greater in males with spastic tetraplegia than in controls. The two SNPs haplotype rs1800796(G) - rs2069837(G) was also associated with a decreased risk of spastic tetraplegia in males (OR=0.619, P=0.009, P=0.027 after Bonferroni correction). CONCLUSION: These data suggest that the genetic variation of the IL6 gene may influence the susceptibility to spastic tetraplegia in males and its role in cerebral palsy deserves further evaluation in a large-scale and well-designed study.

2420F

***ADCYAP1R1* genotype associates with post-traumatic stress symptoms in highly-traumatized African American females.** *L.M. Almli[1], K.B. Mercer[1,2], H. Feng[3], K.N. Conneely[3,4], B. Bradley[1], K.J. Ressler[1,2].* 1) Department of Psychiatry and Behavioral Sciences, Emory University, Atlanta, GA; 2) Howard Hughes Medical Institute, Chevy Chase, MD; 3) Department of Biostatistics and Bioinformatics, Emory University, Atlanta, GA; 4) Department of Human Genetics, Emory University, Atlanta, GA.

Pituitary adenylate cyclase-activating polypeptide (PACAP) and its receptor (PAC1) may be critical mediators of abnormal processes following psychological trauma. Our previous work (Ressler et al., 2011, *Nature*, 470, 492–497) demonstrated that a variant, rs2267735, in the gene encoding PAC1 (*ADCYAP1R1*) is associated with post-traumatic stress disorder (PTSD) in a primarily African American cohort of highly traumatized females. We sought to replicate our previous finding in a replicate sample of 1161 adult African American males and females recruited from the clinics at Grady Memorial Hospital (Atlanta, Georgia). Salivary samples were collected to obtain DNA for genetic analysis. Self-reported psychiatric measures included the PTSD symptom scale (PSS), assessing the extent of PTSD symptom severity over the prior two weeks; Traumatic Events Inventory (TEI), assessing lifetime history of traumatic events; Beck Depression Inventory (BDI), assessing depressive symptoms, and Childhood Trauma Questionnaire (CTQ), assessing three types of child abuse: sexual, physical, and emotional. Genotyping of rs2267735 was conducted using Taqman and Sequenom platforms. Linear regression models were used to test for association under an additive (allelic) model where the number of copies of the “C” allele was allowed to influence the outcome variable. We found no significant main effect for *ADCYAP1R1* genotype on PTS severity in males or females, in contrast to our previous study. However, we found a genotype by trauma interaction on PTS severity in females ($p<0.001$) but not males ($p>0.1$). Moreover, this interaction was significant in females but not males after controlling for depression severity ($p<0.05$), child abuse ($p<0.05$), and depression and child abuse together ($p<0.05$). No significant effects of genotype (or interactions) were found when modeling depression severity when covarying for PTS severity ($p>0.1$), demonstrating the specificity of this variant for PTSD symptoms. In support of our earlier findings, the current results suggest that genetic variation at the *ADCYAP1R1* locus associates with PTS symptom severity, but not depression, among highly-traumatized African American females.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2421W

EFHC2 shows female-specific association with autism. *K.M. Tsang, Y. Cheng, J.F. Quinn, L.A. Weiss.* Department of Psychiatry, UCSF, San Francisco, CA.

Individuals diagnosed with autism spectrum disorders (ASD) suffer debilitating social impairment. Weiss et al. (2007) had previously identified EFHC2 as a QTL for social cognition among subjects with Turner syndrome (45, X). These females were presumed to serve as a model for males as they possess a single X-chromosome. We decided to examine the association of this gene in males and females with known social impairment in order to confirm the proposed sex-specific model. Two datasets of autism SNP genotyping data were used. The datasets considered for this analysis had any shared individuals removed from the replication data and were then sorted into trios featuring either affected males or affected females. Within each data-subset, all SNP markers mapping to the EFHC2 region were analyzed using a Transmission Disequilibrium Test. Linkage disequilibrium of 1000 Genomes CEU population data was used to find proxies for significant SNPs, allowing us to replicate our findings across different genotyping platforms. Our results showed association specific to females affected with ASD. We found within AGRE-NIMH SNP Data (Weiss et al. 2009) that two intronic SNPs in EFHC2 are associated with ASD in females (rs1181065 P<0.0012 OR=1.9 and rs9887166 P<0.0063 OR=0.48, N=310 for both markers) but not in the males of this same study (rs1181065 P>0.5, OR=0.94 Psex difference =0.0017 and rs9887166 P>0.09, OR=0.79, Psex difference =0.0153, N=1200 for both markers). In an independent autism sample (Wang et al. 2009), genotyped on a different platform utilizing different markers, we identified proxy markers for each of the associated SNPs in our discovery sample (r^2 >0.8). These two SNPs replicated our findings with significant P-values and consistent odds ratios in females (rs287779 P<0.005 OR=1.6 and rs5953381 P<0.004 OR=0.47, N=347 for both markers) as well as our insignificant findings in males (rs287779 P>0.6 OR=0.96 Psex difference =0.0062 and rs5953381 P>0.1 OR=0.8333 Psex difference =0.0094, N=1205 for both markers). These two sets of markers direct our attention to specific regions of EFHC2 for further investigation, as the second set of SNPs show moderate to high linkage disequilibrium with SNPs identified by Weiss et al. (2007) but the first set of SNPs are independent from the reported markers for social cognition in Turner syndrome.

2422T

Ataxia with Oculomotor Apraxia Type 2 in the Canadian Aboriginal Population: Expanding the Disease Distribution. *E.R. Dola[1], J.N. Hartley[1,3], F.A. Booth[2,4], A.A. Mhanni[1,2,3].*1) Genetics and Metabolism, Winnipeg Regional Health Authority, Winnipeg, Manitoba; 2) Department of Pediatrics and Child Health, University of Manitoba; 3) Department of Biochemistry and Medical Genetics, University of Manitoba; 4) Section of Pediatric Neurology, Health Sciences Centre, Winnipeg, Manitoba.

Ataxia with oculomotor apraxia type 2 (AOA2) (OMIM 606002) is an inherited autosomal recessive ataxia caused by mutations in the SETX gene, encoding senataxin, a large 2677 amino acid protein belonging to the DNA/RNA helicase super family. This disease was first described in Japanese and Pakistani families as well as ten French Canadian families known as the Quebec cluster. Here we report on two siblings from a consanguineous family of Canadian Aboriginal descent who presented with progressive ataxia, oculomotor apraxia, increased serum alpha fetoprotein and atrophy of the cerebellar vermis on magnetic resonance imaging. Based on the clinical phenotype, sequencing of the SETX gene was undertaken and revealed a homozygous disease causing mutation (c.1406A>G). This mutation has been previously reported in a consanguineous family from southern Italy. This report supports a pan ethnic occurrence of AOA2, hence this condition should be considered in the differential diagnosis in individuals with ataxia regardless of their ethnicity.

2423F

Lamin B1 is a novel regulator of myelin and mediates cell-autonomous neuropathology in a murine model of autosomal dominant leukodystrophy. *M. Y. Heng, L. J. Ptacek, Y. H. Fu.* Neurology, UCSF, San Francisco, CA.

Adult-onset Autosomal Dominant Leukodystrophy (ADLD) is a progressive and fatal neurological disorder characterized by early autonomic dysfunction, cognitive impairment, pyramidal tract and cerebellar dysfunction, and white matter loss in the CNS. ADLD is caused by duplication of the LMNB1 gene with increased Lamin B1 transcripts and protein expression (Coffeen et al., 2000; Padiath and Fu, 2010; Padiath et al., 2006). How over-expression of Lamin B1 leads to myelin defects is unknown. Here, we describe a novel ADLD mouse model that advances our understanding of the pathogenetic mechanism of Lamin B1 over-expression. These mice exhibit cognitive impairment and epilepsy, followed by age-dependent motor deficits. In addition, we found that Lamin B1 is a previously unrecognized regulator of the highly abundant myelin protein, PLP, through interaction with the transcription factor YY1. Mice selectively overexpressing Lamin B1 in oligodendrocytes exhibit marked motor deficits and myelin defects. These studies identify a mechanism by which Lamin B1 overexpression mediates oligodendrocyte cell-autonomous neuropathology in ADLD and implicate Lamin B1 as an important regulator of myelin formation and maintenance during aging.

2424W

Weight loss and low serum glucose levels in Machado-Joseph disease/Spinocerebellar ataxia type 3 (MJD/SCA3). *L.B. Jardim[1, 3, 4, 6], G.N. Souza[1], A.D. Russo[1], T.L. Monte[2], R.M. Castilhos[1], A.F.S. Schuh[2, 4], K.C. Donis[1], R. D Avila[1], T.C. Gheno[4, 7], G.V. Furtado[4, 7], D.O.G. Souza[5], M.L. Saraiva-Pereira[1, 4, 5, 7], L.V.C. Portela[5], C.R.M. Rieder[2, 3], J. Dubois[5], J.A.M. Saute[1, 3].* 1) Medical Genetics Service, Hospital de Clinicas de Porto Alegre, Porto Alegre, Brazil; 2) Neurology Service, Hospital de Clinicas de Porto Alegre, Porto Alegre Brazil; 3) Post-Graduation Program in Medical Sciences, Universidade Federal do Rio Grande do Sul, Brazil; 4) Post-Graduation Program in Genetics and Molecular Biology, Universidade Federal do Rio Grande do Sul, Brazil; 5) Department of Biochemistry, Universidade Federal do Rio Grande do Sul, Brazil; 6) Department of Internal Medicine, Universidade Federal do Rio Grande do Sul, Brazil; 7) Genetic Identification Laboratory, Hospital de Clinicas de Porto Alegre, Brazil.

Machado-Joseph disease, or spinocerebellar ataxia type 3 (MJD/SCA3), caused by a CAG repeat expansion (CAGn) at ATXN3, is usually related to a pure neurological phenotype. We have recently described a low body mass index (BMI) and high insulin sensitivity in MJD/SCA3, either directly related to age at onset (AO) or to the CAGn, in this disorder. **Aims**: to study the body mass composition and fasting glucose levels (FG) (as an indirect measurement of insulin sensitivity), and to correlate them with CAGn and clinical variables of disease. **Methods**: antropometry, bio-impedanciometry and glicemic curves, as well as SARA and NESSCA scales, were performed in 44 MJD/SCA3 patients and in 41 healthy controls. **Results**: patients and controls did not differ in age and gender proportions. MJD/SCA3 patients had a m±sd of 35±10 years of AO, 5.7±2 years of disease duration, and 75±3 CAGn. The following data (m±sd) were obtained in cases and in controls: BMI of 24.3±4.3 and 27.4±6.5kg/m2 (p=0.011), lean mass of 74.3±8.5% and 71.9±8.6% (ns), basal metabolic rate (BMR) of 23.2±2.6 and 22.4±2.6cal/kg (ns), FG of 89.2±8.8 and 94.6±8.2mg/dL (p=0.005) and after an oral glucose tolerance test (GTT, 120min), of 88.3±28 and 105.6±32.6mg/dL (p=0.011). Four patients presented symptomatic hypoglycemia at GTT(with <50mg/dL). BMI was significantly correlated to CAGn (R=-0.475, p=0.001), AO (R=0.485, p=0.001), NESSCA (R=-0.35, p=0.03), but not to sub-item of NESSCA dysphagia, or to SARA. On regression analysis, the main factors determining BMI were CAGn and chronological age. On regression analysis, GTT was determined only by AO. **Discussion**: The loss of weight affects both lean and fat body masses and is detectable since the first years of disease. We have confirmed that low BMI was mainly related to CAGn. GTT levels, used to estimate insulin sensitivity, was partially related to AO. Therefore, whereas weight of loss seem to be a primary manifestation in MJD/SCA3, the increase in peripheral sensitivity to insulin rest to be explained. We speculate if they might act as modifiers of the neurological progression, in MJD/SCA3.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2425T

Spinocerebellar ataxias with hypogonadism: Unraveling a rare group of neuroendocrinologic syndromes. *C. Lourenco[1], C. Sobreira[1], V. Cintra[1], R. Giugliani[2], W. Marques, Jr.[1].*1) Neurology, Univ Sao Paulo, Ribeirao Preto, Brazil; 2) Medical Genetics Service, UFRGS, Porto Alegre, Brazil.

Background: The association between cerebellar ataxia and hypogonadism was initially described at the beginning of the XX century in four sibs by Holmes, and named after him "Gordon Holmes syndrome" (OMIM%212840). At that time, it was not possible to determine whether hypogonadism was hypogonadotropic or hypergonadotropic. Although the clinical picture may be similar, the pathogenic mechanisms involved in the two types of hypogonadism are probably different. Here, we present clinical and molecular/biochemical data on nineteen Brazilian patients with cerebellar ataxia and hypogonadism. Material and Methods: Patients were clinically evaluated at our Neurogenetics clinics . Brain MRI, ophtalmological and audiological exams, EMG/NCV studies, hormone and biochemical tests, screening for CDG, karyotype, muscle biopsy with chain respiratory enzyme assays and measurement of coenzyme Q10, molecular tests for Friedreich ataxia and for the most common forms of SCAs were performed. Results: All patients had cerebellar ataxia as the main reason for referral, with a variable age at onset, but most of them (17 out of 19) developed neurologic symptons before 20 years of age. Parents' consanguinity was present in two families; 13 patients had hypergonadotropic hypogononadism.Molecular tests for Friedreich and SCAs were normal in all patients. Optic atrophy and retinochoroidal degeneration were found in five patients; motor axonal neuropathy was present in four patients; two patients had features of lower motor neuron disease in their EMG studies. Cerebellar atrophy with pons or prominent vermis involvement was a constant feature in 16 patients; 3 patients also had leukodystrophy detected at MRI. One family with 4 affected patients have features consistent with a rare neurological disorder, Boucher-Neuhäuser syndrome ; three other unrelated patients had coenzyme Q10 deficiency detected in muscle biopsy; Congenital Disorder of Glycosylation type Ia was identified in one adult patient. The remaining patients had features that may fit in the Gordon-Holmes and Perrault phenotype (ovarian failure with deafness). Conclusions: The association between cerebellar ataxia and hypogonadism comprise hetereogenous entities and better characterize for the first time this disorder in the Brazilian population. Clinical investigation of further cases with this association may help to define the pathological basis of these fascinating neuroendocrinological syndromes.

2426F

A neuronal cellular model of 22q11.2 DS exhibits disruptions to the miRNA regulatory pathway and may increase the rate of schizophrenia in individuals with 22q11.2 DS. *W. Manley[1], M.R. Ababon[2], P.G. Matteson[2], J.H. Millonig[2], L.M. Brzustowicz[1].* 1) Department of Genetics, Rutgers University, Piscataway, New Jersey 08854–8095, USA; 2) Center for Advanced Biotechnology and Medicine, UMDNJ, Piscataway, New Jersey 08854.

Schizophrenia is a complex and poorly understood disease caused by the interplay of many environmental and genetic factors. The 22q11.2 deletion syndrome is a disorder that is caused by the microdeletion of part of chromosome 22 leading to a 25–fold greater chance of developing schizophrenia in affected individuals versus the general population. The missing 22q11.2 region contains the DGCR8 gene, which is required for the initial step of miRNA biogenesis. However, the 22q11.2 deletion itself is not sufficient to cause schizophrenia, since 75% of individuals with this deletion do not develop the disease. We hypothesize that the 22q11.2 deletion increases the risk of schizophrenia through alterations in miRNA regulatory networks via depletion of multiple miRNAs. These miRNAs may serve as a protective buffer against the accumulation of deleterious mutations at other schizophrenia risk loci. We have developed a neuronal human cellular model of disruptions to the miRNA regulatory network caused as a result of DGCR8 deficiency. The neuroblastoma cell line SH-SY5Y was transfected via electroporation with GIPZ lentiviral shRNAmir constructs specific for DGCR8 (Open Biosystems). The GIPZ constructs were prepared using PureYield Plasmid MaxiPrep System (Promega) and sequenced (GeneWiz) prior to use. For selection purposes, all constructs contained genes for GFP and puromycin resistance. The levels of DGCR8 gene expression were quantified using the 7900HT Fast Real-Time PCR System (Applied Biosystems) along with Taqman Gene Expression Assays to ensure DGCR8 reduction. We have begun to characterize disruptions to the miRNA regulatory network in this SH-SY5Y cell line transfected with the GIPZ constructs specific for DGCR8 using Taqman Array Human Microarray Cards Version 3.0. Here we present the miRNAs that we have identified to be differentially expressed in transfected cells versus untransfected control cells. We predict that these miRNAs could be involved in the elevated risk of developing schizophrenia in individuals with 22q11.2 deletion syndrome versus the general population. Future studies will focus on determining the biological targets of these miRNAs using miRNA target prediction software.

2427W

Analysis of aging phenotypes from long-term expression of the progeria mutation in the brain. *J. Baek[1], E. Schmidt[1], S. Rodriguez[1], C. Strandgren[1], B. Rozell[2], E. Mugnaini[3], M. Eriksson[1].* 1) Department of Biosciences and Nutrition, Karolinska Institutet, Huddinge, Sweden; 2) Department of Veterinary Disease Biology, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark; 3) Department of Cell Molecular Biology, Feinberg School of Medicine, Northwestern University, Chicago, USA.

Hutchinson-Gilford progeria syndrome (HGPS, also known as progeria) is an extremely rare genetic disorder characterised by an early onset of premature aging. The affected individuals display no signs of disease at birth, but within their first years of life, they gradually develop an appearance often referred to as aged-like. The most common genetic defect for progeria has been found to be a de novo heterozygote point mutation in exon 11 of the LMNA gene (c.1824C>T; p.G608G). This mutation causes an increased activation of a cryptic splice site, which in turn leads to an aberrant splice product of lamin A, called progerin or lamin AΔ50. Although extensive research has been carried out on progeria due to its potential role as a model for human aging, the focus has been mainly on the most profoundly affected tissues and organs. Interestingly, previous clinical neurological examinations of progeria patients have shown that they have no intellectual problems nor show any signs of neurological defects, which may suggest that the brain is protected from this lethal disease. Therefore, the aim of this study was to investigate the effects of long-term progerin expression by using an inducible, tissue specific transgenic mice model with expression of the most common HGPS mutation, LMNA c.1824C>T; p.G608G, in the brain. Haematoxylin and Eosin, and Fluoro-Jade C stainings were carried out in order to examine histopathological changes and the presence of neurodegeneration in the brain, respectively. In addition, immunofluorescence was carried out to determine the localisation of lamin A/C and B expression in the brain. The findings from this study are of importance as the long-term effects of progerin expression on the brain have never been investigated. Moreover, by better understanding the mechanism(s) of this premature aging, it may be possible to minimise the pathological process observed in progeria, and also to develop potential anti-aging treatments that can restore declined physiological functions associated with normal aging.

2428T

Neurodegeneration and associated phenotypes in adults with 22q11.2 deletion syndrome. *N.J. Butcher[1], T.-R. Kiehl[2], E. Chow[1 3], E. Rogaeva[4], A.E. Lang[5 6], A.S. Bassett[1 3].* 1) Clinical Genetics Research Program, Centre for Addiction and Mental Health, Toronto, Ontario, Canada; 2) Department of Pathology, University Health Network, Toronto, Ontario, Canada; 3) Department of Psychiatry, University of Toronto, Toronto, Ontario, Canada; 4) Department of Medicine, Centre for Research in Neurodegenerative Diseases, University of Toronto, Toronto, Ontario, Canada; 5) Toronto Western Hospital Research Institute, University Health Network, Toronto, Ontario, Canada; 6) Department of Neurology, University of Toronto, Toronto, Ontario, Canada.

Background: 22q11.2 deletion syndrome (22q11.2DS) is a multisystem genetic syndrome (~1/4000) associated with a hemizygous 22q11.2 deletion. Features of 22q11.2DS include congenital and later onset conditions, including birth defects, intellectual disability, and schizophrenia. Case reports of L-dopa responsive parkinsonism in individuals with 22q11.2 deletions indicate that neurodegenerative disease may be a late-onset feature of 22q11.2DS. The clinical and neuropathological features of neurodegenerative diseases in later adulthood in adults with 22q11.2DS remain largely uncharacterized. Methods: We assessed older adults (aged 30 to 62 years) in a well-characterized cohort of adults with 22q11.2DS for clinical (motor, cognitive, and olfactory), genetic, and neuropathological features of neurodegenerative disease. We reviewed medical and family history for diagnoses of neurodegenerative diseases. Postmortem brain tissue was investigated in five patients using immunohistochemistry for markers of neurodegenerative processes including neural degeneration and abnormal aggregates such as Lewy bodies and amyloid plaques. Results: Compared with the general population, adults with 22q11.2DS (≥50 years) had an elevated prevalence of neurodegenerative disease ($p<0.006$) that could not be attributed to familial genetic risk factors. Motor, cognitive and olfactory deficits previously associated with neurodegenerative diseases in the general population were common. Neuropathological studies of available brain tissue confirmed antemortem diagnoses of Lewy body associated Parkinson disease in two patients. Conclusions: Adults with 22q11.2DS appear to be at elevated risk of developing neurodegenerative disease in mid-adulthood. Side effects from antipsychotic medications and concomitant features of 22q11.2DS may initially mask presenting features and complicate diagnosis in some patients. Functional neuroimaging may help identify patients in asymptomatic or early stages of the disease. While clinical monitoring of adults with 22q11.2DS for neurodegenerative conditions may be indicated, further characterization of the link between Parkinson's disease and 22q11.2DS is needed to help to inform clinical management of adults with 22q11.2DS. Association of 22q11.2 deletions with Parkinson's and other neurodegenerative diseases may provide important new insights into the genetic architecture and pathogensis of these diseases.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2429F

CXCR4 (Chemokine Receptor) Expression Analysis In Chronic Cuprizone Mouse Model of Multiple Sclerosis. *G. Celikyapi[1], T. Avsar[1], G. Terzioglu[1], S. Uyar Bozkurt[2], G. Yanikkaya Demirel[3], A. Siva[4], T. Altug[5], E. Tahir Turanli[1,6].* 1) Molecular Biology and Genetics, Istanbul Technical University, Istanbul, Turkey; 2) Marmara University, Neurological Sciences Institute, Pathology Lab, Istanbul, Turkey, Basibüyük, 34840; 3) Yeditepe University, School of Medicine, Medical Microbiology Department, Istanbul, Turkey, Kayisdagi, 34755; 4) Istanbul University, Cerrahpasa School of Medicine, Neurology Department, Istanbul, Turkey, Fatih, 34098; 5) Istanbul Bilim University, Medical Biology and Genetics Department, Istanbul, Turkey, Gayrettepe, 34349; 6) Istanbul Technical University, Molecular Biology and Genetics Department, Istanbul, Turkey, Maslak 34469.

OBJECTIVE: Multiple sclerosis (MS) is a chronic, inflammatory and demyelinating disease which affects the central nervous system. Activation of CXCR4 chemokine receptor is associated with regulation of cell proliferation, migration, maturation and survival in the CNS. Cuprizone is a toxin which has been shown to cause demyelination of major myelinated tracts such as corpus callosum (CC) and cerebellar peduncles of animal models. Demyelination is followed by repair of myelin when cuprizone exposure is ended, but longer feeding periods (> 6 week) lead to chronic demyelination. This study focuses on the changes in CXCR4 expression in brain, which may enlighten the relation between CXCR4 expression and chronic demyelination. METHODS: 80 C57BL/6J mice were fed with chow including 0.2 % cuprizone for 12 weeks. Control mice fed with regular mice chow without Cuprizone. Cuprizone has been removed from the diet in the following 4 weeks. 5 mice were sacrificed biweekly, brain tissues have been obtained and RNA isolation was performed from the corpus callosum area. Expression changes of CXCR4 gene were analyzed via quantitative Real Time PCR with Gapdh and Pgk1 as internal control genes. RESULTS: CXCR4 expresion levels increased between the 3–6 week of cuprizone diet (p=0.0011), whereas it decreased between 6–12 weeks (p<0.001), and slowly increased after the 12. week compared to controls. CONCLUSION: Oligodendrocytes are known to build the myelin sheath and CXCR4 gene is directly related with maturation of Oligodencyte Progenitor Cells (OPC) to oligodendrocytes. In this study we have shown that, CXCR4 expression levels are decreasing with longer (>6week) cuprizone diet. Our results revealed that remyelination process is damaged at early stages of chronic cuprizone diet.

2430W

Increased error-related brain activity in youth with obsessive-compulsive disorder and unaffected siblings. *G. Hanna[1], M. Carrasco[2], S. Harbin[1], J. Nienhuis[1], K. Fitzgerald[1], W. Gehring[2].* 1) Dept Psychiatry, Univ Michigan, Ann Arbor, MI; 2) Dept Psychology, Univ Michigan, Ann Arbor, MI.

Background: The pathophysiology of obsessive-compulsive disorder (OCD) involves increased activity in cortico-striatal circuits connecting the anterior cingulate cortex (ACC) with other brain regions. The error-related negativity (ERN) is a negative deflection in the event-related potential following an erroneous response and is thought to reflect ACC activity. This study was done to assess the ERN as a biomarker for OCD by comparing ERN amplitudes in pediatric OCD patients, unaffected siblings of pediatric OCD patients, and healthy controls. Methods: The ERN and correct response negativity (CRN) were measured during an Eriksen flanker task to assess performance monitoring in 40 youth with a lifetime diagnosis of OCD, 19 unaffected siblings of OCD patients, and 40 unrelated healthy comparison subjects ranging in age from 10 to 17 years. ERN and CRN amplitudes were compared between groups using linear regression by the generalized estimating equation method to account for correlated data. Results: Compared to healthy controls, ERN amplitude was significantly increased in both pediatric OCD patients and unaffected siblings. There were no significant group differences in CRN amplitude. ERN amplitude in patients was unrelated to OCD symptom severity, current diagnostic status, or treatment effects. Conclusions: Increased error-related brain potentials were observed not only in pediatric OCD patients but also in unaffected siblings. The results provide evidence that enhanced error-related brain activity may serve as a biomarker for OCD in youth that is independent of the presence of clinical symptoms. The ERN may be a useful quantitative phenotype in genetic linkage and association studies of this complex trait.

2431T

Bipolarity in children with dyslexia. *C. C. Mignot[1], F. X. Coude[1,2].* 1) Centre de Pediatrie, Aix-en-Provence, paca, France; 2) Departement de Genetique,Hopital Necker,Paris,France.

OBJECTIVE:Reading impairment(Dyslexia) is a common problem at school. It is present in 8–10% of children in the general population. Learning impairment in children with dyslexia seems to overpass the consequences of reading difficulties and persisted even after partial correction of these difficulties.The purpose of our work was to determine if bipolarity might account for this gap. METHOD: Forty two children (25 males, 17 females; mean age 9.6 years, SD 1.8) diagnosed with Dyslexia were studied. Fifty control children without learning disorders were recruted in the same schools. Executive dysfunction(ED) and bipolariy disorder(BD) indexes were determined using the parental Bipolar Child Questionnaire. Learning impairment(LI) index was determined using the teacher Vanderbilt questionnaire. Leximetry was measured using the french scale of Lalouette. RESULTS: ED was significantly more frequent in children with dyslexia (10.2 vs 1.3, P<0.01) and BD was also significantly elevated in these patients (45.3 vs 20,2, P <0.01). BD and ED indexes were highly correlated. CONCLUSION: Incidence of LI at school was much more dependent upon the level of ED presented by the children than of the reading defect level. Moreover the ED Index is highly dependent of the level of the BD index which means that not only Bipolarity seems frequent in these children and may be the cause, at least partially, of the academic impairment observed at school in children with dyslexia.

2432F

Sequential processing deficit as a cognitive endophenotype in a multigenerational family with a severe speech sound disorder. *B. Peter[1], K. Chapman[2], W. Raskind[3, 4].* 1) Speech & Hearing Sci, Univ Washington, Seattle, WA; 2) Communication Sciences and Disorders, University of Utah, Salt Lake City, UT; 3) Medicine, University of Washington, Seattle, WA; 4) Psychiatry and Behavioral Sciences. University of Washington, Seattle, WA.

The purpose of this study was to investigate deficits in sequential processing as an endophenotype in severe speech sound disorder (SSD). SSD is a childhood disorder of impaired ability to acquire clearly understandable speech. A proposed subtype is childhood apraxia of speech (CAS), thought to interfere with motor programming and/or planning. Our previous studies support a genetic etiology of CAS, Mendelian inheritance, and a central deficit in sequential motor processes. Here, we investigate sequential processing deficits in domains beyond motor processing in individuals with familial CAS. Of 10 adults and 14 children from a three-generational family (age 15 months to 66 years), 3 adults and 6 children had past or present SSD. Participants completed part or all of a protocol covering articulation and phonology; tasks requiring sequential processing (alternating syllable production and key presses, imitation of multisyllabic real words, nonwords, and tongue twisters, reading nonwords), and tasks involving less sequential processing (repetitive syllable production and key presses, verbal and nonverbal cognition, sentence imitation, rapid automatic naming, reading real words). During articulation testing, two 3–year-old cousins showed CAS-like speech characterized by mis-sequenced speech sounds (e.g., “boos”/spoon, “wipseen”/swimming), vowel errors (e.g., “wawos”/flowers, “doi”/car), and unusual speech sound substitutions (e.g., “thau”/clown, “fiwee”/crying). A 6–year-old girl and a 7–year-old boy had residual speech sound errors. In all other affected individuals, single-word testing and conversational speech showed normalized speech. Consistent with previous results, deficits in alternating syllables and key presses patterned with SSD affectation. Deficits in the multisyllabic real word, nonword, and tongue twister imitation tasks were seen in all but two affected individuals and one married-in person but not in the remaining unaffected individuals. One affected adult and one affected child showed deficits in nonword reading. Conversely, only three affected individuals showed deficits in any tasks with low sequencing loads. These results are consistent with the hypothesis that sequential processing deficits underlie speech development in CAS as well as motor and linguistic abilities. Next steps are to identify the causal gene in this family with a combination of traditional linkage analysis and next-generation sequencing.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2433W

Investigation of the role of serotonin and histamine pathways in the etiology of Tourette's syndrome. *N. Sun[1,2], GA. Heiman[1], L. Deng[1], M. Sheldon[1,3], JC. Moore[1,3], AJ. Toro-Ramos[3], RP. Hart[3,4], Z. Pang[5], JA. Tischfield[1,3].* 1) Department of Genetics, Rutgers University, Piscataway, NJ; 2) Microbiology and Molecular Genetics Program, Rutgers University, Piscataway, NJ; 3) Rutgers University Cell & DNA Repository's NIMH Stem Cell Center, Piscataway, NJ; 4) Department of Cell Biology and Neuroscience, Rutgers University, Piscataway, NJ; 5) Child Health Institute of New Jersey, New Brunswick, NJ.

Tourette's syndrome (TS) is a childhood onset neurodevelopmental and genetic disorder characterized by chronic motor and phonic tics. The etiology of TS is complex and believed to be a result of interactions between multiple genetic and environmental factors. Despite many years of research, identifying risk alleles has been slow. Recent studies have implicated several neurotransmission pathways in the etiology of TS including dopamine, serotonin and histamine pathways. We collected whole blood samples from all five family members of a two-generation pedigree from New Jersey Center for Tourette Syndrome Sharing Repository. We are establishing induced pluripotent stem cell (iPSC) lines from each individual's T cells. We are differentiating individual-specific iPSCs into serotonergic neurons by forced expression of certain transcription factors, for example, Ascl1, Lmx1b and Foxa2 in iPSCs, thus developing neuron-based TS models in vitro. Also, we plan to generate iPSC lines from another TS family, all TS individuals from which have W317X nonsense mutation in their L-histidine decarboxylase (*HDC*) gene. Further differentiation of *HDC* mutant family's iPSCs into histaminergic neurons will enable us to model TS in a histamine pathway-specific manner. Comparison of DNA microarray-generated gene expression patterns of neurons from controls and TS patients within both families may reveal aberrant gene expression in serotonin or histamine neurotransmission pathway in TS and this may lead to the discovery of mutant or epigenetically-modified genes involved in disease causation.

2434T

Is DS22q11.2 distal a frequent etiology of Mathematical Learning disability? *M.R.S. Carvalho[1], G.S. Vianna[1,5], L.F.S. Oliveira[2], B.C. Moreira[1], R. Sturzeneker[5], M.J.B. Aguiar[3], L. Leão[4], G. Paskulin[6], P. Zen[6], V.G. Haase[2].* 1) Departamento de Biologia Geral, ICB, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil; 2) Departamento de Psicologia, FAFICH, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil; 3) Departamento de Pediatria, Faculdade de Medicina, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil; 4) Hospital das Clínicas, Faculdade de Medicina, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil; 5) Laboratório de Diagnóstico Molecular da BIOCOD BIOTECNOLOGIA; 6) Universidade Federal de Ciências da Saúde de Porto Alegre.

Approximately 6% of the school-aged children have low math achievement (MD). Neurogenetic etiology has been suggested due to the presence of MD in some genetic syndromes such as Williams, SD22q11.2, Turner, Sotos, Silver-Russell, and in girls FMR1. MD is also prevalent among persons with fetal alcohol syndrome. Among the genetic conditions SD22q11.2 is the most common. Approximately, 90% of the patients are hemyzygous for microdeletions spanning a 3Mb chromosomal segment and almost all the remainder of the cases result from a 1.5–2 Mb microdeletion. Atypical microdeletions, larger or smaller, overlapping or not the 3Mb interval, have also been occasionally found. This chromosomal segment is prone to rearrangements due to the presence of a set of eight interspersed repeats spanning, the LCR22a-g, which favor non-allelic homologous recombination. Among 22q11.2DS children with normal intelligence, learning disability is a frequent finding with a pattern of poorer performance in arithmetical than in reading and spelling tasks. No studies have been developed in order to investigate the contribution of 22q11.2DS to MD. This is a population-based study aiming to ascertain the frequency of 22q11.2DS among school children with low Mathematical achievement with or without spelling disability. Prior to the starting, the study was approved by the Ethics in Research Committee of the University. Children were ascertained in the schools through a two step approach. Initially, all the children were submitted to a language + math screening test. Those children scoring bellow the 25 percentile in their classes, as well as normal control individuals, were invited to a second phase, which included intelligence, subtizing, reaction time, Nine-holes Peg test, Rey Complex Figure, Mini-mental state, and an evaluation of math and language functions. Presence of 22q11.2 deletions was ascertained with SALSA P250 MLPA kit and by a multiplex qPCR covering TBX1 and UFD1L genes. One out 75 subjects in MD group had an atypical deletion spanning the LCR22 D-E interval corresponding to 22q11.2 distal deletion syndrome. This deletion was not present among the control individuals. These results suggest that DS22q11.2 distal is one the most common genetic etiologies for MD. Grants: CNPq, FAPEMIG, PPSUS/FAPEMIG, CAPES.

2435F

DNA methylation alterations in first-episode schizophrenia patients: comprehensive analysis using peripheral blood cells. *M. Nishioka[1, 2], M. Bundo[1], S. Koike[2], C. Kakiuchi[2], T. Araki[3], K. Kasai[2], K. Iwamoto[1].* 1) Department of Molecular Psychiatry, The University of Tokyo, Tokyo, Japan; 2) Department of Neuropsychiatry, The University of Tokyo, Tokyo, Japan; 3) Department of Youth Mental Health, The University of Tokyo, Tokyo, Japan.

Genetic studies on schizophrenia have identified promising candidate genes, but the etiology of schizophrenia was not fully elucidated. Epigenetic alterations could be an additional explanation for its pathophysiological mechanism and several groups have reported the epigenetic alternations in schizophrenia patients. We analyzed DNA methylation status in the peripheral blood cells derived from 18 first-episode schizophrenia patients and 15 healthy controls. The patients were strictly confined to those at the stage of first-episode psychosis, because we hypothesized that epigenetic alterations were different between disease stages. DNA methylation profiles on 27,578 CpG sites were assayed with the Illumina Infinium HumanMethylation27 beads array. The assayed CpG sites were located mainly on the promoter regions, covering 14,475 genes. We identified differentially methylated CpG sites in the first-episode schizophrenia patients using SAM software. We found that DNA methylation alternations preferentially occurred at inside of the CpG island. Ontology analysis revealed that they were enriched in the genes related to nuclear lumen, transcription and nucleotide binding. We also observed differential DNA methylation on the promoters of *HTR1E* and *COMTD1*, which have functional relation with the genes reported in the previous epigenetic studies. Our study suggests that first-episode schizophrenia patients have site-specific epigenetic alterations, which could have functional significance such as regulation of gene expression, in blood cells.

2436W

Clinical and molecular genetic analysis in Ullrich congenital muscular dystrophy with dominant and recessive inheritance. *B. Bozorgmehr[1], A. Kariminejad[1], S. Nafisi[2], U. Andoni[3], C. Gartioux[4], C. Ledevil[8], V. Allamand[4,5,6,7], P. Richard[8], M. Kariminejad[1].* 1) Dept Clinical Genetics, K-N Pathology & Genetics Ctr, Tehran, Iran; 2) Shariati Hospital, Tehran-Iran; 3) Hospital Marin-Paris France; 4) UPMC Univ Paris 06, IFR14, Paris, F-75013, France; 5) CNRS, UMR7215, Paris, F-75013, France; 6) Inserm, U974, Paris, F-75013, France; 7) Institut de Myologie, Paris, F-75013, France; 8) AP-HP, Groupe Hospitalier Pitié-Salpêtrière, UF Cardiogénétique et Myogénétique, Service de Biochimie Métabolique, Paris, F-75013, France.

: Ullrich congenital muscular dystrophy (UCMD) corresponds to the severe end of the clinical spectrum of neuromuscular disorders due to mutations in the genes encoding collagen VI. It is characterized by early onset proximal joint contractures, muscle weakness and hyperlaxity of distal joints. It had long been considered a recessive condition with homozygous and compound heterozygous mutations in the COL6A1, COL6A2, and COL6A3 genes. However, patients harbouring heterozygous deletions have been increasingly reported over the last years.We investigated four unrelated Iranian families with six affected children with typical Ullrich Congenital Muscular Dystrophy (UCMD). Two affected sibs carried a homozygous nonsense mutation in exon 12 of COL6A2, while the other patient harboured a large heterozygous deletion of exons 5–8 in COL6A2. The two other affected sibs had homozygote mutation in exon 24 of COL6 A2 and the last was homozygote in COL6 A1.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2437T

Epigenomes landscaped by histone H3K9me3 contribute to the pathology of Huntington's disease. *J. Shin[1], J. Lee[2,3], Y. Hwang[4], W. Lee[1], J. Wie[5], K. Kim[4], M. Lee[6], D. Hwang[6], R. Ratan[7], N. Kowall[2,3], I. So[5], J. Kim[1], H. Ryu[2,3,4].* 1) Genome Medicine Institute, Seoul National University College of Medicine, Seoul, South Korea; 2) Veteran's Affairs Boston Healthcare System, Boston, MA 02130, USA; 3) Department of Neurology and Boston University Alzheimer's Disease Center, Boston University School of Medicine, Boston, MA 02118, USA; 4) WCU Neurocytomics Group,Seoul National University College of Medicine, Seoul, South Korea; 5) Department of Physiology and Biomedical Sciences, Seoul National University College of Medicine, Seoul, South Korea; 6) School of Interdisciplinary Bioscience and Bioengineering and Department of Chemical Engineering, POSTECH, Pohang, South Korea; 7) Burke Medical Research Institute, White Plains, NY, USA.

Huntington's disease (HD) is an autosomal dominant neurodegenerative disease caused by an expanded trinucleotide CAG repeat in the gene coding for huntingtin (Htt). Deregulation of chromatin remodeling is linked to the pathogenesis of HD but the mechanism remains elusive. In order to identify what genomes are deregulated by trimethylated histone H3K9 (H3K9me3)-dependent heterochromatin, we performed H3K9me3–ChIP genome-wide sequencing combined with RNA-sequencing followed by platform integration analysis. We found that genomes involving neuronal synaptic transmission including cholinergic receptor M1 (CHRM1), cell motility, and neuronal differentiation pathways are down regulated while their promoter regions are highly occupied with H3K9me3 in HD. Moreover, we validated that CHRM1 gene repression by H3K9me3 impairs Ca2+-dependent neuronal signal transduction in HD. Thus, our data indicates that the epigenetic modifications, such as aberrant H3K9me3–dependent heterochromatin plasticity, directly contribute to the pathogenesis of HD.

2438F

Novel PRRT2 Mutation in an African-American Family with Paroxysmal Kinesigenic Dyskinesia. *M.S. LeDoux[1], J. Xiao[1], P. Hedera[2], A. Puschmann[3], D. Momc{{ak}}ilović[4], S.W. Wu[5].* 1) Neurology, University of Tennessee Health Science Center, Memphis, TN. United States; 2) Neurology, Vanderbilt University, Nashville, TN, United States; 3) Neurology, Skåne University Hospital and Department of Neurology, Lund University, Sweden; 4) Clinic for Child Neurology and Psychiatry, Medical Faculty University of Belgrade, Belgrade, Serbia; 5) Pediatrics, University of Cincinnati, Cincinnati, OH, United States.

Recently, heterozygous mutations in *PRRT2* (Chr 16p11.2) have been identified in Han Chinese, Japanese and Caucasians with paroxysmal kinesigenic dyskinesia. In previous work, a paroxysmal kinesigenic dyskinesia locus was mapped to Chr 16p11.2 - q11.2 in a multiplex African-American family. Sanger sequencing was used to analyze all four *PRRT2* exons for sequence variants in 13 probands (9 Caucasian, 1 Caucasian-Thai, 1 Vietnamese and 2 African-American) with some form of paroxysmal dyskinesia. One patient of mixed Caucasian-Thai background and one African-American family harbored the previously described hotspot mutation in *PRRT2* (c.649dupC, p.R217Pfs*8). Another African-American family was found to have a novel mutation (c.776dupG, p.E260*). Both of these variants are SNindels that arose in regions of mononucleotide repeats and are likely to cause loss-of-function via nonsense-mediated decay of mutant *PRRT2* transcripts. All affected individuals had classic paroxysmal kinesigenic dyskinesia phenotypes. In conclusion, heterozygous *PRRT2* gene mutations also cause paroxysmal kinesigenic dyskinesia in African-Americans. The c.649dupC hotspot mutation in *PRRT2* is common across racial groups.

2439W

Incomplete penetrance versus variable expressivity in del 16p11.2: analysis of cognitive performance and social functioning as quantitative rather than dichotomous traits. *A. Moreno-De-Luca[1,2], S.M. Myers[2], T.D. Challman[2], D.W. Evans[3], R.P. Goin-Kochel[4], E. Hanson[5], R. Bernier[6], L. Green Snyder[7], J.E. Spiro[8], W.K. Chung[9], J.N. Constantino[10], D.H. Ledbetter[1].* 1) Genomic Medicine Institute, Geisinger Health System, Danville, PA; 2) Department of Pediatrics, Geisinger Health System, Danville, PA; 3) Department of Psychology, Bucknell University, Lewisburg, PA; 4) Department of Pediatrics, Baylor College of Medicine, Houston, TX; 5) Department of Psychiatry, Harvard Medical School, Boston, MA; 6) Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA; 7) Children's Hospital Boston, Boston, MA; 8) Simons Foundation Autism Research Initiative, New York, NY; 9) Departments of Pediatrics and Medicine, Columbia University Medical Center, New York, NY; 10) Departments of Psychiatry and Pediatrics, Washington University School of Medicine, Saint Louis, MO.

The recurrent ~600 kb 16p11.2 deletion is one of the most common pathogenic copy number variants (CNVs) among individuals with neurodevelopmental disorders. It was initially identified in patients with autism and/or intellectual disability (ID) and subsequently associated with macrocephaly, obesity, and seizures. Deletion carriers show substantial clinical heterogeneity, including apparently normal individuals, the latter observation interpreted as evidence of incomplete penetrance. To evaluate the impact of this CNV on social functioning and autism features, we studied 20 individuals with *de novo* del 16p11.2 from the Simons Variation in Individuals Project and their non-carrier parents (n=40) and siblings (n=22). Within this group, only 30% (6/20) of probands met full criteria for autism using categorical diagnostic tools (ADI-R and ADOS). We also examined parent reports from the Social Responsiveness Scale (SRS), a quantitative scale that evaluates social awareness, reciprocal social communication, social information processing, and social anxiety, resulting in a T-score ranging from 30 (highly sociable) to 90 (severe social impairment). SRS scores are highly heritable, commonly unrelated to IQ, and continuously distributed in the general population. For the SRS, the mean T-score for probands was 74.5 compared to 47.1 in parents, and 44.5 in siblings. Therefore, this quantitative trait revealed a 2.7 SD shift of mean SRS scores of probands relative to unaffected intrafamilial controls ($p=6.49\times10^{-13}$). These results are very similar to those obtained using full scale intelligence quotient (FSIQ) to assess cognitive functioning in this subset of cases: only 19% met a categorical diagnostic criteria for ID (FSIQ ≤70); however, if viewed as a quantitative trait, FSIQ was 1.9 SD lower in probands compared to their parents ($p=6.72\times10^{-8}$). By using continuous, quantitative traits such as IQ and SRS scores to compare probands with their unaffected, non-carrier relatives, rather than using categorical variables such as DSM diagnoses or qualitative, dichotomous traits (i.e., normal vs. abnormal), we showed that parent-reported social behavior and cognitive function are significantly impacted in a deleterious fashion (2.7 SD on SRS and 1.9 SD on FSIQ) among children with del 16p11.2 when compared to non-carrier relatives. These data may be more consistent with variable expressivity related to genetic/family background rather than evidence of incomplete penetrance.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2440T

How to best phenotype the Broader Autism Phenotype (BAP) in Adults. *T. Nishiyama[1], H. Kishino[2], K. Okada[3], S. Sakai[4], Y. Kamio[5], M. Suzuki[6], K. Adachi[7], S. Sumi[8], M. Kojima[9], S. Suzuki[9].* 1) Clinical Trial Management Center, Nagoya City University Hospital, Nagoya, Japan; 2) Laboratory of Biometry and Bioinformatics, Graduate School of Agriculture and Life Sciences, University of Tokyo,Tokyo, Japan; 3) Department of Psychology, Senshu University, Tokyo, Japan; 4) Department of Child Development, Osaka University, United Graduate School of Child Development, Osaka, Japan; 5) Department of Child and Adolescent Mental Health, National Center of Neurology and Psychiatry, Tokyo, Japan; 6) Department of Psychiatry, Nagoya City University Graduate School of Medical Sciences, Nagoya, Japan; 7) Department of Psychiatric and Mental Health Nursing, Nagoya City University School of Nursing, Nagoya, Japan; 8) Department of Pediatrics, Nagoya Western Rehabilitation Center for Children with Disabilities, Nagoya, Japan; 9) Department of Public Health, Nagoya City University Graduate School of Medical Sciences, Nagoya, Japan.

Genome-wide association studies (GWAS) for SNPs and CNVs have thus far succeeded in explaining only a modest fraction of the genetic components of autism spectrum disorders (ASDs) [Mefford HC et al. 2012]. The previous GWAS for ASDs used categorical approach to diagnosis ignores the view that autistic traits are continuously distributed through the general population. If quantitative trait loci (QTL) analysis is to be a successful tool in the search for autism susceptibility genes, then an instrument that can quantify such traits, called the broader autism phenotype (BAP), are required.

So far, four self-administered questionnaires available for measuring BAP, the Autism Spectrum Quotient(AQ) [Baron-Cohen et al. 2001], the Social Responsiveness Scale—Adult Selfreport(SRS-AS) [pre-publication version of the SRS provided by Western Psychological Services], the Broad Autism Phenotype Questionnaire(BAPQ)[Hurley et al. 2007], and the Subthreshold Autism Trait Questionnaires(SATQ)[Kanne et al. 2007] have been developed. The former two contain 50 and 65 items, respectively, and thus may be too lengthy to be included in large genetic studies, whereas the latter two contain 36 and 24 items, respectively, and seem suitable to use in large studies.

Here, to provide evidence that will assist researchers in selecting an instrument for measuring the BAP in genetic studies, we firstly developed the Japanese version for questionnaires that remained undeveloped (the SRS-AS, BAPQ, and SATQ), and compared the all questionnaires for measuring BAP in terms of reliability and validity in a large sample (about 2000 healthy subjects and about 50 ASD subjects).

Based on 223 subjects currently available, all measures showed good reliability and validity. For example, Cronbach's alpha estimates and their 95% confidence intervals (CIs) are SATQ 0.84 (0.79—0.87), SRS-AS 0.95 (0.94—0.96), BAPQ 0.92 (0.89—0.93), and AQ 0.87 (0.84—0.89). As for criterion-related validity, areas under the ROC curve (AUCs) and their 95% CI are SATQ 0.98 (0.95—1.00), SRS-AS 0.98 (0.97—1.00), BAPQ 0.96 (0.91—1.000), and AQ 0.98 (0.96—1.00).This result suggests that all instruments for measuring BAP may be comparable and thus, shorter one can be a useful alternative to the lengthier one.

2441F

Executive dysfunction and bipolarity disorder in school-age children with academic impairment. *F. X. Coude[1,2], C. C. Mignot[2].* 1) Departement de Genetique, Hopital Necker, Paris, France; 2) Ctre de pediatrie, Aix-en-Provence, France.

OBJECTIVE:The study examined executive dysfunction (ED) and bipolarity disorder (BD) in school-age children (7 to 14 years) with academic impairment (AI). MATERIAL:A clinical sample of children diagnosed with academic impairment (n = 92) was compared to a population sample (n = 101) recruted in the same schools on two questionnaires, parental Bipolar Child Questionnaire and teacher Vanderbilt questionnaire. The prevalence of ED and BD in impaired and unimpaired children was examined. RESULTS:Children with AI were significantly more impaired on measures of ED than children without AI (8.6 vs 1.3, P<0.01) and BD index was also significantly elevated in these patients (45.3 vs 12.4, P<0.01). ED and AI Indexes were highly correlated and there was also a significant correlation between BD and ED in impaired patients. CONCLUSION:Children with AI displayed more problems on neuropsychological measures of ED and BP and ED measures were highly dependent of BD index. Therefore bipolarity seems frequent in children and may be an important cause of the learning impairment observed at school.

2442W

Identification of a dosage-sensitive brain development gene within the critical region of 1q deletion syndrome. *E.A. Erickson[1,2], W.A. Gahl[2], C. Boerkoel[1].* 1) NIH Undiagnosed Diseases Program Translational Laboratory, Bethesda, MD; 2) Section on Human Biochemical Genetics, Medical Genetics Branch, NHGRI, NIH, Bethesda, MD.

Submicroscopic terminal 1q deletion results in a syndrome consisting of severe psychomotor delays, aphasia, hypotonia, microcephaly, corpus callosum abnormalities, and facial dysmorphism. Though previous studies have narrowed the critical region for this phenotype to 1q43–1q44, the deleted genes responsible for the syndrome remain unknown. A potential contributor to this syndrome is the uncharacterized gene SCCPDH, which lies within the 1q44 critical region. We chose to investigate the effects of SCCPDH deletion on early brain development in zebrafish, since their development is easily observed and manipulated. RT-PCR analysis determined that the zebrafish SCCPDH homologue is normally expressed throughout development with its highest expression during the earliest stages. The SCCPDH homologue was knocked down in zebrafish by post-fertilization embryo injection of a synthetic antisense oligomer, or morpholino (MO), designed to bind and block a splice site within SCCPDH pre-mRNA. At 24 and 48 hours post fertilization, the embryos had a dose-dependent defect in midbrain-hindbrain boundary formation and hydrocephalus, respectively. This same phenotype was observed with another MO that inhibits translation of SCCPDH transcripts. Demonstrating that the effects of the MO are specific to SCCPDH knockdown, co-injection of SCCPDH mRNA and the SCCPDH splice-site targeting MO rescued the phenotype; over 70% of the co-injected embryos were morphologically similar to the uninjected control embryos. Interestingly, the overexpression of SCCPDH mRNA alone caused cyclopia and poor brain development. These results indicate that the level of SCCPDH expression is highly regulated and that alterations beyond a threshold can result in abnormal brain development. This work suggests that SCCPDH plays an important role in early brain development and merits further investigation.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2443T

Results of genome-wide analyses on neurodevelopmental phenotypes at four-year follow-up following cardiac surgery in infancy. *D.S.K. Kim[1,2], I.B.S. Stanaway[2], R.R. Rajagopalan[1], G.W. Wernovsky[3,4], J.B. Bernbaum[5], C.B.S. Solot[6], N.B. Burnham[7], E.Z. Zackai[8], R.R.C. Clancy[9], S.C.N. Nicolson[10], M.G. Gerdes[11], D.A.N. Nickerson[2], H.H. Hakonarson[12,13], J.W.G. Gaynor[7], G.P.J. Jarvik[1,2].* 1) Department of Medicine, Division of Medical Genetics, University of Washington School of Medicine, Seattle, WA, USA; 2) Department of Genome Sciences, University of Washington School of Medicine, Seattle, WA, USA; 3) Division of Pediatric Cardiology, Children's Hospital of Philadelphia, Philadelphia, PA, USA; 4) Division of Critical Medicine, Children's Hospital of Philadelphia, Philadelphia, PA, USA; 5) Division of General Pediatrics, Children's Hospital of Philadelphia, Philadelphia, PA, USA; 6) Center for Childhood Communication, Children's Hospital of Philadelphia, Philadelphia, PA, USA; 7) Division of Cardiothoracic Surgery, Children's Hospital of Philadelphia, Philadelphia, PA, USA; 8) Division of Genetics, Children's Hospital of Philadelphia, Philadelphia, PA, USA; 9) Division of Neurology, Children's Hospital of Philadelphia, Philadelphia, PA, USA; 10) Division of Cardiothoracic Anesthesiology, Children's Hospital of Philadelphia, Philadelphia, PA, USA; 11) Division of Psychology, Children's Hospital of Philadelphia, Philadelphia, PA, USA; 12) Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, PA, USA; 13) Department of Pediatrics, University of Pennsylvania School of Medicine, Philadelphia, PA, USA.

Background: Adverse neurodevelopmental sequelae are reported among children who undergo early cardiac surgery to repair congenital heart defects (CHD). *APOE* genotype has previously been determined to contribute to the prediction of these outcomes. Understanding further genetic causes for the development of poor neurobehavioral outcomes should enhance patient risk stratification and improve both prevention and treatment strategies.

Methods: A prospective observational study with neurodevelopmental evaluation between the fourth and fifth birthdays was performed. Attention and behavioral skills were assessed through parental report. Linear regression was performed to determine the effect of genome-wide genetic variation on neurodevelopmental measures in the fifth year of life in 316 subjects who underwent cardiac surgery before six months of age.

Results: This genome-wide association study identified single nucleotide polymorphisms (SNPs) associated with three neurobehavioral phenotypes in the postoperative children: Attention Deficit-Hyperactivity Disorder-IV (ADHD) Scale for Impulsivity/Hyperactivity, and Child Behavior Checklist (CBCL) Pervasive Developmental Problems (PDPs), and Total Problems. The most predictive SNPs for each phenotype were: a *POC5* intronic SNP, rs2303647, associated with ADHD-IV Impulsivity ($P = 7.93 \times 10^{-7}$), a *SEMA6D* intronic SNP, rs16959924, associated with CBCL/1.5–5 PDPs ($P = 4.08 \times 10^{-7}$), and a *CHL1* intronic SNP, rs2293675, associated with CBCL/1.5–5 Total Problems ($P = 8.85 \times 10^{-7}$). 15 SNPs (6 for ADHD-IV Impulsivity, 5 for CBCL/1.5–5 Total Problems, and 4 for CBCL/1.5–5 PDPs) had $p < 10^{-5}$.

Conclusions: No SNPs met genome-wide significance for our three neurobehavioral phenotypes, though 15 SNPs reached a threshold for suggestive significance ($p < 10^{-5}$). Given the unique nature of this cohort, larger studies and/or replication are not possible. Studies to further investigate the mechanisms through which these newly identified genes may influence neurodevelopment dysfunction are warranted.

2444F

BrainSpan Atlas of the Developing Human Brain. *S. Sunkin[1], A. Bernard[1], T. Chen[2], C. Dang[1], S. Ding[1], A. Ebbert[1], O. Evgrafov[2, 3], B. Fischl[4], M. Gerstein[5], D. Geschwind[6], A. Harmanci[5], J. Hohmann[1], S. Horvath[6], T. Hyde[7], A. Jones[1], Y. Kawasawa[5], R. Kitchen[5], J. Kleinman[7], P. Levitt[2, 3], M. Li[5], S. Mane[5], N. Parikshak[6], A. Sboner[8], E. Shen[1], K. Smith[1], D. Weinberger[7], M. Hawrylycz[1], J. Knowles[2, 3], N. Sestan[5], E. Lein[1].* 1) Allen Institute for Brain Science, Seattle, WA; 2) USC, Los Angeles, CA; 3) Zilkha Neurogenetic Institute, Los Angeles, CA; 4) Massachusetts General Hospital; Harvard Medical School, Boston, MA; 5) Yale University, New Haven, CT; 6) UCLA, Los Angeles, CA; 7) NIMH, Bethesda, MD; 8) Institute for Computational Biomedicine, Weill Cornell Medical College, New York, NY.

Systematic efforts to map detailed gene expression patterns in the human brain have been lacking, particularly relating to development. To fill this void, we describe here a consortium project aimed at creating a unique multimodal transcriptional atlas of the pre- and postnatal developing human brain as a publicly accessible online resource for neuroscience, genome and medical research communities. This resource integrates transcriptomic, cellular resolution histology and imaging data in the context of human brain development through a portal for viewing, searching and mining of spatiotemporal gene expression patterns. Illumina RNA-sequencing has been used to generate quantitative transcriptome profiles of hippocampus, amygdala, thalamus, striatum, cerebellum and 11 neocortical regions, spanning early prenatal development through adulthood. This broad spatiotemporal data set is supplemented with a high anatomical resolution microarray analysis of >300 laser microdissected structures from mid-gestational prenatal human brains. Complementing the transcriptome data, a large scale in situ hybridization (ISH) data set was generated to analyze cellular distributions. Finally, de novo prenatal histological reference atlases and MRI and diffusion weighted MRI (DWI) data sets spanning human brain development are being created to provide a neuroanatomical and neurodevelopmental context for understanding spatiotemporally regulated transcriptional programs. As an initial representation of the project, here we present integrated data on the transcriptome analyses by RNA-sequencing and cellular resolution in situ hybridization of the adult human brain. The majority of protein-coding genes are well expressed in the human brain. Comprehensive profiling of microRNA and the discovery of novel transcripts are also demonstrated. Many analyzed transcripts exhibit distinct regional and cellular patterns of expression, providing a neuroanatomical context for understanding transcriptional complexity. We also identified transcripts with allele-specific expression. These public data provide valuable opportunities for relating specific transcriptional programs to processes of human brain development and a normative data set for understanding the genetic basis of neuropsychiatric disease. These data are openly available as the BrainSpan Atlas of the Developing Human Brain via the Allen Brain Atlas data portal (www.brain-map.org) or directly at www.brainspan.org.

2445W

Knockdown of putative human autism genes, SHANK3 and SYNGAP1, produces common behavioral phenotypes in zebrafish. *J.E. Dallman[1], R. Kozol[1], Q. Yan[1], H.N. Cukier[2], V. Mayo[2], A.J. Griswald[2], P.L. Whitehead[2], J.L. Haines[2], J.R. Gilbert[2], M.L. Cuccaro[2], E.R. Martin[2], M.A. Pericak-Vance[2].* 1) Biology, University of Miami, Coral Gables, FL; 2) Hussman Institute for Human Genomics, University of Miami Miller School of Medicine, Miami FL.

Autism Spectrum Disorders (ASDs) affect about 1 in 110 children according to the Centers for Disease Control, yet the functional links between genetic mutation and behavioral changes characteristic of ASD remain unknown. ASDs are a heterogeneous collection of neuro-developmental disorders that can be caused by variations in a diverse array of genes, including transcription, translation, and synaptic proteins. Patients with ASD share three clinical features: deficits in verbal and non-verbal communication, reduced social interactions, and repetitive motor behaviors. Here, we present functional data from four zebrafish models of ASD. To generate these models, we use splice-site targeted morpholinos to disrupt zebrafish orthologs of SHANK3, a synaptic scaffolding protein, and SYNGAP1, a synaptic GTPase activating protein. SHANK3 and SYNGAP1 were chosen because of mounting evidence that mutations in these genes are causal to ASD in humans. Both SHANK3 and SYNGAP1 are duplicated in the zebrafish genome and all four of these transcripts (shank3, shank3–like, SynGAP, and SynGAP-like) are enriched in zebrafish brain over spinal cord as demonstrated by in situ hybridization. Our results identify a common developmental time period, 48–72 hours post-fertilization (corresponding to a dramatic burst of synaptogenesis throughout the nervous system), when knockdown of either duplicate of SHANK3 or SYNGAP1 causes pronounced disruptions in escape behavior, including spontaneous seizures and unproductive swim attempts. Based on gene expression patterns and morphant behaviors, we are currently testing the hypothesis that altered neuronal activity in descending brain neurons accounts for disrupted motor behaviors in these zebrafish models of ASDs.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2446T

Short tandem repeat: A versatile marker in diagnosis of genetic disorders. *S. Muthuswamy, A. Kumar, S. Agarwal.* Genetics, Sanjay Gandhi Post Graduate Inst of Medical Sciences, Lucknow, India.

Short tandem repeats are Simple sequence repeats (SSR) of 1–10bp. They are abundant and evenly distributed throughout the genome; STR is highly stable and polymorphic, they vary in length between the subjects. As well as the fact that several genetic diseases are associated with tri-nucleotide repeats such as Friedreich's ataxia, myotonic dystrophy etc. Therefore these markers are now a day's put-upon in molecular diagnosis to have rapidness in reporting and for the purpose of accuracy. The present study was aimed to provide rapid diagnosis for Down syndrome, Friedreich's ataxia and Myotonic dystrophy with STR markers associated with respective chromosome/gene. QF-PCR with 21st chromosome associated STR markers (D21S11, D21S1411 and D21S1435) are used in diagnosis of trisomy 21 and TP-PCR with primers flanking the triplet repeat of the respective gene for Friedreich's ataxia and Myotonic dystrophy. The heterogeneity of the STR markers was assessed and found suitable for the establishment of QF-PCR. We have started to provide prenatal diagnosis for Down syndrome, in our last one year experience; we tested 95 samples (40 CVS and 55 amniotic fluids) and found none of the fetus positive and results were concordant with karyotyping results. Triplet repeat primed PCR was also accurate in gene level diagnosis of Friedreich's ataxia and myotonic dystrophy, which is confirmed with short and long PCR simultaneously.

2447F

Genomic Rearrangements in Idiopathic Cerebral Palsy. *R. Segel[1,3], H. Ben-Pazi[2,3], S. Zeligson[1], A. Fattal-Valevski[4,7], A. Aran[2,3], V. Gross[2,3], N. Shneebaum[4,7], D. Shmueli[5], D. Lev[6,7], L. Blumkin[6,7], S. Perlberg[1], E. Levy-Lahad[1,3].* 1) Medical Genetics Institute, Shaare Zedek Medical center, Jerusalem, Israel; 2) Pediatric Neurology Unit, Shaare Zedek Medical Center, Jerusalem,Israel; 3) Hadassah Medical School, Hebrew University, Jerusalem, Israel; 4) Institute for Child Development and Pediatric Neurology Unit, Tel Aviv Sourasky Medical Center, Tel Aviv, Israel; 5) Jerusalem Child Development Center, Clalit Health Services, Jerusalem,Israel; 6) Metabolic-Neurogenetic Clinic, Wolfson Medical Center, Holon; Israel; 7) Sackler Faculty of Medicine, Tel Aviv University, Ramat Aviv, Israel.

Background: Cerebral Palsy (CP) is a non-progressive motor impairment that is the most common cause of disability in childhood. The worldwide incidence of CP is 2–2.5 / 1,000 live births. While some risk factors for CP are known, the etiology remains largely obscure. We hypothesized that some idiopathic CP cases have a genetic etiology. Recent studies have shown that genomic rearrangements (GRs) have a role in a number of complex pediatric disorders, e.g. mental retardation, autism and epilepsy. The aim of this preliminary study was to assess the role of GRs in individuals with idiopathic CP, using microarray analysis. Methods: Participants were recruited from CP clinics in Israel. We enrolled individuals who were most likely to have a genetic etiology. Inclusion criteria were: 1) non progressive pyramidal or extra pyramidal symptoms beginning before the age of 3 years; 2) brain imaging not characteristic of ischemic damage: normal or with a developmental brain abnormality. Exclusion criteria were targeted at patients with acquired CP: 1) Periventricular leukomalacia (PVL) in patients born prematurely (24–34 weeks gestational age); 2) History of hypoxic ischemic encephalopathy; 3) Hemispheric lesion corresponding to an acquired lesion on imaging studies (i.e. stroke); 4) History of encephalitis or head trauma. Genomic rearrangements were tested using the Affymetrix CytoScan® HD Array and analyzed using ChAS (Affymetrix® Chromosome Analysis Suite). Rearrangements were considered potentially pathogenic if they were de novo, and not reported in databases of genomic variants (DGV). Results: We report analysis of the first 49 cases enrolled. 16/49 had genomic rearrangements: 7 de novo, 3 inherited and in 6 parental testing is ongoing. 5 samples failed. Thus, among completely analyzed cases 7/44 had de novo rearrangements. Analysis of the loci and the genes involved is ongoing, with the goal of identifying common pathways. Discussion: In this preliminary study, we found a high frequency (16%) of de novo GRs in patients with non-acquired CP. Clinically, identifying CNVs with a specific phenotype can end a long quest for diagnosis, and is important for prognosis and genetic counseling. Scientifically, elucidating the role of GRs in CP may aid in understanding CP pathogenesis and may identify specific genes and pathways important for brain development.

2448W

Gain of function Na_v 1.7 mutations in patients with small fiber neuropathy. *M.M. Gerrits[1], C.G. Faber[2], J.G.J. Hoeijmakers[2], D.M.L. Merckx[1], H.J.M. Smeets[1], S.D. Dib-Haij[4,5], S.G. Waxman[4,5], I.S.J. Merkies[2,3].* 1) Department of Clinical Genetics, Maastricht University Medical Center, Maastricht, Netherlands; 2) Department of Neurology, Maastricht University Medical Center, Maastricht, Netherlands; 3) Department of Neurology, Spaarne Hospital, Hoofddorp, The Netherlands; 4) Yale University School of Medicine, New Haven, USA; 5) Center for Neuroscience and Regeneration Research, Veterans Affairs Medical Centre, Westen Haven, USA.

Background: Small fiber neuropathy (SFN) often occurs without apparent cause. We have recently described the presence of gain-of-function variants in SCN9A, the gene encoding for the $Na_v1.7$ sodium channel, in 8 patients with idiopathic small fiber neuropathy. In this study a cohort of 165 patients with possible SFN was tested for the presence of SCN9A variants.**Methods:** Patients referred with possible SFN, who met the criteria of ≥ 2 SFN-related symptoms, a normal electromyogram, and an abnormal quantitative sensory testing (QST) and/or reduced intraepidermal nerve fiber density (IENFD) were screened for SCN9A gene variants. Electrophysiology was used to test functional effects of variant channels. **Results:** Twelve patients of the 165 patients who underwent SCN9A gene analyses were found to carry a variant in the SCN9A gene. The frequency of these variants in control populations ranged from zero to 0.9%. In each case, the variant was missense, and the patient was heterozygous for the variant. Five variants were found in more than one patient. Voltage-clamp analysis was performed for four of the variants and they all showed gain-of-function (impaired slow-inactivation, depolarized slow- and fast-inactivation, or enhanced resurgent currents). Current-clamp analysis demonstrated that these variants rendered dorsal root ganglion (DRG) neurons hyperexcitable. **Conclusion:** Missense substitutions in the SCN9A gene, encoding the $Na_v1.7$ sodium channel, are present in a substantial proportion (~7%; 12 of 165) of this cohort. Electrophysiological analysis of the variants revealed gain-of-function in the mutant channels and showed that the variants render DRG neurons hyperexcitable. This implies that SCN9A functional variants may predispose to the development of channelopathy-associated SFN. Analysis of the SCN9A gene should be considered for patients with SFN-related symptoms plus a normal electromyogram, and abnormal QST and/or reduced IENFD.

2449T

A New Screening Test for Fragile X Syndrome: Quantitative Measurement of FMRP in Dried Blood Spots. *W.T. Brown, T. Adayev, R. Kascsak, R. Kascsak, C. Dobkin, S. Nolin, G. LaFauci.* Dept Human Gen, NYS Inst Basic Res, Staten Island, NY.

Fragile X syndrome (FXS) is caused by the lack of the Fragile X Mental Retardation 1 protein (FMRP) encoded by the FMR1 gene. In most FX full mutation (FM) subjects, FMRP is not made because of transcriptional inactivation of the gene resulting from an expanded triplet CGG repeat. Early diagnosis of the syndrome is important for early intervention. The laboratory diagnosis of FXS is currently performed using DNA to determine both the length and the methylation state of the CGG repeats in FMR1. This information is then used to infer whether abnormal levels of FMRP are present. The tests often require several days, are expensive, labor intensive, and can be performed only in specialized laboratories. We developed a Luminex-based immunoassay (ELISA) for the direct detection and quantification of FMRP in human and mouse cell extracts (lymphoblastoid cells, fresh human lymphocytes, and mouse brain). FMRP is captured by reacting the samples with X-map microspheres coated with a high affinity anti-FMRP monoclonal antibody (mAb). The microsphere-mAb-FMRP complex is successively reacted with an anti-FMRP rabbit polyclonal antibody (Ab) and with an anti-rabbit-IgG Ab conjugated to phycoerythrin. The fluorescence emitted is a function of the amount of FMRP in the specimen and is detected using a Luminex-200 System. We have adapted our assay for the detection and quantification of human FMRP in Dried Blood Spots (DBS). Drops of blood are spotted onto a collection card, air-dried for few hours. Card preparation requires small blood samples that can be drawn by lancet from the finger of a patient or from the heel or toe of neonate. Furthermore, cards can be stored at room temperature in low gas-permeability plastic bags and send by regular mail to a laboratory for testing. Assays are done in 96–well-plates which allow the simultaneous processing of 40 specimens (in duplicate), and can be completed in 24 hours. Using the minute amount of proteins extracted from DBS, we can accurately quantify FMRP using as the standard a recombinant fusion protein carrying the capture and detection epitope domains of FMRP. The assay can identify male FM FX, and male mosaic FM FX from normal and from premutation (PM) individuals with sensitivity and specificity approaching 100%. Since newborn DBS screening for metabolic diseases is performed routinely, our assay offers for the first time an accurate, simple, economical and fast method to perform neonatal screening for FXS.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2450F

Disease specific pathway analysis in Multiple Sclerosis. *T. Avsar[1], M. Durasi[2], S. Saip[3], O.U. Sezerman[2], A. Siva[3], E. Tahir Turanli[1,4].* 1) Molecular Biology and Genetics, Istanbul Technical University, Istanbul, Turkey; 2) Faculty of Engineering & Natural Sciences, Biological Sciences & Bioengineering, Sabanci University; 3) Department of Neurology, Cerrahpasa School of Medicine, Istanbul University, Istanbul, Turkey; 4) Molecular Biology and Genetics Department, Science and Letter Faculty, Istanbul Technical University, Istanbul, Turkey.

Multiple sclerosis (MS) is a chronic inflammatory demyelinating disorder of the central nervous system (CNS) with various degrees of axonal damage. The aim of this study is to identify clinical subtype specific novel protein biomarkers in the CSF and serum samples of MS patients to reveal disease pathways. We compared CSF and serum samples of 35 CIS, 45 RRMS, 20 PPMS and 18 controls including healthy subjects and other neurological disease samples for novel protein biomarkers. For all samples, 2D-PAGE analysis by using different pH and molecular weight ranges were performed. At least 4 times differently expressed protein spots were identified by PDQuest® software and removed from the gel for MALDI-TOF-MS protein identification step. The results of proteomic studies were confirmed by both ELISA and western blot methods for the selected candidate proteins. Protein levels of clinically different MS subtypes were compared with control groups and with each other by using ANOVA test. Further, KEGG pathways analysis was done with genes corresponding the proteins from proteomic study in order to identify relevant disease specific molecular pathways. Proteomic studies revealed in total 73 different proteins comparing to control groups and disease subtypes. ELISA and western blot results confirmed the results for selected proteins (Transthyretin ($p<0.01$), Chain A-Alpha 1 Anti-trpsin ($p=0.039$), Apolipoprotein A4 ($p<0.01$), Apolipoprotein E ($p<0.01$) alpha 2 macroglobulin ($p=0.023$), human alpha 1 acid glycoprotein ($p=0.041$) and prostaglandin D2 synthase ($p<0.01$). Common differentially expressed proteins within each clinical subtype and in all MS cases are listed (Table-1). Finally, KEGG pathway analysis revealed statistically significant over-representation of the processes of Tryptophan metabolism, B cell receptor signaling pathway, Epithelial cell signaling in Helicobacter pylori infection. Proteomic investigation in CSF and serum samples of clinically different MS patients and controls revealed a number of proteins from different protein families that are differentially expressed in clinically different MS subtypes when compared to controls. Some are common for all disease subtypes but some are specifically expressed in only certain disease subtypes. Furthermore KEGG analysis revealed the disease related pathways indicating B cell contribution to the disease.

2451W

Diagnosing complex neurological disease by next-generation sequencing (NGS). *K. Haugarvoll[1,2], S. Johansson[3], C. Tzoulis[1,2], B.I. Haukanes[3], C. Bredrup[4], H. Boman[3], P.M. Knappskog[3], L.A. Bindoff[1,2].* 1) Dept Neurology, Haukeland Univ Hosp, Bergen, Norway; 2) Department of Clinical Medicine, University of Bergen, Norway; 3) Center for Medical Genetics and Molecular Medicine, Haukeland Univ Hosp, Bergen, Norway; 4) Department of Ophthalmology, Haukeland University Hospital, Bergen, Norway.

A correct diagnosis is pivotal to understand and treat neurological disease. Rare diseases may pose a formidable diagnostic challenge, especially when they present with a complex phenotype including clinical features that may overlap between several disorders. Herein we employ next-generation sequencing (NGS) to identify the cause of disease in two patients. The patients are siblings with a complex adult onset syndrome comprising peripheral neuropathy, epilepsy, relapsing encephalopathy, bilateral thalamic lesions on MRI, type 2 diabetes mellitus, retinal pathology and tremor. We applied clinical and genealogical investigation, homozygosity mapping and exome sequencing in order to achieve a diagnosis in these two patients. We suspected a recessive genetic defect as the cause disease in these two siblings of healthy, albeit genealogically related parents. Homozygosity mapping revealed three shared homozygous regions. Subsequently, we performed exome sequencing. Both patients were homozygous for a novel c.367G>A mutation (p.D123N) in the AMACR-gene. The genetic diagnosis of α-methylacyl-coA racemase deficiency was confirmed by demonstrating markedly increased pristanic acid levels in blood (169 µmol/L, normal <1.5 µmol/L). Revision of sequential MRI images from our patients reveals specific chronic degeneration of cerebellar afferents and efferents. A correct diagnosis of α-methylacyl-coA racemase deficiency, an autosomal recessive peroxisomal disorder with accumulation of pristanic acid, was reached utilizing NGS. Six adult patients with α-methylacyl-coA racemase deficiency have been reported to date. Importantly, patients with this disorder may benefit from restricted dietary phytanic and pristanic acid intake. This study highlights the potential role of NGS in improved diagnosis of neurological disorders.

2452T

Elevations of pro-inflammatory cytokines/chemokines in patients with Aicardi-Goutières Syndrome. *J.L. Schmidt[1, 2], A. Takanohashi[1], M. Prust[2], K.F. Gold[3], J. Wang[3], H. Gordish-Dressman[1], G. Rice[4], H. Gornall[4], M. Bloom[5], Y.J. Crow[4], P. Lebon[6], K. Nagaraju[1], A. Vanderver[1, 2].* 1) Center for Genetic Medicine Research, Children's National Medical Center, Washington, DC, USA; 2) Department of Neurology, Children's National Medical Center, Washington, DC, USA; 3) Department of Biostatistics, Children's National Medical Center, Washington, DC, USA; 4) Genetic Medicine, University of Manchester, Manchester Academic Heath Science Centre, Central Manchester University Hospitals, Manchester, UK; 5) Hospitalist Medicine Division, Children's National Medical Center, Washington, DC, USA; 6) Department of Virology, Hospital St Vincent de Paul, Paris, France.

Background: Aicardi-Goutières Syndrome (AGS) is a heritable neurologic disorder with genetic heterogeneity (TREX1, RNAseH2A, RNAseH2B, RNaseH2C and SAMHD1). Mutations in the involved genes are hypothesized to result in accumulation of endogenous nucleic acids, triggering an IFNα mediated innate immune response in the brain. Cerebrospinal fluid (CSF) elevations of interferon alpha (IFNα) and other isolated cytokines have previously been associated with AGS. However, these biomarkers are established only in CSF and are known to decrease with age. This study explores a panel of cytokines in plasma and CSF in patients at different ages, to establish signatures of cytokines and chemokines most predictive of AGS disease. Methods: Plasma from 22 subjects with mutations in TREX1, RNAseH2 or SAMHD1 were assayed for cytokines and chemokines using the MILLIPLEX™ MAP Immunobead system, and compared to results from 8 age-matched normal controls. Cerebrospinal fluid of 11 subjects with mutation proven AGS was tested in an identical manner and compared to results from, 8 age-matched controls. Results: Significant elevations were seen in a panel of 23 cytokines in AGS CSF, and in Flt-3L, IP-10, IL-12p40, IL-15, TNF and sIL-2ra in both AGS patient plasma and CSF relative to controls. Additionally, this cytokine signature was able to correctly cluster a majority of AGS cases based on CSF values. While the majority of cytokines decreased exponentially with age, a smaller subgroup, including IP-10, had persistent elevations beyond early childhood. Conclusion: AGS patients have plasma and CSF elevations of proinflammatory cytokines/chemokines such as in Flt-3L, IP-10, IL-12p40, IL-15, TNF and sIL-2ra. These proinflammatory cytokines and chemokines may be considered for use as diagnostic and therapeutic markers of disease.

2453F

Strategy for genotyping triplet repeats unstable sequences. *S. Lindo_Samanamud[1,2], O. Ortega[1], M. Cornejo-Olivas[1,3], V. Marca[1], P. Mazzetti[1,4].* 1) Neurogenetics Research Center, Instituto Nacional de Ciencias Neurologicas, Lima, Peru; 2) School of Genetics and Biotechnology, Universidad Nacional Mayor de San Marcos, Lima, Peru; 3) UW International Fogarty Scholars and Fellow's Program in Peru. Lima, Peru; 4) Shool de Medicine. Universidad Nacional Mayor de San Marcos, Lima, Peru.

Introduction In triplet repeat diseases with unstable microsatellites and a high content of cytosine, PCR is not useful because of the generation of alternative structures. As a consequence of this abnormal expansion, the region containing the triplet is methylated, as are other regions nearby, promoter regions for example. **Objectives** To modify cytosine-rich repeat sequences to allow PCR amplification; and further designing specific primers, to differentiate the methylation status of modified regions, the one containing the triplet and the nearby regions.**Methods** After appropriate informed consent, DNA of the FMR1 gene from 8 healthy subjects (four female and four male) was modified with sodium bisulfite. Subsequently, using bioinformatics tools, we performed: 1) A simulation of alternative structures of the unstable microsatellite in the 5′-UTR region, 2) After localization of the CpG islands, we generate specific primers which hybridize with the new (modified) expanded microsatellite sequence of the gene (Primers NM) and specific primers that hybridize to the new sequence containing CpG islands (Primers M) in the promoter. Finally, both of these sequences were amplified by PCR. **Results** Modified DNA was obtained after chemical treatment with sodium bisulfite. Alternative structures of the sequence of the microsatellite were characterized as described. The modified unmethylated diclucleotides CG (now UG) on the FMR1 gene were identified by the NM primers by PCR. Those on the FMR1 promoter were identified by the M primers. Male samples were primer M (promoter) negative and primer NM (triplet repeat) positive and normal female ones were primer M positive and primer NM negative.Primers NM flank the region with the abnormal expansion and allow counting triplet size count **Conclusions** We designed a potential alternative method for amplification of cytosine rich microsatellite sequences. Combining bisulfite modification(CG to UG) and primers binding to the unmethylated gene and the methylated promoter of the FMR1 gen, we were able to recognize the status of the gene and its promoter in normal men and women. This method will potentially allow differentiating between normal, premutation and full mutation status. Further studies are needed in DNA samples from individual with expanded microsatellites.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2454W

Cleavage and polyadenylation specificity factor 1 (CPSF1) regulates alternative splicing of interleukin 7 receptor (IL7R) exon 6. *S.G. Gregory[1,2,3], I. Evsyukova[2,4], S. Bradrick[2,5], M.A. Garcia-Blanco[2,3,5].* 1) Center for Human Genetics; 2) Department of Molecular Genetics and Microbiology; 3) Department of Medicine; 4) Department of Biochemistry; 5) Center for RNA Biology, Duke University Medical Center.

Interleukin 7 receptor, IL7R, is expressed exclusively on cells of the lymphoid lineage and its expression is crucial for the development and maintenance of T cells. While transcriptional regulation of IL7R expression has been widely studied, its posttranscriptional regulation has only recently been uncovered. Alternative splicing of IL7R exon 6, the only exon that encodes the transmembrane domain of the receptor, results in membrane-bound (exon 6 included) and soluble (exon 6 skipped) IL7R isoforms, respectively. Interestingly, the inclusion of exon 6 is affected by a single-nucleotide polymorphism associated with the risk of developing multiple sclerosis, a prototypic demyelinating disease of the central nervous system. Given the potential association of exon 6 inclusion with multiple sclerosis, we investigated the cis-acting elements and trans-acting factors that regulate exon 6 splicing. We identified multiple exonic and intronic cis-acting elements that impact inclusion of exon 6. Moreover, we utilized RNA affinity chromatography followed by mass spectrometry to identify trans-acting protein factors that bind exon 6 and regulate its splicing. These experiments identified cleavage and polyadenylation specificity factor 1 (CPSF1) among protein binding candidates. A consensus polyadenylation signal AAUAAA is present in intron 6 of IL7R directly downstream from the 5′ splice site. Mutations to this site and CPSF1 knockdown both resulted in an increase in exon 6 inclusion. Correspondingly, CPSF1 depletion had no effect on a minigene with a mutation in the intronic polyadenylation site. 3′RACE and RT-PCR experiments on RNA from Jurkat cells suggested that the intronic AAUAAA site is utilized at low frequency by the polyadenylation machinery to produce a novel polyadenylated mRNA isoform. Our results suggest that competing mRNA splicing and polyadenylation may regulate exon 6 inclusion and resultant levels of functional IL7R produced. Since the intronic polyadenylated isoform of IL7R is predicted to be translated into a membrane-bound protein product with a shortened, signal transduction-incompetent cytoplasmic tail, this may be relevant for both T cell biology and development of multiple sclerosis.

2455T

Functional analysis of *ASNS* variant identified in progressive microcephaly and brain atrophy patients. *Y. Hitomi[1], E.K. Ruzzo[1], D.L. Silver[2], Y. Anikster[3,4], D.B. Goldstein[1,2].* 1) Center for Human Genome Variation, Duke University School of Medicine, Durham, NC; 2) Department of Molecular Genetics and Microbiology, Duke University School of Medicine, Durham, NC; 3) Edmond and Lily Safra Children's Hospital, Sheba Medical Center, Ramat Gan, Israel; 4) The Sackler School of Medicine, Tel Aviv University, Tel Aviv, Israel.

We previously analyzed a specific genetic variant in two Israeli families that presented with a seemingly similar and potentially newly described congenital microcephaly condition. Whole-exome sequencing was applied to both families, and the sequence data identified that all three patients shared a homozygous nonsynonymous variant located in the *asparagine synthetase*(*ASNS*) gene that was absent in over 300 sequenced controls. The protein encoded by this gene is involved in the synthesis of asparagine from aspartate. In this study, we sought to characterize the functional consequences of the *ASNS* variant detected in the patients. Variant *ASNS* cDNA clones were generated by mutagenesis and transfected into monkey kidney cell line COS-7, human kidney cell line HEK-293, and human neuroblastoma SH-SY5Y. Although quantities of exogenous *ASNS* mRNA were similar between variant and wildtype alleles, the variant form of the ASNS protein product was drastically diminished in these transfectants in a lysosomal degradation system-dependent manner. Patient's fibroblast from the skin also showed no ASNS protein product. These results strongly suggested that the nonsynonymous variant in *ASNS* accounts for the pathogenesis. In this presentation, we will also discuss the morphological abnormalities of the brain in the *Asns* deficient mice. Additionally, the efficiency of glutamate metabolism and glutamate-related cytotoxicity in induced pluripotent stem cells (iPSc)-derived neuron which has been obtained from patient's skin fibroblast will be also discussed. These analyses illustrate that straightforward functional work can significantly facilitate the interpretation of whole-exome sequence data, and facilitates the discovery of novel therapies for such central nervous system (CNS) disorders by identification of the molecular mechanism of pathogenesis.

2456F

Impaired PPARδ function underlies mitochondrial demise and neurodegeneration in Huntington's disease: mechanistic basis and therapeutic implications. *A.S. Dickey[1], V.V. Pineda[2], A. Buttgereit[1], M.M. Torres[2], T. Tsunemi[1], A.R. La Spada[1,3,4].* 1) Pediatrics, University of California, San Diego, La Jolla, CA; 2) Laboratory Medicine, University of Washington, Seattle WA; 3) Cellular & Molecular Medicine, University of California, San Diego, La Jolla, CA; 4) Neurosciences, University of California, San Diego, La Jolla, CA.

Huntington's disease (HD) is a progressive autosomal dominant neurodegenerative disorder characterized by involuntary movements, cognitive decline, and psychiatric illness. HD arises from a CAG - polyglutamine repeat expansion in the coding region of the huntingtin (htt) gene. Recent studies have shown that the mitochondrial dysfunction in HD results from transcriptional dysregulation of peroxisome proliferator-activated receptor gamma coactivator-1–alpha (PGC-1α). As a transcriptional regulator, PGC-1α positively modulates nuclear receptor transcription factors including the three PPARs, α, γ, and δ. Of these, PPARδ is the most abundantly expressed in the CNS, although its functional relevance in this tissue is not defined. With the discovery that retinoic acid binds to PPARδ to mediate pro-survival effects, PPARδ may be responsible for promoting a variety of survival / repair processes in neurons. When an unbiased screen for htt interacting proteins yielded PPARs as candidate interactors, we evaluated the different PPARs, and demonstrated an interaction between htt and PPARδ in co-transfection studies. We found that htt and PPARδ interact in HD striatal neurons, and in the brains of BAC-HD transgenic mice, and we mapped the interaction domain on PPARδ to the “hinge” domain. PolyQ-htt repressed PPARδ-mediated transactivation in ST-Hdh cells and primary neurons from BAC-HD mice, but could be rescued by over-expression of PGC-1α or PPARδ. These studies revealed that both a PPARδ-selective ligand and retinoic acid elicited PPARδ-mediated transactivation in neurons. Furthermore, activation of retinoic acid receptor (RAR) transactivation was favored in HD striatal-like neurons and BAC-HD primary cortical neurons. Indeed, studies in BAC-HD primary neurons confirmed that modulation of retinoid binding protein expression levels promoted cell death and mitochondrial dysfunction when RAR signaling was agonized, while augmentation of PPARδ-mediated transactivation boosted cell survival and mitochondrial function. Microarray analysis of PPARδ-regulated genes in the striatum of HD patients and controls confirmed that PPARδ targets are altered in early symptomatic HD patients. Our results indicate that impaired PPARδ function in the CNS contributes to HD neurodegeneration, and suggest that modulation of the PGC-1α - PPARδ pathway is an attractive therapeutic strategy for HD and related neurodegenerative disorders characterized by mitochondrial dysfunction.

2457W

NHEJ1 deficiency causes abnormal development of the cerebral cortex. *B. EL WALY[1,2], E. BUHLER[2,3,4,5], M.R. HADDAD[1,2], F. WATRIN[2,3,4,5], L. VILLARD[1,2].* 1) Inserm, U910, Marseille, France; 2) Aix-Marseille Université, Faculté de Médecine Timone, Marseille, France; 3) Inserm, U901, Marseille, France; 4) INMED, Parc Scientifique de Luminy, Marseille, France; 5) Plateforme postgénomique INMED INSERM, Parc Scientifique de Luminy, Marseille, France.

The NHEJ pathway repairs the DNA double strand breaks (DSBs), specifically in postmitotic cells. The NHEJ1 gene is a DNA repair factor essential in that pathway. Mutations in NHEJ1 cause immunodeficiency associated with microcephaly and growth retardation in humans. We previously identified a mutation in the NHEJ1 gene in a patient presenting a malformation of the cerebral cortex consisting of diffuse polymicrogyria (PMG) combined with heterotopic nodules, and showed that the NHEJ1 mRNA was preferentially expressed in the telencephalic ventricular and subventricular zones during the development in humans (Cantagrel et al. 2007). Using shRNAs and in utero electroporation, we now studied the impact of Nhej1 dowregulation on cell survival and cerebral cortex development. We show that Nhej1 downregulation leads to massive cell death in vivo and in vitro. We also show that Nhej1 downregulation disrupts neuronal migration visualized at embryonic stages by a massive accumulation of heterotopic neurons in the intermediate zone. We also show that Nhej1 downregulation cause abnormal development of cortical layers visualized at postnatal stages by reduction of cortical superficial layers width. The excessive cell death observed when Nhej1 is underexpressed likely explains the severe microcephaly phenotype of patients presenting a mutation in the NHEJ1 gene, and the disruption of neuronal migration and cerebral cortex development may explain the cortical malformation presenting in our PMG patient. Taken together, our results reveal that the Nhej1 protein is necessary for cortical neurons survival and proper cortical development.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2458T

***KCTD13* is a major driver of mirrored neuroanatomical phenotypes associated with the 16p11.2 copy number variant.** *C. Golzio[1], J. Willer[1], M. E. Talkowski[2,3], E. C. Oh[1], Y. Taniguchi[5], S. Jacquemont[4], A. Reymond[6], M. Sun[2], A. Sawa[5], J. S. Gusella[2,3], A. Kamiya[5], J. S. Beckmann[4,7], N. Katsanis[1,8].* 1) Center for Human Disease Modeling and Dept of Cell Biology, Duke University, Durham, NC; 2) Molecular Neurogenetics Unit, Center for Human Genetic Research, Massachusetts General Hospital, Boston MA, USA; 3) Departments of Neurology and Genetics, Harvard Medical School, Boston MA, USA; 4) Service de Génétique Médicale, Centre Hospitalier Universitaire Vaudois, Lausanne, Switzerland; 5) Department of Psychiatry, Johns Hopkins University School of Medicine, Baltimore MD, USA; 6) Center for Integrative Genomics, University of Lausanne, Lausanne, Switzerland; 7) Department of Medical Genetics, University of Lausanne, Lausanne, Switzerland; 8) Department of Pediatrics, Duke University, Durham, NC.

Copy number variants (CNVs) are major contributors to genetic disorders. We have dissected a region of the 16p11.2 chromosome— which encompasses 29 genes—that confers susceptibility to neurocognitive defects when deleted or duplicated. Overexpression of each human transcript in zebrafish embryos identified *KCTD13* as the sole message capable of inducing the microcephaly phenotype associated with the 16p11.2 duplication, whereas suppression of the same locus yielded the macrocephalic phenotype associated with the 16p11.2 deletion, capturing the mirror phenotypes of humans. Analyses of zebrafish and mouse embryos suggest that microcephaly is caused by decreased proliferation of neuronal progenitors with concomitant increase in apoptosis in the developing brain, whereas macrocephaly arises by increased proliferation and no changes in apoptosis. A role for *KCTD13* dosage changes is consistent with autism in both a recently reported family with a reduced 16p11.2 deletion and a subject reported here with a complex 16p11.2 rearrangement involving *de novo* structural alteration of *KCTD13*. Our data suggest that *KCTD13* is a major driver for the neurodevelopmental phenotypes associated with the 16p11.2 CNV, reinforce the idea that one or a small number of transcripts within a CNV can underpin clinical phenotypes, and offer an efficient route to identifying dosage-sensitive loci.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2459F

miR-137 expression in schizophrenia and bipolar disorder brains. *I. Guella, A. Sequeira, B. Rollins, F. Torri, M.P. Vawter, F. Macciardi.* Department of Psychiatry and Human Behavior, University of California Irvine, Irvine, California, United States of America.

MicroRNAs (miRNAs) are small non-coding RNAs that act as potent regulators of gene expression. In a recent GWAS the rs1625579 SNP, located downstream of miR-137, was reported as the strongest new association with schizophrenia (Ripke S et al., 2011). Two independent SZ imaging-genetic studies (Potkin SG et al., 2010) found miR-137 to be 1 of 3 miRNAs with target genes significantly enriched for association. Strikingly, this miRNA has been implicated in regulating both adult neurogenesis and neuronal maturation. We investigated the miR-137 expression levels in the dorsolateral prefrontal cortex (DLPFC) of postmortem brain tissue from 2 independent cohorts: 1) 27 subjects (10 control (CTR), 8 schizophrenia (SZ), 9 bipolar disorder (BD)) collected at the UCI Brain Bank; and 2) 99 subjects (33 CTR, 35 SZ, 31 BD) obtained from the Stanley Medical Research Institute (SMRI). Within the UCI cohort, we observed a statistically significant difference in the miR-137 expression across the 3 groups, with the BD having a higher expression (2.01 fold increase) and the SZ a lower expression (0.69 fold decrease) compared to the control group. We were not able to replicate the BD significant results in the SMRI collection, but it is interesting that in both cohorts we noticed a reduced miRNA-137 expression in SZ. The rs1625579 genotypes of 107 samples were next investigated for an association with miR-137 expression. MiR-137 was increased 1.18 fold in heterozygous TG compared to the homozygous TT subjects in a joint analysis of the 2 cohorts, although this difference did not achieve statistical significance (p=0.19). When controls only were analyzed, in the joint analysis, the miR-137 expression was 1.41 fold higher in TG subjects compared to TT (p=0.036). In conclusion, expression changes were observed for miR-137 in BD and SZ, and an effect of rs1625579 on expression was noted, but only in the control group. The lack of significance in case-control analysis of miR-137 expression is likely due to the variability of the expression levels across cohorts, and to the small sample size of the UCI cohort. We have thus far investigated the DLPFC where our results suggest a possible association between the T risk allele of rs1625579 and decreased miR-137 expression. Future studies are needed to better explain the within subject expression variability, to explore in other brain regions the miR-137 expression, and the genotype association with miR-137 expression.

2460W

Common gene co-expression network analysis in schizophrenia and bipolar disorder. *S. Hong[1], X. Chen[3], L. Jin[1], M. Xiong[2].* 1) School of Life Science, Fudan University, Shanghai, Shanghai, China; 2) School of Public Health, The University of Texas Health Science Center at Houston; 3) Departments of Psychiatry and Human Genetics, Virginia Commonwealth University.

Schizophrenia and bipolar disorders were recognized as two main diagnostic categories of the most severe psychiatric disorders. Though the psychotic features and clinical diagnosis of these two diseases were different, there were more and more studies supporting sharing some genetic mechanisms between them. In the previous studies, most of the co-expression network construction methods only focused on using a Gaussian graphical model to estimate a single co-expression network for each type of tissue samples. However, those methods cannot unmask the underlying heterogeneity between networks with some shared structures and estimate a co-expression network shared by two types of tissue samples or two diseases. Recently, a Joint Estimation of Multiple Graphical model to preserve the common structure of networks from heterogeneous data has been developed. We adapted this model to our study and constructed the co-expression networks underlying both schizophrenia and bipolar disorder. RNA-seq was applied to 31 schizophrenia, 25 bipolar and 26 normal samples in this study. We identified 76 differential expressed genes between metal disorders (schizophrenia and bipolar) and normal with q-value less than 0.01. Then, we used a joint multiple graphical model to identify a co-expression network shared by both schizophrenia and bipolar disorder. Most of genes in the common network were previously reported to be suspected genes to both schizophrenia and bipolar, such as SNAP25, PVALB, RGS4, GABRA1 and NPAS3. Our results showed that this method was very useful to reveal the shared network between diseases which were supposed to share the same etiology.

2461T

Delusion-associated SNPs in NRG3 Show Regulatory Potential, Dysregulate NRG3 Splicing and Differentially Bind to Nuclear Proteins. *M. Zeledón[1,2,3], M. Taub[5], N. Eckart[1,2], R. Wang[1], M. Szymanski[1,2], P. Chen[6], A. Pulver[3,4], J.A. McGrath[3,4], P. Wolyniec[3,4], D. Avramopoulos[1,3], A. Sawa[3], D. Valle[1].* 1) McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD; 2) Predoctoral Training Program in Human Genetics, Johns Hopkins University School of Medicine, Baltimore, MD; 3) Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore, MD; 4) Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD; 5) Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD; 6) Department of Medical Genetics, National Taiwan University Hospital, Taipei City, Taiwan.

Schizophrenia (SZ) is a severely disabling psychiatric disease that affects 1% of the world's population. Symptoms include dysfunctions in cognition, thought/belief (delusions), perception/sensation (hallucinations) and affect. Previously, our group reported a linkage peak for SZ (NPL of 4.7) at 10q22 in the Ashkenazi Jewish (AJ) population and, in a follow-up fine mapping association study in the AJ, found strong evidence of association between a quantitative phenotypic trait “delusion” and three intronic SNPs in the 5' end of *NRG3* (Fallin *et al*, 2003; Chen *et al*, 2009). Two other independent groups have replicated our findings, making *NRG3* a strong candidate gene for a subtype of schizophrenia with delusions. To identify causative variants, we sequenced the 162 kb LD block covering the 5' end of *NRG3* and containing the three associated SNPs in 47 AJ SZ patients at either extreme of the delusion quantitative trait. We identified 5 SNPs with minor alleles on the implicated haplotype and significantly overrepresented in high delusion patients. We tested these for regulatory potential. In transient transfection assays with dual luciferase reporters, we found that the delusion associated alleles of rs10883866 and rs60827755 showed significantly and reproducibly decreased and increased (3–4 fold) expression of a reporter gene respectively, compared to the common allele. This effect was consistent in a variety of cell types including primary rat cortical neurons. To test whether these changes are due to differential binding of DNA regions containing the alleles to nuclear protein, we performed electrophoretic mobility shift assays. We determined that the 21 bp including and surrounding the SNPs consistently bind to nuclear proteins and binding is influenced by the genotype of the SNP. These results are consistent with data from Eckart *et al*. (abstract submitted to this conference) that suggest SNPs in *NRG3* perturb expression of *NRG3* isoforms. In summary, we have identified cis-acting regulatory motifs that modify *NRG3*expression and that account for the association with delusions, have regulatory potential, and bind to nuclear proteins. We are currently addressing the identity of the nuclear proteins that bind these regulatory motifs. We anticipate that these experiments will allow us to move from genetic observations to deciphering the mechanistic pathways that may lead to risk for SZ or modifications of the SZ phenotype.

2462F

USF1 regulates sleep and circadian traits in humans. *H.M. Ollila[1], K.M. Rytkönen[2], K. Silander[1], M. Perola[1], V. Salomaa[1], M. Härmä[3], S. Puttonen[3], T. Porkka-Heiskanen[2], T. Paunio[1].*1) National Institute for Health and Welfare, Finland, Helsinki, Finland; 2) Institute of Biomedicine, Physiology, University of Helsinki, Finland; 3) Finnish Institute for Occupational Health, Helsinki, Finland.

OBJECTIVES: Epidemiological studies show association between sleep duration and lipid metabolism. Our aim was to characterize gene variants that regulate both metabolism and sleep. One such candidate is USF1 that regulates the expression of several lipid genes and has a similar structure as canonical circadian transcription factors. We studied the polymorphisms in USF1 gene region with sleep duration and coping with circadian stress in Finnish population based samples combined with gene expression and USF1 knock out animal model. MATERIALS AND METHODS: The known USF1 variants that had previously been associated with lipid traits were studied with sleep and coping with circadian stress. The analyses were performed in a Finnish population based sample with no sleep problems (N=1085) and Finnair workers (N=1415). RNA expression from mononuclear leucocytes was measured in additional 584 individuals. EEG recordings from USF1 knock-out animals were performed. RESULTS: The genetic analysis identified that the USF1 polymorphisms associated with sleep duration in the healthy individuals. The same variant associated with coping with circadian stress in Finnair workers. Gene expression analysis showed that those individuals carrying minor allele had higher USF1 expression. Finally, EEG analysis showed distinct changes between USF1 knockout and wild type littermates. CONCLUSIONS: Our results show that the allelic variants of USF1 associate with sleep duration and coping with circadian stress. The finding may reflect the shared roots of sleep and metabolism. The shared genetic background between sleep and metabolism may at least partially explain the mechanism behind the well-established connection between diseases with disrupted metabolism and sleep.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

† Trainee Award Finalist

2463W

Sodium channel mutations and epilepsy. *J.E. O'Brien, M.H. Meisler.* Department of Human Genetics, University of Michigan, Ann Arbor, MI.

Voltage-gated sodium channels are members of a large gene family that includes the neuronal channels *SCN1A*, *SCN2A*, *SCN3A*, *SCN8A*, *SCN9A* and *SCN10A*. Mutations of *SCN1A* are a major cause of Dravet syndrome, a severe early-onset epileptic encephalopathy. *SCN2A* has been implicated in generalized epilepsy with febrile seizures plus and benign familial neonatal-infantile seizures, *SCN3A* in cryptogenic pediatric partial epilepsy, and *SCN9A* in other febrile seizures. A heterozygous null mutation in *SCN8A*, one of the most abundant sodium channels in the brain, was found in a family with cognitive impairment (Trudeau et al, J Med Genet 2006), but no seizure phenotypes were previously attributed to human *SCN8A*. We studied a proband with epileptic encephalopathy characterized by early-onset, features of autism, intellectual disability, ataxia, and SUDEP (sudden unexplained death in epilepsy) (Veeramah et al, AJHG 2012). Whole genome sequencing of the proband, her parents, and her unaffected brother identified a *de novo* heterozygous missense mutation in *SCN8A*, c.5302A>G, (p.Asn1768Asp), that alters an evolutionarily conserved residue. The mutant channel demonstrated hyperexcitability consistent with the dominant phenotype in the heterozygous proband. *SCN8A* contains an alternatively-spliced, neuron-specific exon that is essential for generation of active channel protein. We found that the Rbfox splice factors contribute to neuron-specific generation of full length channel protein (O'Brien et al, Mol Cell Neurosci 2012). We also identified a novel interaction of the cytoplasmic N-terminus of *SCN8A* with the microtubule-associated protein Map1b that appears to function in channel trafficking (O'Brien et al, J Biol Chem 2012). *SCN8A* is the fifth sodium-channel gene found to contrubute to seizure disorders. Generation of a mouse model carrying the p.Asn1768Asp mutation is in progress. (Supported in part by NS34509 and GM007544.).

2464T

Phosphorylation and Regulation of FMRP. *M. Santoro[1], S. Warren[1,2].* 1) Department of Human Genetics, Emory University School of Medicine, Atlanta, GA; 2) Departments of Biochemistry and Pediatrics, Emory University School of Medicine, Atlanta, GA.

Fragile X syndrome, the most common form of inherited intellectual disability, results from a lack of Fragile X Mental Retardation Protein (FMRP). FMRP is a selective mRNA-binding protein and plays a key role in synaptic plasticity. In neurons, FMRP represses translation of its target mRNAs but allows translation to proceed upon receipt of specific synaptic signals. In the hypothalamus, activation of mGluR receptors causes FMRP to derepress translation of its target mRNAs, resulting in a burst of local protein synthesis, internalization of AMPA receptors, and long term depression. In vitro experiments indicate that FMRP's activity is regulated by the phosphorylation status of a specific serine (S499 in mice; S500 in humans). In order to more fully investigate the translational regulation of FMRP we are developing Fmrp knock-in mice containing phosphomimic forms of FMRP. A Ser499Ala mutation results in an amino acid replacement that structurally resembles unphosphorylated FMRP, whereas a Ser499Asp mutation structurally mimics the phosphorylated serine. The Ser499Ala strain has been created while the Ser499Asp strain is currently under development. Ser499Ala is expected to phenocopy a null mutation - increased translation of FMRP targets, excessive AMPAR internalization, and overactive mGluR signaling. The Ser499-Asp protein is expected to show the opposite phenotype but, since the corresponding mutation has not been seen in humans, it is difficult to predict what the phenotype might be. We will report biochemical and neurobiological aspects of these two novel Fmr1 mutant mice.

2465F

The *cis*-regulatory effect of an Alzheimer's disease-associated polymorphic poly-T locus on expression of *TOMM40* and *APOE* genes. *O. Chiba-Falek[1,2,3], C. Linnertz[1], W. Gottschalk[1,2], D. Crenshaw[1,2], M. Lutz[1,2], J. Allen[3], S. Saith[3], M. Mihovilovic[1], L. Anderson[1], K. Welsh-Bohmer[2], A. Roses[1,2,4].* 1) Dept Med/Neurology, Duke Univ, Durham, NC; 2) Joseph and Kathleen Bryan ADRC, Duke Univ, Durham, NC; 3) Institute for Genome Sciences & Policy, Duke Univ, Durham, NC; 4) Zinfandel Pharmaceuticals, Chapel Hill, NC.

Differential gene expression controlled, in part, by polymorphisms in non-coding regions of the genome, may contribute to the pathogenesis of complex human traits including late-onset Alzheimer's disease (LOAD). We studied the TOMM40-APOE genomic region that has been associated with the risk and age of disease onset of LOAD to determine if a highly polymorphic intronic poly-T within this region (rs10524523, hereafter `523') affects the expression of APOE and TOMM40 genes. Alleles of this poly-T variant are classified as short (S), long (L) or very long (VL) based on the number of T residues. We evaluated differences in APOE-mRNA and TOMM40-mRNA levels as a function of `523' genotype in two brain regions from APOE ε3/3 Caucasian autopsy confirmed LOAD cases and normal controls. Expression levels of both genes were significantly increased with disease. In temporal cortex (TC) from both Normal (n = 42) and LOAD (n = 66) subjects, APOE and TOMM40 mRNA levels were higher in `523'VL-homozygotes compared to S-homozygotes. The association between higher TOMM40 and APOE expression and the VL allele was also observed in matched occipital (OCC) tissues (n=34 Normal and 69 LOAD subjects). Differences in mean expression according to the `523' genotype in the TC and OCC samples for both TOMM40 and APOE mRNAs approached statistical significance. Expression levels for both mRNAs decreased with age, but the contribution of age to the variation in transcript level was not statistically significant. Gender, tissue source or PMI were not associated with transcript level. We used a luciferase reporter system to test the effect of the `523' locus, in the native genomic context, on gene expression. In agreement with the human brain mRNA analysis, the VL (34 T residues) `523' polyT resulted in significantly higher expression than the S (16 T residues) polyT in both HepG2 cells and SH-SY5Y neuroblastoma cells. This effect was greater in the SH-SY5Y cells than in the hepatoma cells, hinting at tissue-specific modulation of the `523' polyT locus. These results suggest that the `523' locus may contribute to LOAD susceptibility by modulating the expression of TOMM40 and/or APOE transcription. Elucidating the mechanism by which this locus affects LOAD may lead to the identification of novel, tractable drug targets or modalities for treating, preventing or delaying onset of this devastating disease.

2466W

Apolipoprotein E ε4 as a factor in the evaluation of the genetic risk of Alzheimer's dementia in Filipinos. *M. Daroy[1], C. Casingal[1], C. Mapua[1], D. Florendo[1, 2], R. Matias[1, 3], J. Dominguez[2], St. Luke's Dementia Study Group.* 1) Research and Biotechnology Div, St. Luke's Medical Center, Quezon City, NCR, Philippines; 2) Memory Center, International Institute of Neurosciences, St. Luke's Medical Center, Quezon City, NCR, Philippines; 3) Medical Affairs Division, United Laboratories, Inc., Mandaluyong City, NCR, Philippines.

Alzheimer's dementia is a growing threat to the rapidly increasing elderly population worldwide. In the Philippines, there are 3.8 million people above 65 years old who are at risk of developing AD. This report presents the results of a cross-sectional study to determine the genotypic distribution of ApoE in affected and unaffected groups of Filipinos, who either consulted in the hospital for Alzheimer's dementia, or were screened in a community setting. Peripheral blood was collected from a total of 600 subjects, composed of 124 patients who consulted at the Memory Center of St. Luke's Medical Center, Quezon City and 476 individuals who participated in a community-based cohort. DNA from the buffy coat was analyzed for ApoE genotype by PCR-RFLP (Hha I) method. The most frequent genotype for both groups was ε3/ε3: 51.6% for the hospital group and 66.6% for the community group. The clinical profile of all the subjects is: 96 with Alzheimer's dementia, 110 with mild cognitive impairment, 369 with no dementia, and 25 with diagnoses other than dementia. In this group of Filipinos, the ApoE ε4 allele is significantly associated with risk for Alzheimer's dementia, compared with the no dementia group, with a p-value of 0.0481 and an odds ratio of 1.611. The frequencies of the risk ε4 allele were as follows: 14.1% for those diagnosed with Alzheimer's dementia, 9.1% for those with mild cognitive impairment, 9.2% for those with no dementia, and 6.0% for those with diagnoses other than dementia. These results suggest that the detection of the presence of the ApoE ε4 allele could be a useful tool for the assessment of genetic risk for Alzheimer's dementia in Filipinos when integrated with the patient's clinical data and family history.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2467T

Functional Testing of Schizophrenia Associated Variants in *NOS1AP*. *CT. Ramirez[1], N. Wratten[1], Y. Huang[2], A. Dulencin[3], J. Millonig[3], VJ. Vieland[2], AS. Bassett[4], LM. Brzustowicz[1].* 1) Dept Gen, Rutgers Univ, Piscataway, NJ; 2) Battelle Ctr Math Med, Res Inst Nationwide Child Hosp, Columbus, OH; 3) Department of Neuroscience and Cell Biology, University of Medicine and Dentistry of New Jersey, Piscataway, NJ; 4) Department of Psychiatry, University of Toronto and Centre for Addiction & Mental Health, Toronto, Ontario, Canada.

Schizophrenia is a major neuropsychiatric disorder that affects approximately one percent of the general population. Although the etiology of schizophrenia is unknown, family, twin and adoption studies have demonstrated that schizophrenia is a predominantly genetic disorder with a high heritability. We previously reported highly significant evidence of linkage to chromosome 1q21–23 in a set of Canadian families of European descent, identified *NOS1AP* as the candidate gene from this region contributing to risk of schizophrenia in our sample, and identified three SNPs with strong evidence of association to schizophrenia. These SNPs were tested for regulatory functions using a Dual Luciferase Reporter Assay with the native *NOS1AP*-promoter. rs12742393 demonstrated regulatory activity and produced significant allelic expression differences in two different human neural cell lines, SK-N-MC and PFSK-1. Statistical analysis under the PPL framework of eight SNPs in strong LD with rs12742393 revealed a haplotype block of four SNPs in complete LD with each other (rs4557949-rs1158803-rs4145621-rs12084492) and with strong evidence of association to schizophrenia. Additional statistical analysis revealed that a haplotype consisting of rs12742393 and this block produced even stronger evidence for association than rs12742393 or this block alone, suggesting that an interaction between two or more SNPs could modulate NOS1AP gene expression. rs4145621 has been tested for regulatory activity and produced no evidence of being a functional variant in the human cell lines. Preliminary data suggest that rs4557949 is also not a functional variant.Further testing of the two remaining SNPs in the haplotype and all four SNPs using different cell confluency levels may reveal potential contributions to regulatory function that could be important in schizophrenia expression.

2468F

***MIR137*, a candidate gene for schizophrenia: identification of targets and downstream effects.** *A.L. Collins[1], R.J. Bloom[1], Y. Kim[1], P.F. Sullivan[1,2].* 1) Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC; 2) Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, NC.

Findings from the Psychiatric GWAS Consortium (PGC) have identified significant and replicated association for a region on chromosome 1 near *MIR137* (rs1625579; $p=1.6\times10^{-11}$; the second top hit in the PGC GWAS). Unpublished studies increase the significance of this finding to $p=1.3\times10^{-12}$. Genome-wide significant associations have also been identified for genes with confirmed miR-137 target sites (*CACNA1C*, *TCF4*, *CSMD1*, and *C10orf26*). miR-137 is implicated in other psychiatric disorders (intellectual disability with autism spectrum disorders, Alzheimer's, and mouse models of Rett syndrome). Combined with indications that miR-137 plays an important role in neuronal differentiation and proliferation, it is imperative that we understand miR-137 function. Therefore, we aim to identify potential (direct or indirect) targets of miR-137, to elucidate a pathway altered in schizophrenia. Human neural stem cells (ReNcell-VM) were transduced with lentiviral vectors for overexpressing and inhibiting miR-137 (and the respective controls). Transduction efficiencies approached 100% for all vectors. For overexpression, miR-137 levels are increased by ≥ 40X. For inhibition, miR-137 is sequestered, therefore testing of established targets of miR-137 (e.g., *MIB1* and *MITF*) is in progress to verify the functionality of the inhibition. We are now testing RNA and protein levels of candidate genes (predicted targets of miR-137 plus genes implicated by schizophrenia GWAS) to determine the impact of miR-137 dysregulation. To identify novel downstream effects, we completed RNA-seq from cells transduced with the miR-137 overexpression vector or control (analysis in progress). RNAseq for inhibition is in progress. Comparison of predicted binding sites and direction of alteration will yield a candidate list of direct versus indirect targets. miRNAs are intriguing candidates in complex disease, as the ability to regulate hundreds of genes may allow identification of many other contributors to disease pathogenesis. Moreover, identification of members of the network regulated by miR-137 could identify critical participants in schizophrenia pathogenesis. Additional work to identify translationally inhibited targets will be needed to further our understanding of this micro-RNA. Nonetheless, elucidation of a pathway or pathways regulated by miR-137 could provide a myriad of potential targets for pharmaceutical intervention.

2469W

Schizophrenia mir137 Risk Genotype is Associated with DLPFC Hyperactivation. *T.GM. van Erp, I. Guella, M.P. Vawter, F. Torri, J. Fallon, F. Macciardi*, S.G. Potkin*.* Department of Psychiatry and Human Behavior, UCI, Irvine, CA.

Background. A recent large genome-wide association study (Ripke et al. 2011) confirmed findings from a prior gene set enrichment analysis (Potkin et al. 2009) in implicating MIR137-dysregulation in the etiology of schizophrenia, but the functional sequelae of MIR137 dysregulation remain to be determined. **Methods.** Functional magnetic resonance imaging scans on 50 schizophrenia patients and 66 age and sex-matched healthy volunteers were acquired while subjects performed a Sternberg Item Response Paradigm with memory loads of 1, 3, and 5 items. Functional image processing was performed as previously reported (Potkin et al. 2009). We extracted dorsolateral prefrontal cortex (DLPFC) retrieval activation for the working memory load of 3 items, for which hyperactivation had been shown in schizophrenia patients compared with controls. The implicated MIR137 SNP (rs1625579) was genotyped (schizophrenia: GG n=0, GT n=10, TT n=40; healthy volunteers: GG n=2, GT n=15, and TT n=49). Frequency distributions of MIR137 (rs1625579) alleles were examined using the Fisher's Exact Test. Mixed model regression analyses were performed to examine the effects of diagnosis and genotype on DLPFC activation while controlling for effects of sex, age, and site. **Results.** Patients did not have a higher percentage of MIR137 risk (rs1625579 T) alleles than controls. Patients showed significantly higher DLPFC activations compared with controls. Individuals with a MIR137 TT genotype showed significantly higher DLPFC activations than those with a GG/GT genotype, independent of diagnosis. **Conclusion.** Compared with individuals with the MIR137 GG or GT genotype, individuals with the MIR137 TT genotype show DLPFC hyperactivation as measured with functional magnetic resonance imaging during working memory. The MIR137 TT schizophrenia risk genotype is associated with the schizophrenia risk phenotype DLPFC hyperactivation commonly considered a measure of brain inefficiency.

* shared senior author.

2470T

A Loss-of-Function Variant in the Human Histidyl-tRNA Synthetase (HARS) Gene is Neurotoxic In Vivo. *A. Antonellis[1,2], A. Vester[1], G. Velez-Ruiz[3], H.M. McLaughlin[1], J.R. Lupski[4,5,6], K. Talbot[7], J.M. Vance[8], S. Zuchner[8], R.H. Roda[9], K.H. Fischbeck[9], L.G. Biesecker[10], G. Nicholson[11,12], A.A. Beg[3], NISC Comparative Sequencing Program, NIH, Bethesda, MD.*1) Human Genetics, University of Michigan, Ann Arbor, MI; 2) Neurology, University of Michigan, Ann Arbor, MI; 3) Pharmacology, University of Michigan, Ann Arbor, MI; 4) Molecular and Human Genetics, Baylor College of Medicine, Houston, TX; 5) Pediatrics, Baylor College of Medicine, Houston, TX; 6) Texas Children's Hospital, Houston, TX; 7) Department of Clinical Neurology, University of Oxford, Oxford, UK; 8) Hussman Institute for Human Genomics, University of Miami Miller School of Medicine, Miami, FL; 9) Neurogenetics Branch, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD; 10) Genetic Disease Research Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD; 11) Northcott Neuroscience Laboratory, ANZAC Research Institute and Molecular Medicine Laboratory, Concord Hospital, Concord, New South Wales, Australia; 12) Faculty of Medicine, University of Sydney, Camperdown, New South Wales, Australia.

Aminoacyl-tRNA synthetases (ARSs) are ubiquitously expressed enzymes responsible for ligating amino acids to cognate tRNA molecules. Mutations in four genes encoding an ARS have been implicated in inherited peripheral neuropathy with an axonal pathology, suggesting that all ARS genes are relevant candidates for disease in patients with related phenotypes. Here, we present results from a mutation screen of the histidyl-tRNA synthetase (*HARS*) gene in a large cohort of patients with peripheral neuropathy. These efforts revealed a rare missense variant (p.Arg137Gln) that resides at a highly conserved amino acid, represents a loss-of-function allele when evaluated in yeast complementation assays, and is toxic to neurons when expressed in*C. elegans*. In addition to the patient with peripheral neuropathy, p.Arg137Gln *HARS* was detected in three individuals by genome-wide exome sequencing. We therefore conclude that p.Arg137Gln *HARS* is a pathogenic allele with incomplete penetrance. These findings suggest that *HARS* is the fifth ARS locus associated with axonal peripheral neuropathy. Implications for identifying ARS alleles in human populations and assessing them for a role in neurodegenerative phenotypes will be discussed.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2471F

Neuronal expression of FIG4 and neurodegeneration. *G.M. Lenk[1], C.J. Ferguson[2], M.H. Meisler[1].* 1) Human Gen, Univ Michigan, Ann Arbor, MI; 2) Neuroscience Program, Univ Michigan, ANn Arbor, MI.

FIG4 is a ubiquitously expressed phosphatase that regulates the level of the signaling lipid PI(3,5)P2. Null mutation of *Fig4* in the mouse results in spongiform degeneration of brain and peripheral ganglia, defective myelination, and juvenile lethality. Recessive mutations of human *FIG4* result in a severe form of Charcot-Marie-Tooth neuropathy (CMT4J). Most patients with CMT4J are compound heterozygotes with one null allele and the missense mutation Ile41Thr. This mutation impairs interaction with the scaffold protein VAC14, destabilizing the FIG4 protein (Lenk et al, PloS genetics, 2011). Neurons from *Fig4* mutant mice contain enlarged vacuoles derived from the endosome/lysosome pathway and astrocytes accumulate proteins involved in autophagy. We used transgenic mouse models to distinguish the contributions of neurons and glia to spongiform degeneration in the *Fig4* null mouse. Wildtype *Fig4* expression was directed by the neuron-specific NSE promoter and the astrocyte-specific GFAP promoter in *Fig4* null transgenic mice. Neuron specific expression of *Fig4* was sufficient to rescue the cellular and neurological phenotypes including spongiform degeneration, gliosis, and juvenile lethality (Ferguson et al, HMG, 2012). CNS hypomyelination was also rescued by the neuron-specific transgene (Winters et al, J Neurosci, 2011). Expression of Fig4 in astrocytes prevented accumulation of autophagy markers in acstrocytes, but this did not prevent spongiform degeneration or lethality. To confirm the role of neurons in the spongiform degeneration, we generated a floxed allele of *Fig4* and crossed it with mice expressing the Cre recombinase from the neuron-specific synapsin promoter. Mice with conditional inactivation of *Fig4* in neurons developed spongiform degeneration by 30 days and exhibit premature lethality. The data demonstrate that expression of *Fig4* in neurons is necessary and sufficient to prevent spongiform degeneration. Therapy for patients with FIG4 deficiency will therefore require correction of the deficiency in neurons. (Supported by GM24872 and The Hartwell Foundation).

2472W

Global mapping of FUS-binding sites and global profiling of FUS-mediated RNA metabolisms in the mouse brain. *K. Ohno[1], A. Masuda[1], S. Ishigaki[2,4], Y. Fujioka[2], Y. Iguchi[2], M. Katsuno[2], A. Shibata[1], F. Urano[3], G. Sobue[2,4].* 1) Neurogenetics, Nagoya University Graduate School of Medicine, Nagoya, Japan; 2) Neurology, Nagoya University Graduate School of Medicine, Nagoya, Japan; 3) Program in Gene Function and Expression, University of Massachusetts Medical School, Worcester, MA; 4) CREST, Japan Science and Technology Agency, Kawaguchi, Japan.

Aberrations of FUS are causally associated with familial and sporadic amyotrophic lateral sclerosis and frontotemporal lobar degeneration (ALS/FTLD). FUS is a multifunctional protein that has been involved in several steps of gene expression regulation especially for transcription and RNA splicing. FUS belongs to the FET family of RNA-binding proteins, which includes FUS, Ewing's sarcoma (EWS), and TATA-binding protein-associated factor (TAFII68). FUS is present in polymerase II transcription complexes that function in the transcription process. Interestingly, FUS inhibits the acetyltransferase activities of CREB-binding protein (CBP) and p300 on cyclin D1 (*Ccnd1*) in HeLa cells. Both CBP and p300 are co-activators of multiple classes of signal-dependent transcription factors and the FUS/CBP/p300 interactions result in the inhibition of histone-acetyltransferase activities followed by repression of transcription. Moreover, FUS is involved in the splicing machinery by cooperating with other splicing factors like serine/arginine-rich (SR) proteins, SRm160, and PTB. These observations suggest that compromised effects of FUS on transcription and alternative splicing could lead to neuronal cell degeneration in FUS-associated ALS and FTLD.

In an effort to understand the global roles of FUS on RNA metabolisms in neuronal cells, we analyzed exon arrays of mouse primary cortical neurons after knocking down *Fus*. We also identified FUS-binding RNA segments in the mouse cerebrum using the high-throughput sequencing of RNA isolated by crosslinking immunoprecipitation (HITS-CLIP). We observed no distinct FUS-binding motifs in the CLIP tags. We, however, found that FUS-binding sites tend to form stable secondary structures. Positional analysis of FUS-binding sites in relation to alternative splicing events disclosed scattered binding of FUS to and around the alternatively spliced exons with noticeable FUS-binding peaks in the downstream introns. We found that FUS are bound to the antisense RNA strand at the promoter regions. Global analysis of these FUS-tags and the expression profiles disclosed that binding of FUS to the promoter antisense strand downregulates transcriptions of the coding sense strand. Our analysis revealed that FUS regulates alternative splicing events and transcriptions in a position-dependent manner.

2473T

***Sortilin-related receptor 1* is associated with cognitive impairment in Filipinos.** *C.R. Casingal[1], M.L.G Daroy[1], C.A. Mapua[1], D.J.A. Florendo[1,2], J.C. Dominguez[1], St. Luke's Dementia Study Group.* 1) Research and Biotechnology Division; 2) Memory Center, Institute of Neuroscience, St. Luke's Medical Center, Quezon City, Philippines.

This study investigated the association of 6 *sortilin-related receptor 1 (SORL1)* variants with cognitive impairment. SORL1 SNPs 8, 9, 10, 13, 19 and 23 were genotyped in 484 Filipinos, including 100 Alzheimer's disease (AD) cases, 109 mild cognitive impairment (MCI) cases, 257 with no dementia and 18 with other cases of cognitive impairment, using TaqMan® SNP Genotyping Assays. Cognitive impairment (CI), which includes AD, MCI and other cognitive impairment cases, was associated with SNP 23 (p=0.041). None of the *SORL1* variants was associated with AD. It may be that the association was not detected because of inadequate sample size. SNPs 8, 9, 10 and 23 were associated with MCI, (p=0.028, p=0.034, p=0.034 and p=0.025, respectively). Based on these results, *SORL1* may be used as a biomarker in the early detection or diagnosis of cognitive impairment and other dementias. This study also found an association between SNP 23 and female sex in MCI group, (p=0.033). This suggests that *SORL1* may affect cognitive impairment and the interplay of sex-specific risk factors, sex-specific disease course, and sex-specific survival of a disease, through a female-dependent mechanism. The subjects were also evaluated based on the presence of *APOE-ε4* allele. The findings showed that in the *APOE-ε4* noncarrier group, SNPs 8, 9, 10, 19 and 23 were associated with MCI group, (p=0.030, p=0.050, p=0.050, p=0.031 and p=0.007, respectively). These results support the findings that *SORL1* contributes to an increased risk for cognitive impairment even when the *APOE-ε4* allele is not present.

2474F

Identifying neuropsychological and cognitive endophenotypes of schizophrenia-associated exonic variants of NPAS3 and COMT. *L.M. Luoma[1], G. Macintyre[1], F.B. Berry[1,2], D. LaFreniere[3,4], S. Purser[3,4], A. Beierbach[4], P. Tibbo[3,5], D.W. Cox[6], S. Purdon[3,4].* 1) Department of Medical Genetics, University of Alberta, Edmonton, AB, Canada; 2) Department of Surgery, University of Alberta, Edmonton, AB, Canada; 3) Department of Psychiatry, University of Alberta, Edmonton, AB, Canada; 4) Neuropsychology, Alberta Hospital Edmonton, Edmonton, AB, Canada; 5) Department of Psychiatry, Dalhousie University, Halifax, NS, Canada; 6) Department of Medicine, University of Alberta, Edmonton, AB, Canada.

Schizophrenia is a disorder of psychosis characterized by deficits in neurodevelopment, neurotransmission and adult neurogenesis. Although schizophrenia is highly heritable, GWASs to date have identified many variants with low effect size, often with little to no replication. This indicates that genetic heterogeneity may contribute to variability in clinical presentation and therapeutic response of persons with schizophrenia. Identification of specific sub-clinical quantifiable phenotypes (endophenotypes) associated with specific genetic variants would thus be invaluable to investigations of pathogenesis and outcome. Studies of schizophrenia-associated genetic variants have demonstrated links to personality traits, cognitive skills, and discrete behaviours in normal populations. These characteristics have further been associated with vulnerability to psychopathologies such as psychosis. Variants of catechol-O-methyltransferase (*COMT*), involved in metabolism of catecholamine neurotransmitters, affect dopaminergic neurotransmission, cortical thickness, cognitive flexibility and working memory, each of which have been found disrupted in schizophrenia. NPAS3 (neuronal PAS domain containing 3) is a schizophrenia-associated transcription factor implicated in neurodevelopment, adult neurogenesis and neurosignalling. Loss of Npas3 in mice causes altered adult neurogenesis and behaviours which mimic features of schizophrenia. Furthermore, *NPAS3* variants are linked to cross-disorder pleiotropic effects in humans. Endophenotypes of *NPAS3* variants have not yet been identified. Eighty-seven healthy teenagers (mean age 17.68 years) were recruited for cognitive, neuropsychological and genetic analyses to identify endophenotypes linked to schizophrenia-associated exonic variants of *COMT* and *NPAS3*. Individuals carrying the schizophrenia-associated allele of the *COMT* rs4680 (p.V158M) and/or *NPAS3* (p.A552P) variants were identified. The *NPAS3* risk genotype was associated with verbal working memory (p=0.004), and the general index score of the Screen for Cognitive impairment in Psychiatry (SCIP, p=0.039). The *COMT* risk genotype was associated with social anhedonia (p=0.008) and errors on a continuous performance test (p=0.032). In summary, we have identified associations between *NPAS3* and *COMT* risk genotypes and schizophrenia-associated neuropsychological and cognitive parameters that appear to be good candidates for endophenotype markers for psychosis.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2475W

Production of IL-6 is altered in primary human macrophages carrying 255Adel *PARK2* gene variation. *R.H.M. Sindeaux[1], A.C. Senegaglia[2], R.P. Munhoz[3], H.A. Teive[3], S. Raskin[1], M.M.A. Stefani[4], M.T. Mira*[1], V.S. Sotomaior*[1].* 1) Core for Advanced Molecular Investigation, Pontifícia Universidade Católica do Paraná/PUCPR, Brazil; 2) Laboratory of Experimental Cell Culture, Pontifícia Universidade Católica do Paraná/PUCPR, Brazil; 3) Department of Neurology, Hhospital de Clínicas, Universidade Federal de Goiás/UFPR, Brazil; 4) Tropical Patholgy at Public Health Institute, Universidade Federal de Goiás/UFG, Brazil.

Variations of the *PARK2* gene that encodes parkin, an E3 ubiquitin-protein ligase, have been implicated as causative of an autosomal recessive early-onset form of Parkinson's disease (AR-EOPD). Multiple biological roles have been assigned to parkin, including modulation of inflammatory responses. Studies have demonstrated altered circulating levels of cytokines, including IL-6, in Parkinson's disease (PD) patients, suggesting a possible involvement of inflammatory/immune-mediated mechanisms in PD pathogenesis. However, direct evidence linking*PARK2* variations to immune response in the context of PD is yet to be produced. The aim of this initiative was to use monocyte-derived macrophages (MDMs) from AR-EOPD patients with*PARK2* spontaneous variations as a model to investigate aspects of immune response. Primary peripheral blood mononuclear cells (PBMCs) isolated from three AR-EOPD patients with*PARK2* variations (P1, presenting exclusively intronic variations; P2, homozygous and P3, heterozygous for a severe, nonsense exon 2 255Adel mutation) and four healthy controls were cultured in vitro for 5 days in order do generate MDMs (approximately 1.5×10^5 cells/ well). An aliquot of the cells was treated with IFNγ + LPS for 24 hours before collection of culture supernatants for IL-6 quantification by ELISA. Data were expressed as the mean ± SD of two assays. MDMs carrying *PARK2* variant 255Adel presented increased IL-6 basal levels (P2: 168.96 ± 1.97 pg/mL; P3: 385.93 ± 2.11 pg/mL) as compared with cells obtained from patient P1 and controls (mean ± SD: 8.91 ± 7.45 pg/mL). When MDMs were activated with IFNγ + LPS, the production of IL-6 by P1 and controls cells increased and IL-6 concentrations became similar across all samples (mean ± SD of all samples: 301.43 ± 18.08 pg/mL). In conclusion, we obtained the first evidence of a biological function for 255Adel *PARK2* variation in the context of AR-EOPD pathogenesis. * These authors share senior authorship; This study was supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Brazil.

2476T

Broadening the role of FMRP targets across autism spectrum disorders. *J. Steinberg[1,2], C. Webber[1].* 1) MRC Functional Genomics Unit, University of Oxford, Oxford, United Kingdom; 2) Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom.

Autism spectrum disorder (ASD) is a complex neurodevelopmental condition with an early age of onset and a prevalence of 0.3%–1%. Although ASD is highly heritable, with studies documenting a role of both *de novo* point mutations and structural variation, the underlying disease mechanism is still unknown. Recently, four studies used exome sequencing to identify *de novo* mutations in patients with autism. Of the identified candidate genes, at least 20% are regulated by FMRP, a protein thought to play a key role in neuroplasticity and neuronal translation. FMRP has been shown to regulate a total of 842 genes via stalling of ribosomal translocation and a significant overlap of these targets with ASD candidate genes has been reported. Furthermore, a loss of function mutation in the gene encoding FMRP itself causes Fragile X syndrome, the leading monogenic cause of ASD.

We have further explored the role of FMRP targets in ASD gene disruptions. Firstly, we use human brain expression data to divide the 842 FMRP targets into four robust modules with distinct spatio/temporal expression patterns, and carry out a functional analysis of each module. Subsequently, we show that both genes disrupted in ASD via balanced chromosomal abnormalities and genes implicated via *de novo* point mutations are enriched in FMRP targets that belong to a particular module showing developmental stage-specific expression. Secondly, we develop a novel test to analyse case/control datasets for disruptions in gene sets of interest. Importantly, this test is sensitive to a multi-hit disease aetiology, statistically robust and has increased power to detect pathway associations compared to existing methods. Applying this test to rare copy number variation (CNVs) in 872 ASD cases and matched unaffected siblings, we provide further evidence for the implication of FMRP targets and a second module enriched in synaptic functioning in ASD. We replicate this result in an expanded CNV dataset of 1124 ASD cases compared to their parents, and in an independent CNV dataset of 561 cases and 1146 unrelated controls.

In summary, we show that different biologically-meaningful modules of FMRP targets play roles in ASD via *de novo* point mutations, balanced chromosomal abnormalities and CNVs.

2477F

Expanded Hexanucleotide GGGGCC Repeat RNA associated with ALS/FTD is sufficient to cause neurodegeneration. *Z. Xu[1,2], M. Poidevin[1], H. Li[2], T. Wingo[3], P. Jin[1].* 1) Department of Human Genetics, Emory University School of Medicine, Atlanta, GA, USA; 2) Division of Histology and Embryology, Department of Anatomy, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, People's Republic of China; 3) Department of Neurology, Emory University School of Medicine, Atlanta, GA, USA.

Expansion of GGGGCC hexanucleotide repeat in the noncoding region of chromosome 9 open reading frame 72 (C9ORF72) gene has been found among many families with frontotemporal dementia (FTD) and/or amyotrophic lateral sclerosis (ALS) linked to chromosome 9 (c9FTD/ALS), however, the molecular pathogenesis underlying this expanded repeat remains unknown. Here we developed both mammalian cells and Drosophila model of this expanded hexanucleotide repeat, and showed that the expression of the expanded GGGGCC repeat RNA (30) is sufficient to cause neurodegeneration, suggesting that rGGGGCC-repeat-binding proteins (RBPs) may be sequestered from their normal function by rGGGGCC binding. We further identified Pur α, β and γ as RBPs of rGGGGCC repeat, and show that Pur α and rGGGGCC repeats interact in vitro and in vivo with a sequence-specific fashion that is conserved between mammals and Drosophila. Overexpression of Pur α in mammalian cells and Drosophila could suppress rGGGGCC-mediated neurodegeneration. These results suggest that the expanded rGGGGCC repeat is sufficient to cause the neurodegeneration, and Pur α could potentially play a role in the pathogenesis of ALS/FTD.

2478W

An Integrated and Systematic Approach to Autism Pathway Analysis. *Z. Lin[1], SJ. Sanders[2], JK. Lowe[4], DH. Geschwind[4,5], M. State[2], H. Zhao[3].* 1) Computational biology and bioinformatics, Yale University, New Haven, CT; 2) Program on Neurogenetics, Child Study Center, Department of Psychiatry, Yale University School of Medicine, 230 South Frontage Road, New Haven, CT; 3) Department of Epidemiology and Public Health, Department of Genetics, Yale University, New Haven, CT; 4) Program in Neurogenetics and Neurobehavioral Genetics, Department of Neurology and Semel Institute, David Geffen School of Medicine, University of California, Los Angeles, CA; 5) Department of Human Genetics, University of California Los Angeles, Los Angeles, CA.

Introduction Autism spectrum disorder (ASD) is a common, highly heritable neurodevelopment syndrome characterized by fundamental impairments in social communication. Driven by the ability to examine genome-wide expression levels and DNA variations, researchers have used various approaches to identifying the genes contributing ASD risk, including: differential gene expression in autistic and normal brains; common single nucleotide polymorphisms (SNPs) disease association; and recurrent rare de novo single nucleotide variants (SNVs). In this study, we hypothesize that genes carrying risk for ASD are functionally related and thus annotated to a smaller number of pathways, which is present in data across all sample methodologies.**Methods Summary** We have examined exome-sequencing data from 547 families (200 quartets and 347 trios) from the Simons Simplex Collection, microarray genotyping derived from 1273 families (5445 subjects) from Autism Genetic Resource Exchange, and gene expression microarray of 36 post-mortem brain tissue samples (19 autistic cases and 17 controls) from Autism Tissue Project and the Harvard brain bank. We have performed hypothesis testing on annotated pathways instead of individual genes to increase detection power as a result of reduced hypothesis testing. Permutation test based on the number of hits by rare de novo SNVs was performed on each pathway. For the microarray genotyping data, Transmission Disequilibrium Test was performed to acquire p-value for each SNP. The p-value of a gene was based on the SNP with lowest p-value within 50kb. Permutation test based on the mean of p-values of genes in the pathway was performed. Gene Set Enrichment Analysis (GSEA) on the brain expression microarray data was performed to acquire p-value for each pathway. P-values from the three independent datasets were integrated by Fisher's method and a summarized p-value was acquired for each pathway. **Results and Discussion** Our results suggest that signals from the three independent data sources show evidence of significant overlap. After correction for multiple testing, six pathways are identified as showing enrichment. Of these `FocalAdhesion' showed the most significant association (p=0.0067). Genes from the other five pathways showed significant overlap with genes annotated to the focal adhesion pathway and removing these genes left no remaining association. We conclude that disruption of focal adhesion is a primary event in the causation of ASD.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2479T

RNA sequencing in iPSC-derived neurons identifies gene expression changes associated with 22q11.2 microdeletion syndrome. *Y. Tian[1,2], S. Pasca[3], J. Ou[1], A. Krawisz[3], J. BernsteinJon[4], J. Rapoport[5], J. Hallmayer[6], S. Horvath[2,7], M. Pellegrini[2], R. Dolmetsch[3], D. Geschwind[1,2,7,8], ACE Genetics Consortium.* 1) Neurogenetics Program, Department of Neurology, David Geffen School of Medicine, UCLA, Los Angeles, CA; 2) Interdepartmental Ph.D. Program in Bioinformatics, UCLA, Los Angeles, CA; 3) Deptartment of Neurobiology, Stanford University, School of Medicine, Stanford University, Stanford, CA; 4) Deptartment of Pediatrics, Stanford University, Stanford, CA; 5) Child Psychiatry Branch, National Institute. of Mental Health, Bethesda, MD; 6) Department of Psychiatry and Behavioral Science, Stanford University, Stanford, CA; 7) Department of Human Genetics, David Geffen School of Medicine, UCLA, Los Angeles, CA; 8) Center for Autism Research and Treatment and Semel Institute, David Geffen School of Medicine, UCLA, Los Angeles, CA.

Genetically defined neurodevelopmental disorders provide a powerful window into the pathogenesis of mental disease. The 22q11.2 microdeletion syndrome is caused by the deletion of a 3 million base-pair region of chromosome 22, and is associated with increased risk of autism spectrum disorders (ASD) and schizophrenia. In this study, we performed genome-wide measurement of RNA expression via deep RNAseq on cortical neuronal precursor cells (NPC) and neurons from induced pluripotent stem cells (iPSC) derived from 22q11.2 deletion cases and normal controls to identify dysregulated molecular pathways at key developmental epochs. We found 453 differentially expressed (DE) genes in neurons and 216 genes in NPCs, which by hierarchical clustering clearly distinguished 22q11.2 deletion cases from controls. Remarkably, 21 DE genes in 22q11.2 deletion neurons and 14 DE genes in NPCs have been previously implicated in idiopathic ASD. Gene Ontology (GO) analysis revealed enrichment of alternative splicing and calcium ion binding. Additionally, by comparing the pattern of gene expression changes that occur during neural differentiation in controls vs. 22q11.2 deletion carriers, we can see an apparent defect in neuronal differentiation in the deletion carriers. We applied Weighted Gene Co-expression Network Analysis (WGCNA) to provide a higher order view of transcriptional dysregulation. We identified 4 gene co-expression modules associated with 22q11.2 deletion, including one having GRIN2B, an autism candidate gene as a hub, which was downregulated in 22q11.2 deletion neurons. These data characterize the network of genes dysregulated in iPSC-derived cells from 22q11. 2 deletion patients during neuronal differentiation and maturation. These patterns provide insight into the pathogenesis of 22q11.2 deletion syndrome and a first step in designing molecularly targeted therapy.

2480F

The effects of inherited deletions on neurocognition in Canadian children with 22q11.2 deletion syndrome. *E. Chow[1,2,3], A. Ho[1], D. Young[2], S. Langlois[4,5], C. Cytrabaum[6], A. Rideout[3], S. Dyack[3,7].* 1) Clinical Gen Service, Ctr Addiction & Mental Hlth, Toronto, ON, Canada; 2) University of Toronto, Toronto, ON, Canada; 3) IWK Health Centre, Halifax, NS, Canada; 4) BC Children's Hospital, Vancouver, BC, Canada; 5) University of British Columbia, Vancouver, BC, Canada; 6) Hospital for Sick Children, Toronto, ON, Canada; 7) Dalhousie University, Halifax, NS, Canada.

Background: Because there are no exclusively Canadian data on the neurocognition in children with 22q11DS in the literature, we carried out the current study. Samples: We assessed a total of 56 subjects affected by 22q11DS (25 M, 31 F) and between the ages of 8 to 12 years (mean=9.8 SD =1.9 years) and a total of 26 siblings (14 M, 12 F) also in the same age range (mean= 10.5 SD=1.8 years) from Ontario (n=20 affected and 9 sibling subjects), British Columbia (n=18 affected and 7 sibling subjects), and the Maritimes provinces (n=18 affected and 10 sibling subjects). The 22q11.2 deletion in 11 of the affected children was inherited (all but one maternally), although only 45% were raised by their biological parents. Methods: All subjects were assessed with tests of intellect (WISC-III or IV), academic achievement (WRAT-3), attention/working memory (Connor's CPT, digit span), memory (Rey auditory verbal learning test, CMS-III), motor skills (Purdue pegboard, Trials A), visual spatial skills (Line orientation), and social cognition (Theory of Mind). Test performances were compared between the 22q11DS and siblings subjects, and between 22q11DS subjects with and without an inherited deletion. Results: 22q11DS children performed significantly worse than their siblings on all neurocognitive tests: WISC-III or IV FSIQ (67.5, SD 12.2 vs. 96.4, SD 16.6, $p<0.001$), academics (WRAT3 reading, spelling and arithmetic $p<0.001$), attention (Connor's CPT $p=0.016$, Digit Span $p<0.001$), verbal memory (RAVLT total score $p<0.001$, delayed recall $p= 0.003$; CMS immediate $p=0.016$, delayed recall $p=0.011$), visual spatial memory (CMS immediate $p=0.002$ and delayed recall $p<0.001$), motor skills ($p<0.001$), visual spatial perception ($p=0.012$), and social cognition ($p<0.001$). Children with an inherited deletion had lower FSIQ (71.9 vs 65.8, $p=0.027$), and worse performance in digit span ($p=0.099$), symbol search ($p=0.030$), WRAT-3 spelling ($p=0.035$), and motor skills (non-dominant hand $p=0.041$) than children with a de novo deletion. Although sample sizes were small, the data suggests that children with inherited deletions raised by their affected parent had lower IQ and verbal learning skills than affected children raised by adoptive parents. Conclusions: 22q11DS is associated with a wide range of neurocognitive impairments in childhood and they may be worse in children with an inherited deletion, especially when raised by their affected parent.

2481W

Brain-Derived Neurotrophic Factor (BDNF) Val66Met and Adulthood Chronic Stress Interact to Affect Depressive Symptoms. *R. Jiang, B.H. Brummett, M.A. Babyak, I.C. Siegler, R.B. Williams.* Duke University Medical Center, Durham, NC.

Objective: To examine whether the association between *BDNF* Val66Met and depressive symptom severity is modified by adulthood chronic stress in two independent samples.**Background:** *BDNF* Val66Met polymorphism has been found associated with depression, but the results are mixed, which may be partially attributed to failure to consider environmental factors, such as chronic stress. *BDNF* Val66Met by chronic stress interaction has been studied using childhood stress as a moderator, but has not been widely studied using chronic stress in adulthood. **Methods:** Two independent samples were used: Duke-CG (238 Caucasians) and MESA (5524 Caucasians, African Americans and Hispanics). In the Duke-CG sample chronic stress was operationalized as having the primary caregiving responsibility for a spouse or relative with diagnosed Alzheimer's disease or other major dementia; chronic stress in MESA was defined using a chronic burden score constructed from self-reported problems of health (self and someone close), job, finance and relationships. In both samples, the Center for Epidemiologic Studies of Depression (CES-D) scale was the measure of depressive symptoms. The *BDNF* Val66Met by adulthood chronic stress interaction predicting CES-D was examined using linear regression, adjusted for race, gender, age, BMI, and anti-depressant use in both samples, and data collection site in the MESA sample. **Results:** The main effect of *BDNF*Val66Met genotype on CES-D scores was non-significant ($ps > 0.607$) but the adulthood chronic stress indicator was significant ($ps < 0.001$) in both samples. The *BDNF* Val66Met genotype by adulthood chronic stress interaction was also significant ($ps < 0.039$) in both samples. The impact of chronic stress in adulthood on CES-D scores was significantly larger in individuals with the Val/Val genotype compared to Met carriers. **Conclusion:** We found in two independent samples that depressive symptom levels increased significantly more as a function of adulthood chronic stress among Val/Val carriers than Met carriers that was in a different direction from the moderation observed for childhood stress. If the association of the Val/Val genotype with increased depressive symptoms in adults exposed to chronic stress is confirmed in future studies, persons with the Val/Val genotype and chronic stress exposure could be targeted for interventions designed to reduce risk of depression.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2482T

miRNA Expression in the Prefrontal Cortex of Suicide Completers. *J.P. Lopez[1, 2], R. Lim[3], B. Labonte[2], C. Cruceanu[1,2], J.P. Yang[2], V. Yerko[2], C. Ernst[1,2], N. Mechawar[2], P. Pavlidis[3], G. Turecki[1,2].* 1) Human Genetics, McGill University, Montreal, Quebec, Canada; 2) McGill Group for Suicide Studies, Douglas Mental Health University Institute, McGill University, Montreal, Quebec, Canada; 3) Department of Psychiatry, University of British Columbia, Vancouver, B.C, Canada.

Suicide is a major public health problem. It was estimated in 2007 that suicide accounts for 1.5% of the total deaths in Canada. Over the last decades, a large body of evidence has shown that individuals who commit suicide have a predisposition that is mediated by neurobiological factors. MicroRNAs (miRNAs) are small non-coding RNAs that regulate gene expression by means of RNA degradation or translational repression. A growing body of evidence supports the role of miRNAs in neuropsychiatric disorders such as schizophrenia and bipolar disorder. Based on this evidence we believe that miRNAs have an important role in the neurobiology of suicidal behaviours. As such, we sought to identify miRNA differences in the brain of suicide completers as compared to controls. We profiled the expression of miRNAs in the prefrontal cortex of suicide completers and controls using miRNA microarrays and quantitative real-time PCR. We identified target genes of the differentially expressed miRNAs using bioinformatics analysis. Finally, we performed functional experiments using miRNA mimics and target protectors on HEK-293T cell lines to confirm the interaction between miRNAs and their targets. Our results revealed an interaction between miR-1202 and the Metabotropic Glutamate Receptor 4 (GRM4) gene in the prefrontal cortex of suicide completers. These results suggest that the relationship between miR-1202 and GRM4 in postmortem brain tissue may be associated with the pathophysiology of suicidal behaviour. Ultimately, our results provide new evidence of the role of miRNAs in neuropsychiatric disorders and a better understanding of the etiology of suicide.

2483F

Distinct gene expression and DNA methylation perturbations in PTSD with severe child abuse versus no PTSD with no child abuse. *D. Mehta[1], T. Klengel[1], K.N. Conneely[4], A.K. Smith[3], A. Altmann[1], M. Rex-Haffner[1], A. Loeschner[1], M. Gonik[1], K.B. Mercer[2], B. Bradley[3,5], B. Müller-Myhsok[1], K.J. Ressler[2,3,6], E.B. Binder[1,6].* 1) Max Planck Institute for Psychiatry, Max Planck Institute for Psychiatry, Munich, Germany; 2) Howard Hughes Medical Institute, Chevy Chase, Maryland, USA; 3) Department of Psychiatry and Behavioral Sciences, Emory University School of Medicine, Atlanta, USA; 4) Department of Human Genetics, Emory University School of Medicine, Atlanta, GA, USA; 5) VA Medical Center, Atlanta, GA, USA; 6) Yerkes National Primate Research Center, Atlanta, GA, USA.

The aim of the current study was to interrogate the influences of the environment on gene expression profiles by characterizing biological differences in PTSD after severe child abuse versus PTSD after adult trauma. A total of 396 trauma-exposed individuals were included in the study. Whole blood gene expression and DNA methylation was measured on Illumina Human-HT12v3 and HumanMethylation 450k arrays. Analysis was performed in R. Expression levels of 741 transcripts were significantly associated with current PTSD severity. However, only 2% of these transcripts were associated with PTSD in both groups after accounting for adult trauma severity. Expression differences were also reflected in DNA methylation differences. DNA methylation differences were observed in 69.3% of the transcripts significantly regulated by PTSD with child abuse but for only 33.6% transcripts selective for PTSD without child abuse. Functional annotations revealed distinct biological pathways enriched among the two distinct expression profiles. The observations that non-overlapping transcriptional profiles and biological pathways seem to be affected in the two PTSD groups and that changes in DNA methylation appear to have a much greater impact in the early trauma group may reflect differences in the pathophysiology of PTSD after exposure to early trauma. Stratification by environmental variables such as timing, type and severity of trauma might therefore be essential for the identification of robust biomarkers for PTSD. T.

2484W

Frequency of polymorphisms of the alcohol dehydrogenase enzyme (ADH1B and ADH1C) and aldehyde dehydrogenase enzyme (ALDH) in the Brazilian population. G*.P.M. Antonialli, B. Coprerski, R.M. Minillo, M.P. Migliavacca, D. Brunoni, A.B.A. Perez.* Clinical genetics, Federal University of Sao Paulo, Sao Paulo, Brazil.

Introduction: Alcohol dependence is a major social and economic problem worldwide. The influence of the polymorphisms in the code regions of the alcohol dehydrogenase (ADH) and aldehyde dehydrogenase (ALDH) enzyme is well characterized as a risk factor for alcohol dependence in some populations. There are different classes of ADH, the classes I (ADH1A, ADH1B e ADH1C) and II are responsible for the great part of the metabolism of the ethanol ingested. The ALDH is presented in the mitochondrial form (ALDH2), which has two isoforms, and in the cytosolic form (ALDH1). Until this moment few studies were performed to try to establish the frequency of these polymorphisms in the Brazilian population. Purpose: To determine the frequency of ADH1B (ADH1B*2 - Arg48His and ADH1B*3 - Arg370Cys), ADH1C (ADH1C*2 - Ile350Val) and ALDH2 (ALDH2*2) polymorphisms in a healthy Brazilian population and also compare with the frequency of other populations. Methods: DNA was isolated from peripheral blood of 202 healthy patients, of which 101 were male and 101 were female divided by ethnicity. DNA was amplified by PCR reaction, digested by specific restriction enzyme and analyzed by gel electrophoresis. Results: The following genotype frequencies were found for each polymorphism: ADH1B*2: 87,13% ADH1B*1/1 and 12,87% ADH1B*1/2. ADH1B*3: 89,69% ADH1B*1/1 and 10,31% ADH1B*1/3. ADH1C*2: 40,10% ADH1C*1/1; 49,01% ADH1C*1/2 and 10,89% ADH1C*2/2. ALDH2*2: 98,02% ALDH2*1/1 and 1,98% ALDH2*1/2. The genotype distribution of the ADH1B, ADH1C and ALDH2 enzyme polymorphism of the population studied were similar to those described in some white European and American population, but quite different from the frequency described in Asians. The frequencies found for this ADH polymorphism were similar to those described in previous Brazilian studies. The only allele that showed a statistically significant association by ethnicity was the ALDH2*2, which was more frequent in the descendants of Eastern ($p<0,0001$). There were no significant differences in the genotype distribution between both sexes. Conclusion: These results will be fundamental to future Brazilian epidemiological studies on the influence of these polymorphisms in several diseases, including alcoholism and Fetal Alcohol Syndrome.

2485T

Serotonin transporter (SLC6A4) and receptor (5-HT2A) gene polymorphisms in healthy Brazilian individuals with diverse ethnic backgrounds. *B. Coprerski[1,2], M. P. Micliavacca[1], A. B. A. Perez[1], A. P. P. Costa[2], D. Brunoni[1,2].* 1) Clinical Genetics, Federal University of Sao Paulo, Sao Paulo, Brazil; 2) Biology and health department, Makenzie Presbiterian University, Brazil.

Introduction: Autism is a heterogeneous syndrome defined by the commitment in three areas (social interaction, language and communication) and the presence of repetitive and stereotyped patterns of behavior and interests. Intense genetic-clinical and systematic research sets up to 30% of causes. The heritability of idiopathic cases showed 90% and are explained by epistatic multifactorial mechanism. Hyperserotonemia in 30% of patients and the improvement of symptoms with drugs that inhibit serotonin reuptake put this system as possibly involved. In this line have been investigated genes of the transporter (SLC6A4, located at 17q11.2) and receptor 2A (5HT2A, located at 13q12-q21) of serotonin. Objective: To determine allele frequencies and the hypothesis of preferential association of four polymorphisms of serotonin in patients with PDD and controls. Material and Methods: Three polymorphisms of the SLC6A4 gene were studied: The polymorphism deletion/insertion of 44 bp in the promoter region, known as 5-HTTLPR, a polymorphism of 17 bp variable number of tandem repeats (VNTR) in intron 2 and a single base polymorphism (SNP) in the 3 `untranslated (UTR) of 5-HTT gene. For the 5HT2A gene T102C SNP was studied. DNA was extracted and genotyped for each polymorphism. We analyzed 116 patients with PDD for the four polymorphisms and 239 controls. The cases were analyzed by the instruments ASQ and ABC as to the higher susceptibility to repetitive behaviors and stereotyped. Results and Conclusion: There was no statistically significant evidence for the distribution of genotypes in both cases and controls. Regarding the number of stereotypes, both in the ASQ and the ABC there was no significant difference in genotype distribution. The sample is in Hardy-Weinberg equilibrium. Further studies with larger sample and better definition of some endophenotype will be required to demonstrate the possible implication of these polymorphisms in the etiology of autism.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2486F

Regulatory effects of genes associated with Schizophrenia. *N. Eckart[1], R. Wang[2], J. McGrath[2], P. Wolyniec[2], M. Zeledòn[1], M. Szymanski[1], A. Pulver[2], D. Valle[1], D. Avramopoulos[1,2].*1) Institute of Genetic Medicine, Johns Hopkins University, Baltimore, MD; 2) Department of Psychiatry, Johns Hopkins University, Baltimore, MD.

Schizophrenia (SZ) is a common complex psychiatric disorder with high heritability. GWAS experiments have begun to identify associated SNPs; however, the true causative variants and their pathophysiological mechanisms are largely unknown. As in other disorders, the majority of the SZ associated SNPs are in non-coding regions and likely reflect regulatory variation. We hypothesized that many reported disease-associated variants disrupt the regulation of gene expression and explored this hypothesis in 187 postmortem brain samples from the temporal lobe of donors without brain pathology. We focused our study on a total of 14 SNP markers, 2 associated with SZ, 6 with both SZ and bipolar disorder, and 6 SNPs in *NRG3*, a gene we previously reported to be associated with certain phenotypic features of SZ. Using transcript specific primers and RT-qPCR, we measured the levels of known alternative transcripts of the gene closest to each of the 14 SNPs in RNA extracted from our samples. We used a generalized linear model including age, sex, and plate to find correlations between the genotypes and the total and relative expression of the gene's alternative mRNA transcripts. We found significant correlations for three SNPs (rs1006737 in *CACNA1C*; rs10748842 in *NRG3*; rs7899151 in *NRG3*). For*CACNA1C*, we found the genotypes of rs1006737, previously reported to be associated with both SZ and BP, were correlated with expression of all *CACNA1C* transcripts. For *NRG3*, we found the genotype of rs10748842 was associated with the relative abundance of *NRG3* alternative transcripts. Previously, our group showed that rs10748842 genotype correlated strongly with the “delusion” phenotypic factor in SZ. An effect of rs10748842 on relative the same alternative transcript of *NRG3* has been independently reported by others in the prefrontal cortex. Finally rs7899151 which we studied because of its correlation with the “disorganization” factor of SZ showed a correlation between genotype and a different *NRG3* alternative transcript. We conclude that non-coding variants associated with disease often mark regulatory haplotypes that can also involve alternative splicing. Further studies to examine the functional consequences of variation within these haplotypes are important to better understand the mechanisms of pathogenesis.

2487W

Genome-wide association study of 32,143 individuals reveals several novel associations in schizophrenia. *C. O'Dushlaine[1], S. Ripke[1,2], J. Moran[1], K. Chambert[1], P. Sklar[3], S. Purcell[1,2,3], C. Hultman[4], S. McCarroll[1], P. Sullivan[5], Swedish Schizophrenia Consortium.* 1) Stanley Center for Psychiatric Research, Broad Institute of Harvard and MIT, Cambridge, MA, 02142; 2) Analytical and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA, 02114; 3) Division of Psychiatric Genomics, Department of Psychiatry, Mount Sinai School of Medicine, NY, NY 10029; 4) Department of Medical Epidemiology, Karolinska Institutet, Stockholm, Sweden; 5) Departments of Genetics, Psychiatry, and Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, NC, 27599.

BackgroundSchizophrenia is a genetically complex and clinically heterogeneous disorder. Several large-scale genome-wide association meta-analyses have discovered 14 independent loci with genome-wide significant association. Here, we present results from the latest and largest such study, representing an increase in sample size of 59% over previous studies. Our primary study consisted of 5,001 cases and 6,243 controls ascertained from Sweden. The genetic data were imputed into HapMap3 and 1000 Genomes data and analyzed using logistic regression with MDS covariates. **Methods/ Results** We compared the results to already published data from the collection from the Psychiatric GWAS Consortium (excluding the Swedish cohorts, 8,832 cases and 12,067 controls of European ancestry). First, we performed polygene testing (Purcell, 2009) to assess variants of lesser effect. Second, we first compared directionality in SNPs with high-significance ($p<10-4$) via a sign test. After showing excellent concordance in both directions (sign test), we meta-analyzed these two result-sets in a sample of 13,833 cases and 18,310 controls. We found increased evidence of association for many previously implicated loci in addition to evidence for novel loci: in total we find 14 LD-independent SNPs with genome-wide significance. An additional 31 LD-independent SNPs with $p<10-6$ are now under investigation in a replication dataset. Finally, we conducted gene-based and pathway analyses to test for aggregate-level associations, possibly in the presence of allelic and/or locus heterogeneity. Our gene-based tests recapitulate many of the top GWAS hits but also highlight additional loci of interest. Pathway analysis using KEGG, Gene Ontology and TargetScan datasets, revealed several highly significant associations, most notably relating to postsynaptic density in addition to the previously-reported Cell Adhesion Molecule Pathway. Intriguingly, of 87 microRNA families tested (range of 2–200 genes), 26 had a pathway p-value <10-4, potentially implicating additional microRNAs in schizophrenia. We confirm enrichment of association among targets of mir137 ($P=1.8e-9$) and find that 45 of 299 targets with at least one SNP have a gene p-value ≤ 0.05. **Discussion** Our results confirm previous associations in schizophrenia and yield new insights at both the gene and pathway level, together representing major steps forward in understanding the molecular pathophysiology of schizophrenia.

2488T

Genome Wide Association Study Reveals Genes Influencing the Rate of Cognitive Decline Among Persons With Alzheimer's Disease are Distinct From Those Associated With Disease Risk. *R. Sherva[1], Y. Tripodis[1], D. Bennett[2], L.A. Chibnik[3], P.K. Crane[4], P. de Jager[3], L.B. Farrer[1], J. Shulman[3], R.C. Green[3].* 1) Boston University, Boston, MA; 2) Rush University Medical Center; 3) Harvard Medical School; 4) University of Washington.

Substantial inter-individual variability exists in the longitudinal disease trajectories of Alzheimer's disease (AD) patients, some of which may be due to genetic factors. In a discovery sample of 303 AD cases recruited in the AD Neuroimaging Initiative (ADNI), we performed a genome wide association study on the outcome of rate of change in the Alzheimer's Disease Assessment Scale-cognitive subscale (ADAS-cog) measured longitudinally at six month intervals for up to four years after disease onset. Genetic association was assessed with interaction terms between SNPs and time with AD using linear regression models solved by generalized estimating equations to account for correlated observations within individuals. Several SNPs were significantly associated with rate of progression after correcting for multiple tests and genomic inflation, including SNPs in α-catenin 2 (CTNNA2) (rs6738962, $P=1.2\times10-8$), which regulates the stability of synaptic contacts in the hippocampus, and spondin 1 (SPON1) (rs11607537, $P= 7.2\times10–11$), which binds to the amyloid precursor protein and inhibits its cleavage by beta secretase. Minor alleles at both of these SNPs were associated with slower decline in ADAS-cog. A SNP 4KB upstream of rs11607537 was associated with more rapid decline in an independent sample of 323 AD cases from the Religious Orders Study and Rush Memory and Aging Project (rs11606345, $P=0.002$) using a different cognitive phenotype and analytic model. The SNP was also associated with a slower rate of amyloid β-40 (but not Aβ-42) accumulation in cerebrospinal fluid ($P = 0.001$), and slower rates of decline measured by the Rey Auditory Verbal Learning Test ($P = 0.008$) and the Mini-Mental State Examination ($P = 0.003$). None of the previously identified variants that predict the risk of developing AD were associated with rate of progression after correcting for multiple tests, indicating genes influencing the rate of cognitive decline after disease onset are different from those which influence mechanisms leading to disease.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2489F

Genome-wide scan for Copy Number Variation association with Age at Onset of Alzheimer's Disease. *K. Szigeti[1], D. Lal[1], Y. Li[2], R. Doody[2], K. Wilhelmsen[3], L. Yan[4], S. Liu[5], C. Ma[4], Texas Alzheimer Research and Care Consortium.* 1) Dept Neurology, University at Buffalo, Bufalo, NY; 2) Depts of Molecular and Human Genetics, Neurology Baylor College of Medicine, Houston, TX; 3) Dept Neurology, University of North Carolina, Chapel Hill, NC; 4) Department of Bioinformatics, University at Buffalo, Buffalo, NY; 5) Roswell Park Cancer Institute, Buffalo, NY.

Alzheimer's disease is a progressive neurodegenerative disease with high prevalence imposing a substantial public health problem. The heritability of AD is estimated at 70 % forcasting potential of using genetic biomarkers for risk stratification in the future. Several large scale GWAS studies using high frequency variants identified 10 loci accountable for only a fraction of the estimated heritability. To find the missing heritability systematic assessment of the various mutational mechanisms will need to be performed. This copy number variation genome-wide association study with age at onset of AD identified 14 CNVs corresponding to 11 chromosomal regions that may contribute to the heritability of AAO of AD. Two CNV events are intragenic causing a deletion in ACTG2 and CPNE4. In addition, to further study the mutational load at the 10 known susceptibility loci, CNVs overlapping with these loci are also catalogued. We identified rare small events overlapping CR1 and BIN1 in AD and NC with opposite CNV dosage. The CR1 events are consistent with previous reports. Larger scale studies with deeper genotyping specifically addressing copy number variation are needed to evaluate the significance of these findings.

2490W

Utilizing eQTL networks to identify genes relevant to autism in the 16p11.2 region. *Y. Cheng, W. Lauren.* Department of Psychiatry and Institute for Human Genetics, University of California, San Francisco, San Francisco, CA.

Copy number variants (CNVs) in the 16p11.2 region were identified by Weiss et al. 2008 as associated with autism in three independent datasets, and repeated shortly thereafter by Kumar et al. 2008 using a different approach with a subset of the same discovery sample. To date, none of the 25 annotated genes in the 16p11.2 region have proven responsible for the autism phenotype. Here, we utilized expression quantitative trait loci (eQTL) networks to identify autism genes in the 16p11.2 region. We first used control lymphoblast cell line (LCL) genome-wide expression (21 16p11.2 genes with expression data available) and genetic data in SCAN (Gamazon et al. 2009) to map 1) the single nucleotide polymorphisms (SNPs) associated with expression level of each of the 21 genes in trans, and 2) the SNPs located in the 16p11.2 region that were associated with the expression of other genes in trans. The resulting SNPs were organized into eQTL regions (with regions considered the same across networks if they had >100kb overlap) and subsequently tested for association with three SNP datasets and seven CNV datasets in idiopathic autism (16p11.2 CNV carriers excluded). We discovered significant overlap between eQTL networks of SPN, CORO1A, PPP4C, CDIPT, C16orf54, TBX6, KCTD13, ALDOA, MAZ and autism-specific CNVs in at least two CNV datasets. In addition, significant SNP sets were identified for the eQTL networks of SPN, CDIPT, MVP, YPEL3, TBX5, HIRIP3, KIF22, and MAZ in at least one SNP dataset. 16p11.2 genes shared upstream eQTL regions. For example, putative eQTL regions that were associated with the expression of at least five 16p11.2 genes include the following genes: CSDM1, CBLN2, FAM135B, NRG3, A2BP1, LINC004659, MACROD2, PTPRT, ROBO2, GRIN2B, and several multi-gene regions: USH2A / ESRRG, C3orf19 / C3orf20, RAB31 / TXNDC2 / VAPA, 19q13.32 (eleven genes). A similar analysis of the non-overlapping 16p11.2 eQTLs (that could be more confidently assigned to a single gene) suggested that only ALDOA showed significant association with autism in two CNV datasets, and thus might possess an independent regulatory network. In summary, our results indicated that a number of 16p11.2 genes were co-regulated by common upstream genes critical to neurodevelopment and their expression network was associated with common and rare variants in autism. Our data also suggest that multiple genes in the region may function together or in similar processes to contribute to autism.

2491T

A genomic instability model of autism: global copy number burden associated with autism. *N. KATIYAR[1], S. Girirajan[1,2], A. Srikanth[1], K.H. Yeoh[1], S.J. Khoo[1], F. Tassone[3], R. Hansen[3], I. Hertz-Picciotto[3], E.E. Eichler[2], I.N. Pessah[3], S.B. Selleck[1], M.D. Ritchie[1].* 1) Bioinformatics and Genomics, The Pennsylvania State University, STATE COLLEGE, PA; 2) Department of Genome Sciences, University of Washington Medical School, Seattle, WA; 3) The Medical Investigations of Neurodevelopmental Disorders Institute, University of California at Davis Medical Center, Sacramento, CA 95817, USA.

Recent work has demonstrated an elevated frequency of large, rare copy number variants (CNVs) in a number of neurodevelopmental disorders, including autism spectrum disorder (ASD), schizophrenia, and developmental delay. However, the global load of deletions or duplications per se, and the relationship between total copy number burden and the clinical manifestation of autism/ASD has not been well documented. We examined CNVs in 554 children aged 2–5 years who either had full syndrome autism (AU) or were typically developing (TD) controls from the population-based CHildhood Autism Risks from Genetics and Environment (CHARGE) study. All cases were confirmed clinically on both the Autism Diagnostic Inventory-Revised and the Autism Diagnostic Observation Schedule. Similarly, all controls were confirmed to have typical cognitive and adaptive function and screened negative for ASD on the Social Communication Questionnaire. We interrogated 107 genomic regions flanked by segmental duplications (genomic hotspots) for events >50kbp and the entire genomic backbone for variants >300 kbp using a custom targeted DNA microarray. This analysis was complemented by a separate study of five highly dynamic, hotspot regions previously associated with autism or developmental delay syndromes using a finely-tiled array platform providing a high resolution for CNV detection (>1 kbp) in 142 gender and ethnicity-matched AU and TD children. In both studies, a significant increase in total copy number burden (measured as total base pairs affected by burden) was associated with AU/ASD, represented principally as large duplication events. Significantly elevated levels of burden remained even after removal of rare, known and likely, pathogenic events. These findings were replicated using data from previously published work on CNVs measured for a number of neurodevelopmental disorders, including ASD (Simons Simplex Cohort, SSC). Our data suggest a mechanistic bias toward duplications in the etiology of ASD and emphasizes the importance of genomic instability in neurodevelopmental disorders. While we were not able to identify the parent of origin or the extent to which these duplications were amplified de novo, this study does demonstrate how higher resolution can provide more refined results. This study provides an exciting new approach to pursue in the exploration of the complexity of autism and other neurodevelopmental disorders.

2492F

The impact of glutamate and gamma-aminobutyric acid receptor signaling in non-syndromic autism. *A. Mohamed-Hadley[1], D. Hadley[1], Z. Wu[1], C. Kao[1], A. Kini[1], J. Glessner[1], R. Pellegrino[1], C. Kim[1], K. Thomas[1], H. Hakonarson[1,2], AGP Consortium.* 1) The Center for Applied Genomics, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA; 2) Division of Genetics and Department of Pediatrics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania, USA.

Autism is a complex heterogeneous disorder that is strongly heritable. While several variants have been associated with autism, they only explain a small fraction of the disease heritability. Here, we compare genome-wide copy number variation in over 6.7K patients diagnosed with autism spectrum disorders (ADSs) to over 12.5K neurologically normal controls. The results demonstrate enriched CNV burden that most significantly impacts specific gamma-aminobutyric acid (GABA) receptors and receptors within the metabotropic glutamate receptors (mGluR) pathway impacting ~6% of patients with ASDs. Interestingly, postsynaptic metabotropic glutamate receptors and GABA-B receptors, which are known to functionally interact with each other to regulate glutamate mediated excitatory transmission in the brain, have recently been implicated in autism and genetic defects in multiple GABA receptor genes have also been shown to be associated with the common form of autism. More recently, mGluR5 functional activity was shown to be central in defining the axis of synaptopathophysiology in rare syndromic autism. Although the importance of mGluRs has been repeatedly demonstrated in ADHD, a highly co-incident symptom within the autism spectrum, our study is the first to demonstrate defective mGluR network in the more common non-syndromic autism. We provide evidence of a role for both GABA and mGluR receptor activity in ASDs, with both pathways serving as attractive targets for therapeutic rescue in patients with ASDs that harbor these specific mutations. Functional studies underway are anticipated to better define their individual roles and the interaction of their component genes within these important genetic networks in the pathogenesis of autism.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2493W

Chromosome 15q11.2 is associated with schizophrenia risk in a set of patients. *D. Rudd[1], M. Axelsen[2], S. Casey[2], S. Mugge[2], N. Andreasen[2], T. Wassink[2].* 1) Dept Genetics, Univ Iowa, Iowa City, IA; 2) Dept of Psychiatry, University of Iowa, Iowa City.

Schizophrenia is characterized by positive and negative symptoms that include hallucinations, delusions and disorganized thought as well as flat affect and decreased speech. Genomic copy number variation (CNV) has been found to underlie and increase risk for a large number of neuropsychiatric disorders like schizophrenia. This project focuses on identification and characterization of pathogenic CNVs in a sample set of 457 schizophrenia patients and 290 psychiatrically healthy controls. To date, 167 probands and 101 controls have been genotyped using the Affymetrix Genome-Wide Human SNP 6.0 Arrays and analyzed for CNVs with Genotyping Console and PennCNV. CNVs were compared between cases and controls and cross-referenced with the Database of Genomic Variants. The proximal arm of chromosome 15 contains canonical segmental duplications that predispose the region to CNV formation, but only recently have disease-associated CNVs been reported between the two most proximal breakpoints (BPs), BP1 and BP2. Here we report three case subjects that have a ~1Mb deletion of the BP1–BP2 region at chromosome 15q11.2. The BP1–BP2 deletion encompasses the four genes CYFIP1, NIPA1, NIPA2 and GCP5, is present in three probands, two of which were maternally inherited, but not in controls. All CNVs of interest were validated by qPCR or NimbleGen 385K whole-genome array CGH. The four genes within the CNV have previous association with Prader Willi syndrome, Angleman syndrome, autism and schizophrenia and are often inherited from a mildly or unaffected parent. CYFIP1 is of primary interest due to its direct interaction with the Fragile X Mental Retardation Protein (FMRP) and its reported role in regulation of neuronal translation during brain development. NIPA1, NIPA2 and GCP5 are also expressed in the brain. To see if the CNV unmasked any recessive variants on the remaining allele, that allele was sequenced and no such variants were identified. It is unlikely that the 15q11.2 CNV alone is pathogenic because it is such a common region of variation. Therefore each case was investigated further for other compelling CNV hits, based on a “multi-hit” CNV hypothesis where multiple CNVs as a group are pathogenic to schizophrenia. The two probands with maternal inheritance of the deletion were further analyzed for sporadic secondary CNVs. Examples of other compelling CNVs are a de novo duplication of four exons of ATXN2 de novo duplication of two exons CACNA1B and DLGAP2.

2494T

Genome-wide Copy Number Analysis on Schizophrenia in Han Chinese. *E. WONG[1], A.W. Butler[1,5], Q. Wang[4], S.S. Cherny[1,2,3], T. Li[4], P.C. Sham[1,2,3].* 1) Department of Psychiatry, the University of Hong Kong, Hong Kong SAR, China, Hong Kong; 2) Centre for Genomic Sciences, the University of Hong Kong, Hong Kong SAR, China, Hong Kong; 3) State Key Laboratory for Brain and Cognitive Sciences, the University of Hong Kong, Hong Kong SAR, China, Hong Kong; 4) The Mental Health Center & Psychiatric laboratory, West China Hospital, Sichuan University, Chengdu, Sichuan, China; 5) MRC SGDP Centre, Institute of Psychiatry, King's College London, UK.

Schizophrenia is a highly heritable, severe psychiatric disorder affecting approximately 1% of the world population. Apart from common variants observed in genome-wide association studies, a number of studies suggested that rare copy number variations (CNV) are rare high-penetrant mutations that increase risk of schizophrenia, and also of a range of other psychiatric disorders including autism and mental retardation. The majority of copy number variation studies on schizophrenia were performed on population with European ancestry. To examine if there are population specific loci to confer risk of schizophrenia, we performed a genome-wide CNV analysis on 381 Han Chinese patients with schizophrenia and 869 controls with no major disorder. We observed a higher global burden of duplication in schizophrenic patients than in controls (All CNVs: 2.3 fold, with empirical p-value = 1 x10-8; rare CNVs: 1.4 fold, with empirical p-value = 3.2 x10-5). As rare CNV events are more likely to be pathogenic, we performed rare CNV analyses in the genic and non-genic regions. There is a non-genic region on 10p14 showed moderate significance, with an empirical adjusted p-value 0.000136. Duplications with average length of 124kb were observed in 8 schizophrenia patients but not in controls. For those CNV regions previously reported associated to schizophrenia, we did not observe similar increased burden of microdeletions or microduplications. Instead of conferring risk to schizophrenia, some of those regions are highly variable in the Han Chinese, suggesting them to be common CNV regions in the population. One of these regions was on 16p11.2, in which 674 CNVs in total were observed in our cases and 873 CNVs in controls, consisting of two common CNV regions. Apart from examining the previously reported CNV regions, we also evaluated genes that are suggested to be associated to schizophrenia.

2495F

Rare copy number variants in schizophrenia and bipolar disorder in a Latino Population. *C. Xu[1], A.N. Blackburn[2], S. Gonzalez[1], E.C. Villa[3], M. Ramirez[1], J. Zavala[1], M. Rodriguez[1], C. Camarillo[1], J. Ordonez[1], R. Armas[2], S.A. Contreras[4], R.J. Leach[5], D. Flores[6], D. Jerez[7], A. Ontiveros[8], H. Nicolini[9], D. Lehman[2], M. Escamilla[1].* 1) Psychiatry and Neurology, Texas Tech University Health Sciences Center, El Paso, TX; 2) University of Texas Health Science Center San Antonio TX; 3) University of Texas at El Paso, TX; 4) Department of Psychiatry, University of Texas Health Science Center at San Antonio, San Antonio, TX; 5) Department of Cellular and Structural Genetics, University of Texas Health Science Center at San Antonio, San Antonio, TX; 6) Los Angeles Biomedical Research Center at Harbor, University of California Los Angeles Medical Center, Torrance, CA; 7) Centro Internacional de Trastornos Afectivos y de la Conducta Adictiva, Guatemala, Guatemala; 8) Instituto de Información e Investigación en Salud Mental AC, Monterrey, N.L., México; 9) Grupo de Estudios Médicos y Familiares Carracci, S.C., México, D.F., México.

Purpose: Rare copy number variants (CNVs) have been implicated in neuropsychiatric disorders. We aimed to establish whether burden of CNVs was increased in schizophrenia (SC) and bipolar disorder (BP), and to investigate whether identified CNVs were enriched for loci previously identified for SC and BP in the Latino population. Material and Methods: We undertook a genome-wide analysis of CNVs in 843 subjects including 208 BP, 195 SCZ and 440 unrelated matched controls of Mexican or Central American ancestry from the Southwest United States, Mexico, and Guatemala using Illumina HumanOmni1M-Quad array with 1.1 million markers. A number of quality assurance steps were conducted to filter problematic samples entirely and correct the Log R Ratio data. Two algorithms, SNP and Variation Suite 7 (SVS) and PennCNV, were used to identify CNVs. After quality control, we retained 799 subjects for further analysis. Results: Among high confidence CNV calls defined by both calling algorithms and overlapping for at least 95% of their length, we observed 18 rare CNV deletions between 10 kb to 131 kb in length that were present solely or more often in SC, BP cases than in controls. We confirmed 6 CNV regions that had previously been reported in BP and SC (1q41, 3q26.31, 3q27.1, 4q34.3, 7q31.1, and 7q36.3), and observed 10 novel deletions on 1q31.1, 2q, 5q, 10q26.13 and 11q23.2 in this Hispanic population. Interestingly, two CNV regions at 17q21.31 that were observed in 4 cases and 1 control (202 kb duplication) and 2 BP subjects (121 kb deletion), were located within 2Mbp of our previous linkage peak for SC in the same population (Escamilla et al., 2009). Using SVS algorithm alone, we also identified 9 deletions (ranges between 113.5–1900 kb in at least three subjects) in cases but not controls on 1p36.21, 2p11.2, 3p12.3, 7q11.21, 10q21.1, 15q11.2, 18p11.32, and 22q11.22-23. We also observed duplications at least 3 subjects, ranges between 147–464 kb in cases only on 1p21.1, 10q11.22, 12q14.2, 15q11.2–11.3, 16p11.2-11.1, and 22q11.23. We are performing independent methods to confirm the CNV calls in this population. Conclusion: To our knowledge, this would be the first study of the discovery of CNVs associated with BP and SC in the Latino population. Future replication in the same population and target genome sequencing on these novel CNV regions may finally provide a definitive genomic location for genes that may predispose to BP and SC.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2496W

Genome-wide CNV (copy number variation) analysis of narcolepsy in the Japanese population. *M. Yamasaki[1], T. Miyagawa[1], H. Toyoda[1], K. Seik Soon[1], A. Koike[2], T. Sasaki[3], Y. Honda[4], M. Honda[4,5], K. Tokunaga[1].* 1) Department of Human Genetics, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan; 2) Central Research Laboratory, Hitachi Ltd, Tokyo, Japan; 3) Department of Physical and Health Education, Graduate School of Education, The University of Tokyo, Tokyo, Japan; 4) Japan Somnology Center, Neuropsychiatric Research Institute, Tokyo, Japan; 5) Sleep Research Project, Department of Psychiatry and Behavioral Sciences, Tokyo Metropolitan Institute of Medical Science, Tokyo, Japan.

Human narcolepsy is a sleep disorder, characterized by sleepiness, cataplexy and rapid eye movement (REM) sleep abnormalities. Narcolepsy is caused by a reduction of neurons that produce orexin (hypocretin) neuropeptide. The disease usually begins during adolescence and affects 0.16–0.18% of the general population in Japan and 0.02–0.06% of the populations in the United States and Europe, with men and women equally affected. It is reported that narcolepsy was strongly associated with the human leukocyte antigen (HLA) class II region: *HLA-DRB1*1501-HLA-DQB1*0602* haplotype. The HLA association is thought to be a necessary condition but not sufficient. Recently, rare and large copy number variations (CNVs) have been reported to play a role in the etiology of several neuropsychiatric disorders. Since narcolepsy is considered as a neurological disorder, we aimed to assess the possible association with rare and large CNVs. DNA microarray data (Affymetrix 6.0) and two CNV detection software applications (Birdsuite and PennCNV-Affy) were utilized to detect CNVs in 462 Japanese narcoleptic patients and 562 healthy individuals. Overall, we found significantly enriched rare and large CNVs (frequency ≤1%, size ≥30kb) in the patients (case-control ratio of CNV count = 1.39, $P = 3.80\times10^{-4}$ by Birdsuite). The enrichment of rare and large CNVs in the patients was confirmed by another CNV detection software application (PennCNV-Affy). *PARK2* (parkinson protein 2, E3 ubiquitin protein ligase) region showed a significant association with narcolepsy. Four patients were assessed to carry duplications of the gene region, while no controls carried the duplication, that was further confirmed by quantitative PCR assay. Moreover, a pathway analysis of rare and large CNVs revealed enrichments in lipid catabolic process, ion transport, immune response and acetyltransferase activity, which partially agree with the analysis of our SNP-based GWAS data. This study provided the first report for the risk of multiple rare CNVs in the pathogenesis of narcolepsy.

2497T

Integrating GWAS and protein-protein interaction network identifies a subnetwork implicated in Alcohol Dependence. *S. Han[1], B.Z. Yang[1], H. Kranzler[2,3], F. Lindsay[4], H. Zhao[5], J. Gelernter[1].* 1) Department of Psychiatry, Yale University School of Medicine, New Haven, CT; 2) Department of Psychiatry, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania, USA; 3) VISN 4 MIRECC, Philadelphia VAMC, Philadelphia, Pennsylvania, USA; 4) Departments of Medicine (Biomedical Genetics), Neurology, Ophthalmology, Genetics & Genomics, Biostatistics, and Epidemiology, Boston University Schools of Medicine and Public Health, Boston, Massachusetts, USA; 5) Departments of Genetics, Biostatistics, Statistics, Yale University School, New Haven, CT.

Family and twin studies have established a genetic contribution to alcohol dependence (AD). However, few AD susceptibility genes have been identified and successfully replicated. In the current study, we aimed to integrate genomewide association studies (GWAS) and human protein-protein interaction (PPI) networks to identify a subnetwork of interacting genes that may work together to contribute to the risk of AD. Using two GWAS datasets from the Study of Addiction: Genetics and Environment (SAGE) and the Collaborative Study on the Genetics of Alcoholism (COGA), we first performed a dense module search within the human PPI network to construct a subnetwork enriched for genes associated with AD, followed by association tests between the subnetwork and AD in five independent case-control samples, including the European-American (EA) (1453 AD and 1217 controls) and African-American (AA) (708 AD and 442 controls) samples from the merged SAGE and COGA dataset, randomly selected samples (795 AD and 1433 controls) from the genomewide association study of alcohol use and alcohol use disorder in Australian Twin-Families (OZ-ALC), and our collected EA (1021 AD and 473 controls) and AA (1641 AD and 1367 controls) samples. We established a subnetwork of 39 genes that are not only enriched for genes associated with AD, but also are collectively associated with AD in both EAs ($P<0.0001$) and AAs ($P=0.0008$) from SAGE and COGA samples. We further replicated the association in three other independent case-control samples, including the Euroepan-ancestry Austalian twin sample ($P=0.0003$), our EAs ($P=0.004$) and and our AAs ($P=0.005$). Functional enrichment analysis of the genes in the subnetwork revealed biological mechanisms that may underlie some of the genetic risk for AD, such as cation transport, ion transport, synaptic transmission and transmission of nerve impulse. In conclusion, we identified a subnetwork implicated in AD that is biological meaningful and highly reproducible, and may provide important clues to future research into the etiology and treatment of AD.

2498F

APP associated with late-onset Alzheimer disease in autopsy-confirmed dataset. *T. Montine[1], G.W. Beecham[2], R. Rajbhandary[2], K.L. Hamilton[2], A.C. Naj[2], E.R. Martin[2], R. Mayeux[3], J.L. Haines[4], L. Farrer[5], G. Schellenberg[6], M.A. Pericak-Vance[2], The Alzheimer Disease Genetics Consortium.* 1) Department of Pathology, University of Washington, Seattle, WA; 2) John P. Hussman Institute for Human Genomics, University of Miami, Miami, FL, USA; 3) Taub Institute of Research on Alzheimer's Disease, Columbia University, New York, NY, USA; 4) Center for Human Genetics Research, Vanderbilt University, Nashville, TN, USA; 5) School of Medicine, Boston University, Boston, MA, USA; 6) School of Medicine, University of Pennsylvania, Philadelphia, PA, USA.

Background: Late-onset Alzheimer disease (LOAD) is a highly heritable neurological disease with several known genetic risk loci (APOE, CR1, etc). With the exception of APOE, these loci were identified using GWAS in large datasets and have small effects on risk (e.g., odds ratios ~= 1.1). However, phenotypic heterogeneity may obscure the true effects of these and other loci. To investigate this concern we performed a GWAS between autopsy-confirmed cases (ACC) and controls and compared the results to a GWAS of clinically identified cases (CIC) and controls. Methods: Samples were derived from 12 cohorts that are part of the Alzheimer Disease Genetics Consortium. Genotyping was performed on high-density genotyping chips and imputed to a 1000 Genome Project map. Logistic regression was performed within each cohort and METAL was used to meta-analyze across cohorts. Datasets consisted of 4,854 neuropathology confirmed cases and 8,264 controls (ACC set) and 5,190 clinically identified cases and 7,400 controls (CIC set). Both control sets included a mix of clinically identified and neuropathology confirmed controls. Results: Most known LOAD loci were confirmed in both the ACC and CIC analyses with at least nominal association ($p<0.05$); the lone exception being EPHA1, whose effect trended in the reported direction, but did not reach statistical significant association in the autopsy dataset (p-value=0.20). There were no novel associations that met a genome-wide correction for multiple-testing in either analysis. There were however several loci with strong association ($p <5\times10\text{-}6$) in the ACC analysis that were not associated in the CIC analysis. Notably, common variants in the amyloid precursor protein (APP) were associated with autopsy cases (rs4817090, OR=0.85, $p=2.4\times10\text{-}6$), but not with clinic-based cases (OR=0.98, p=0.55). Conclusion: While it is known that rare variation in APP can lead to early-onset forms of AD, this study is the first to implicate common variants in APP in a large case-control study. It suggests common variation in APP may influence LOAD pathology in a way that a more heterogeneous clinic-based sample may not be adequately powered to detect, and confirms the utility of an autopsy-confirmed dataset. To further follow-up we are examining association with more detailed neuropathology features, and investigating more stringently defined autopsy-confirmed cases and controls.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2499W

An Association Study Comparing Children with Extreme High and Extreme Low Obsessive-Compulsive Traits in the General Population. *P.D. Arnold[1], V. Sinopoli[2], J. Crosbie[1], L. Park[2], A. Dupuis[3], A. Paterson[2], R.J. Schachar[1].* 1) Dept Psychiatry, Hosp Sick Children, Toronto, ON, Canada; 2) Program in Genetics and Genome Biology, Hospital for Sick Children, Toronto, ON, Canada; 3) Program in Child Health Evaluative Sciences. Hospital for Children, Toronto, ON, Canada.

Objective: Obsessive- Compulsive disorder (OCD) is a common, debilitating and phenotypically heterogeneous disorder which has been shown to be more highly heritable when symptoms begin in childhood or adolescence. OC symptoms are quantitative traits, continuously distributed in the general population and ignoring the quantitative nature of OCD may reduce the power of genetic association studies comparing OCD patients to controls, particularly if controls are unscreened. The goal of the study was to utilize a novel and cost-effective design in which we conduct a GWAS comparing children on the extreme high and extreme low ends of a quantitative distribution of OC traits measured in the general population. Methods: The sample consisted of 16,099 children and adolescents recruited from the community. Self- and/or parent-report data on obsessive-compulsive (OC) symptoms was collected from all participants. From 7545 unrelated children and adolescents of European Caucasian descent, we selected individuals with the highest (top 10 per cent, N=755) and extreme lowest (bottom 10 per cent, N=755) scores on OC symptoms after controlling for age and gender. Genotyping is being conducted using the Illumina Omni 2.5 beadchip. In addition, we have genotyped the serotonin transporter linked polymorphic region (5HTTLPR) including an insertion/deletion polymorphism and rs25531, variants which in combination have been shown to influence SLC6A4 expression. Quality control analyses were conducted using PLINK, including multidimensional scaling (MDS) for population structure and the Cochrane-Armitage trend test for detection of association. Results: To date, we have analyzed the 5HTTLPR and identified no significant association (p=0.15) between high- and low- expressing genotypes of promoter variants in this region and OC traits. The GWAS is ongoing and results will be available at the time of the presentation. Quality control analyses on an initial set of 191 individuals resulted in 7/191 exclusions due to low call rate; (n=2), sex discrepancy (n=1) and non-European ancestry based on MDS plots (n=4). Conclusions: This research has the potential to increase our ability to identify candidate variants for future biological investigation, to facilitate the understanding of the mechanism by which genetic risks result in OCD. This will be the first report of the utility of performing a genome-wide study of quantitative OC traits.

2500T

Irish high-density schizophrenia families are enriched for common polygenic effects. *T.B. Bigdeli[1], S.A. Bacanu[1], B.T. Webb[1], B.S. Maher[2], D.L. Thiselton[1], B. Wormley[1], F.A. O'Neill[4], D. Walsh[5], A.H. Fanous[3,1], B.P. Riley[1], K.S. Kendler[1].* 1) Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, Richmond, VA; 2) Department of Mental Health, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD; 3) Washington VA Medical Center Washington, D.C; 4) Department of Psychiatry, Queens University, Belfast, UK; 5) The Health Research Board, Dublin, Ireland.

Although the combined evidence of family, twin, and adoption studies supports a substantial aggregate genetic component for schizophrenia, there is no evidence of either Mendelian inheritance in affected families nor genes of large effect in the general population. Recent genome wide association studies (GWAS) of schizophrenia provide strong support for a substantial common polygenic contribution of a large number of small effects. However, the respective contributions of rarer or structural variation have not yet been realized. Of interest herein is the supposition that the occurrence of multiple cases in a family supports segregation of a transmissible genetic risk factor, as opposed to de novo mutations having given rise to sporadic cases. We sought to assess the predictive value of polygenic findings reported by the Psychiatric GWAS Consortium (PGC), as applied to a sample of Irish high-density Schizophrenia families (ISHDSF). Using genome-wide SNP data available for affected and unaffected relatives, as well as a larger panel of imputed SNP genotypes, per-individual polygenic risk scores were constructed based on the PGC analysis. Per-family averages, calculated from the normalized scores for all family-members were shown to be elevated (mean Z-score=0.424, SD=0.761). This enrichment for common effects remained significant following permutation of simulated PGC datasets and recalculation of per-individual scores ($P_{permutation}$=0.01). Polygenic scores were found to be increased among probands (mean Z-score=0.425) compared to their non-psychotic relatives (mean Z-score=0.341, SD=0.891). However, comparison of mean scores by categorical diagnoses of psychotic illness revealed greater enrichment among less-severely affected relatives of probands (mean Z-score=0.649, SD=1.007). Because families were clustered geographically, we addressed whether the observed enrichment could exhibit regional differences. These findings indicate a non-random distribution of family means across the historical counties of the Republic of Ireland and Northern Ireland. In summary, we have demonstrated empirically that the unaffected members of multiply-affected schizophrenia families may be enriched for widely-replicated, common polygenic effects. Furthermore, between-family differences in patterns of enrichment among psychotic and non-psychotic relatives highlights an apparent heterogeneity of genetic architecture in schizophrenia.

2501F

Genome Wide Association Study for Maternal Effects in Autism. *S. Buyske[1], E.S. Stenroos[2], W.G. Johnson[2].* 1) Statistics & Biostatistics, Rutgers Univ, Piscataway, NJ; 2) Neurology, Univ of Medicine & Dentistry of New Jersey-Robert Wood Johnson, Medical School, Piscataway, NJ.

Maternally acting gene alleles (MAGAs) are hypothesized to act in mothers prenatally to alter the fetal environment and affect the phenotype in offspring, independently of whether or not they are inherited by the fetus. Some evidence of MAGAs in autism has been found using a candidate gene approach, but there appear to be no published genome-wide association studies for any trait on MAGAs. We conducted a genome-wide analysis for MAGAs using existing data from the Autism Genetic Resource Exchange (AGRE) on the Illumina Hap550 array in 825 families and identified a region of chromosome 3p24.3 that might contain one or more MAGAs. A log-linear analysis based on asymmetries in parental genotypes showed rs12487874, intronic in the *RTFN1* gene, as genome-wide significant (p=8.01e-09) for a maternal effect with no evidence of child effect. Analysis of existing data from the Autism Genome Project (AGP) in 948 families (excluding AGRE subjects) on the Illumina 1M array showed that rs12636481, 88kb from rs12487874, was highly significant (p = 7.89e-15) for a maternal effect with no evidence of child effect; rs12487874 did not pass quality control in the AGP study. A second locus was genome-wide significant for a maternal effect with no evidence of child effect in the AGP study: rs2741849 in an intronic region of the *PPP1R12B* gene on 1q32.1 (p = 4.03e-10); this region is poorly covered by the Hap550 array. Follow-up work based on imputated genotypes using a 1000 Genomes Project reference panel is underway.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2502W

Analysis of putative pathogenic CNVs in autistic patients of Portuguese origin. *I.C. Conceição[1,2,3], B.A. Oliveira[1,2,3], C. Correia[1,2,3], J. Coelho[1], C. Café[4], J. Almeida[4], S. Mouga[4,5], F. Duque[4,6], G. Oliveira[4,5,6,7], A.M. Vicente[1,2,3].* 1) Instituto Nacional de Saúde Dr Ricardo Jorge, Lisbon, Portugal; 2) Center for Biodiversity, Functional & Integrative Genomics, Lisboa, Portugal; 3) Instituto Gulbenkian de Ciência, Oeiras, Portugal; 4) Unidade Neurodesenvolvimento e Autismo, Centro de Desenvolvimento, Hospital Pediátrico (HP); Centro Hospitalar e Universitário de Coimbra (CHUC), Coimbra, Portugal; 5) Instituto Biomédico de Investigação em Luz e Imagem, Faculdade de Medicina da Universidade de Coimbra, Coimbra, Portugal; 6) Faculdade de Medicina da Universidade de Coimbra, Coimbra, Portugal; 7) Centro de Investigação e Formação Clinica do HP-CHUC, Coimbra, Portugal.

Recent reports by the Autism Genome Project (AGP) consortium and other groups show that Copy Number Variants (CNV), while individually rare, collectively may explain a large fraction of the etiology of Autism Spectrum Disorders (ASD). The goal of this study was to establish the relevance for ASD etiology of potentially pathogenic CNVs identified in a Portuguese population sample by the AGP whole genome CNV analysis. A total of 14218 CNVs were identified in 342 Portuguese probands (35 females, 307 males), genotyped by the AGP using the Illumina Infinium1M SNP microarray. From these, we selected 936 CNVs that did not overlap more than 50% with approximately 8000 controls in available databases. These CNVs range from about 5 Kb to 1.9 Mb, and 52% are deletions (del), while 48% are duplications (dup). Larger CNVs (>500 Kb) are more frequently duplications than deletions; also, the ratio of dup/del for CNVs that do not include any genes is approximately 40/60, while for CNVs with more than three genes the ratio is inverted (60/40). These 936 CNVs were inspected for recurrence rates in this population and inclusion of ASD-implicated or candidate genes, or other genes of relevance for brain function. Particularly interesting CNVs were validated by qPCR in patients and relatives, and breakpoint mapping was performed by Long-range PCR. ASD genes like *SHANK3*, *NRG1*, *CNTNAP2* were identified in, respectively, 1, 2 and 1 probands. Genes such as *ATRX* and *DPYD*were duplicated and deleted, respectively, in 10 and 22, patients, but not in controls. The same is true for *NRP1*, *VAULTRC2* (a vault RNA) and the 5q13.2 region. Two genes from the*Annexin* family (*ANXA1* and *ANXA7*) were also exclusively duplicated or deleted in autistic patients. *ANXA1*, known to have a neuroprotective role against inflammation, is present in 4 patients but not in controls from databases or in 410 Portuguese controls. Analysis of clinical presentation did not identify any specific symptom patterns or associated co-morbidities that might specifically characterize any of these CNVs. Segregation analyses in families and correlation with the presence of mild broad spectrum signals in relatives is ongoing. This exhaustive analysis of CNVs for clinical interpretation is necessary for the efficient translation of this knowledge into clinical practice, aiming at the development of molecular tools that may assist behavioral screening procedures for early diagnosis in ASD as well as genetic counseling.

2503T

GWAS Identifies Biologically Relevant SNP Associations with Sexual Partnering Behavior. *J. Gelernter[1, 2], H.R. Kranzler[3], R. Sherva[4], R. Koesterer[4], L. Almasy[5], H. Zhao[1], L.A. Farrer[4].* 1) Yale University School of Medicine, New Haven, CT; 2) VA CT Healthcare Center, West Haven, CT; 3) University of Pennsylvania School of Medicine, Philadelphia, PA; 4) Boston University School of Medicine, Boston, MA; 5) Texas Biomedical Research Institute, San Antonio, TX.

The specific factors influencing human sexual partnering are poorly understood. Arguably, in the pre-modern era, multiple mating may have been tied to selection for traits related to survival including resistance to infection and starvation, strength, and certain behaviors. Recently, we completed a GWAS using the Illumina Omni-Quad microarray in ~5800 African- and European-American (AA and EA) participants in genetic studies of alcohol, cocaine, and opioid dependence. Subjects were interviewed using the Semi-Structured Assessment for Drug Dependence and Alcoholism (SSADDA) - an instrument that covers all major DSM-IV diagnoses as well as other numerous psychiatric and lifestyle traits. One of these is the response to: “How many sexual partners have you had in your life?” Association of age-adjusted residuals of this variable with more than 3 million SNPs reliably imputed using the 1K Genomes reference panel was tested in each sex*population subgroup using generalized estimating equations. Results from subgroup analyses were combined by meta analysis. SNPs with p-values <1E-06 were genotyped in a replication sample of ~2300 subjects. Genomewide-significant results were obtained for 13 SNPs including ones that map to genes coding proteins involved in reproductive-related functions (e.g., rs74738626 in KCNU1 which encodes a testes-specific K+ channel [p=1.2E-12], rs78227383 in NME5, a nucleoside diphosphate kinase which may have a specific function in the phosphotransfer network involved in spermatogenesis [p=4.0E-11 in EAs only], and rs76221611 in CCND2 which encodes cyclin D2, shown to be highly expressed in ovarian and testicular tumors [p=3.3E-11 in AAs only]), immune response (e.g., rs2709778 in GARS which encodes gylcyl-tRNA synthetase shown to be a target of autoantibodies in human autoimmune diseases [p=1.0E-10 in males only]), and other genes of biological interest (e.g., rs10849971 in ALDH2, an alcohol-metabolizing enzyme that is also an alcohol dependence risk locus [p=9.6E-09 in females only]). These findings have clear implications with respect to normal sexual function and potentially for risk of sexually transmitted disease.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2504F

A genome wide association study (GWAS) of alcohol dependence (AD) using a hybrid design. *L.M. Hack[1], A.E. Adkins[1], B.T. Webb[2], B.S. Maher[3], D.G. Patterson[4], D. Walsh[5], C.A. Prescott[6], D.M. Dick[1,2], K.S. Kendler[1,2], B.P. Riley[1,2].* 1) Department of Human and Molecular Genetics, Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, Richmond, VA; 2) Department of Psychiatry, Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, Richmond, VA; 3) Department of Mental Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA; 4) Shaftesbury Square Hospital, Belfast, Northern Ireland, United Kingdom; 5) Health Research Board, Dublin, Ireland; 6) Department of Psychology, University of Southern California, Los Angeles, CA, USA.

Background: The powerful, systematic, and unbiased GWAS has been successful in identifying replicated susceptibility variants for numerous complex diseases. Compared to several other psychiatric disorders, the total sample size for GWAS of AD is more modest with individual samples in the range of 500–2500 individuals. Our sample is comparably sized (total N= 1934) and consists of related cases and unrelated, lightly screened controls. **Methods:** The cases are from the Irish Affected Sib Pair Study of Alcohol Dependence and were diagnosed using DSM-IV criteria. Genotyping was conducted using the Affymetrix V6.0 array by three separate genotyping core facilities. Because artifacts are a known issue when combining samples genotyped at multiple sites, genotypes were called using BeagleCall, which considers both allele signal intensities and LD information. After quality control filtering, the sample contained 712 cases, 1762 controls, and 676,736 SNPs for analysis. The within-site and cross-site duplicate error rates were 0.0043% and 0.0054%, respectively. Methods for analysis using the hybrid design are not widely implemented and the optimal analysis approach is being investigated. Here we present results from an independent case control analysis consisting of 433 cases and 1762 controls. The analysis was conducted in PLINK using logistic regression with age and sex as covariates. **Results:** While no SNP reached genome-wide significance (p<7.39×10^{-8}), there were 4 SNPs in high LD within chromosome 15 in the suggestive range (p=10^{-7}-10^{-5}) that fall within *RYR3*, which encodes an intracellular calcium channel. Of the three ryanodine isoforms in humans, RyR3 is the most highly expressed in the brain (Giannini et al., 1995). The ryanodine family of receptors has been shown to modulate BK channels, which were implicated in AD and alcohol-related traits in previous GWA as well as model organism studies of ethanol behaviors (Chavis et al., 1998; Davies et al., 2003; Edenberg et al., 2010; Kendler et al., 2011). Furthermore, proteins in the ryanodine family have been found to interact with proteins in the CLIC family, intracellular chloride channels implicated in acute ethanol behaviors in model organisms (Song et al., 2010; Bhandari et al., 2012). **Discussion:** We have identified a locus associated with AD that should be followed-up in independent samples. Studies in model organisms of this gene are currently being pursued within the VCU Alcohol Research Center.

2505W

Genetic analysis of tobacco use in the Swedish Twin Registry: Genome wide association studies of cigarette and snus consumption. *PA. Lind[1], D. Cesarini[2], DJ. Benjamin[3], PD. Koellinger[4], PKE. Magnusson[5], M. Johannesson[6], SE. Medland[1].* 1) Quantitative Genetics, Queensland Inst Med Res, Herston, Australia; 2) Center for Experimental Social Science, Department of Economics, New York University, New York, NY, USA; 3) Department of Economics, Cornell University, Ithaca, NY, USA; 4) Department of Applied Economics, Erasmus Universiteit Rotterdam, Rotterdam, Netherlands; 5) Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden; 6) Department of Economics, Stockholm School of Economics, Stockholm, Sweden.

To identify genetic risk variants associated with increased nicotine intake through cigarette smoking and Swedish snus (oral smokeless tobacco) use, we performed genome-wide association studies (GWAS) for the number of cigarettes smoked per day (CPD) and boxes of snus consumed per week (BOXWK) in 4,100 (51% male) and 1,054 (male) unrelated individuals respectively from the Swedish Twin Registry. Association was conducted using both genotyped (Illumina OmniExpress) and imputed (1KGP-EUR-vI.3, MAF >1%) single nucleotide polymorphisms (SNPs) using both PLINK and Merlin, adjusting for age and sex, followed by gene and pathway enrichment analyses. Although no SNP reached genome-wide significance in either analysis, fifteen achieved P-values < 1 × 10^{-5} for CPD, including 10 variants located near or within the *CHRNA5-CHRNA3-CHRNB4* locus on chromosome 15q25. Fourteen SNPs achieved P-values < 1 × 10^{-5} for BOXWK, with 8 located on chromosome 4q21. There was little or no evidence for association between snus use and 15q25 variants (P ≥ 0.05). Secondary pathway analyses found several addiction-related pathways were enriched for SNPs associated with CPD and BOXWK. We report the first GWAS of snus use for which the underlying genetic basis is largely unknown. While both behaviours are important sources of nicotine intake within the Swedish population, we have shown that the pattern of association with CPD and BOXWK is different at the individual SNP level. While further replication is required, particularly for snus use, the identification of susceptibility loci and biological pathways will be important in characterizing the biological mechanisms that underpin both snus consumption and cigarette smoking.

2506T

Genome-wide association study for intracranial aneurysm in the Japanese population. *S. Low[1], A. Takahashi[1], M. Kubo[2], Y. Nakamura[3].* 1) Laboratory for Statistical Analysis, Center for Genomic Medicine, RIKEN, Tokyo, Japan; 2) Laboratory for Genotyping Development, Center for Genomic Medicine, RIKEN, Yokohama, Japan; 3) Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan.

Aneurysmal subarachnoid hemorrhage (aSAH) is the most serious subtype of stroke. Genetic factors have been known to play an important role in the development of intracranial aneurysm (IA), some of which further progress to subarachnoid hemorrhage (SAH). In this study, we conducted a genome-wide association study (GWAS) to identify common genetic variants that are associated with the risk of IA, using 1383 aSAH subjects and 5484 control individuals in the Japanese population. We selected 36 single-nucleotide polymorphisms (SNPs) that showed suggestive association (P <1 × 10^{-4}) in the GWAS as well as additional 7 SNPs that were previously reported to be associated with IA, and further genotyped an additional set of 1048 IA cases and 7212 controls. We identified an SNP, rs6842241, near *EDNRA* at chromosome 4q31.22 (combined P-value = 9.58 × 10^{-9}; odds ratio = 1.25), which was found to be significantly associated with IA. Additionally, we successfully replicated and validated rs10757272 on *CDKN2BAS* at chromosome 9p21.3 (combined P-value = 1.55 × 10^{-7}; odds ratio = 1.21) to be significantly associated with IA as previously reported. Furthermore, we performed functional analysis with the associated genetic variants on *EDNRA*, and identified two alleles of rs6841581 that have different binding affinities to a nuclear protein(s). The transcriptional activity of the susceptible allele of this variant was significantly lower than the other, suggesting that this functional variant might affect the expression of EDNRA and subsequently result in the IA susceptibility. Identification of genetic variants on *EDNRA* is of clinical significance probably due to its role in vessel hemodynamic stress. Our findings should contribute to a better understanding of physiopathology of IA.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2507F

Genome-wide association study identified susceptibility loci associated with nicotine dependence in a Japanese population. *D. Nishizawa[1], K. Shinya[1], J. Hasegawa[1], N. Sato[2,3], F. Tanioka[4], H. Sugimura[3], K. Ikeda[1].* 1) Addictive Substance Project, Tokyo Metropolitan Institute of Medical Science, Tokyo, Japan; 2) Department of Clinical Nursing, Hamamatsu University School of Medicine, Hamamatsu, Japan; 3) Department of Tumor Pathology, Hamamatsu University School of Medicine, Hamamatsu, Japan; 4) Department of Pathology, Iwata City Hospital, Iwata, Japan.

Objectives: Many genetic and environmental factors can be involved in the etiology of nicotine dependence (ND). To date, several candidate genetic variations have been identified to be associated with smoking behaviors and vulnerability to ND. These were found mostly by investigating single nucleotide polymorphisms (SNPs) on each gene related to phenotypes of concern or by genome-wide association studies (GWAS), which treat genetic markers in the whole genome, conducted for subjects with European ancestry. However, genetic factors have not been reportedly investigated for Japanese population utilizing whole-genome genotyping arrays. We comprehensively explored genetic contributors to ND by GWAS in Japanese. **Methods**: Subjects were 300 patients who visited Iwata City Hospital. A number of participants involved in this study had various smoking habits and filled in a questionnaire leaflet containing various questions. Whole-genome genotyping was performed for 300 samples with iScan System (illumina, San Diego, CA) and the BeadChip HumanCytoSNP-12 v2.0. A two-stage GWAS was conducted using 150 samples for each stage to investigate the association between genotype distributions and severity of ND, which was measured by scores of the Fagerström Test for Nicotine Dependence (FTND), a test that yields a continuous measure of ND, the Tobacco Dependence Screener (TDS), consisting of 10 questions, and the numbers of cigarettes smoked per day (CPD). **Results and Conclusion**: In association study between over 200,000 marker SNPs and FTND, TDS, and CPD, none of the SNPs were found to reach the genome-wide significant level and the *P*-values were no less than 10^{-6} in all analyses. However, several potent SNPs were found in loci that have not been highlighted, as well as in loci that include known candidate genes resulted from previous GWAS. Among them was *CSMD1*, CUB and Sushi multiple domains 1, which appeared in top ranks of our GWAS results for all the three phenotypes examined. Although future studies with larger sample size is required, these results will serve to discover genetic factors contributing to ND and smoking behavior specific to Japanese population in addition to those common to other populations.

2508W

Genome-wide Meta-analysis of Autism Spectrum Disorders (ASD) Supports *A2BP1* as an ASD Candidate Gene. *Y.S. Park[1,2], J. Jaworski[1], I. Konidari[1], P. Whitehead[1], M.L. Cuccaro[1], E.R. Martin[1,2], J.L. Haines[3,4], M.A. Pericak-Vance[1,2].* 1) John P. Hussman Institute for Human Genomics, University of Miami, Miami, FL, USA; 2) Dr. John T. Macdonald Department of Human Genetics, University of Miami, Miami, FL, USA; 3) Vanderbilt University School of Medicine, Vanderbilt University, Nashville, TN, USA; 4) Center for Human Genetics Research, Vanderbilt University, Nashville, TN, USA.

Recent genetic studies have identified both common and rare genetic variations contributing to autism spectrum disorders (ASD) in multiple genes including *RELN*, *FOXP2*, and *SHANK3*. However, much of the genetic contribution of this highly heritable disorder (h^2>0.9) remains unidentified. We conducted a genome-wide association study of 2,995 affecteds and 7,057 unaffecteds in order to identify novel genetic associations of common SNPs to ASD. Our discovery data set included 620 families (694 affected individuals) ascertained and genotyped at John P. Hussman Institute for Human Genomics and Vanderbilt Center for Human Genetics Research, as well as 696 families (1,240 affected individuals) ascertained and genotyped at Autism Genetics Resource Exchange. The analysis of the discovery data set alone has been published elsewhere (Ma et al., *Annals of Human Genetics*, 2009). Our replication data set included 1,065 families (1,061 affected individuals) ascertained and genotyped by Simons Simplex Collection. Sample genotyped data were imputed up to a common set of 7,560,685 SNPs with IMPUTE2 v.2.1.2 using the 1,000 Genomes Project June 2011 Release as a reference. Generalized estimating equations were used with covariate adjustments for sex and population substructure. While no SNP associations met the threshold for genome-wide significance ($p<5\times10^{-8}$), 10 SNPs demonstrated associations of p<0.001 from the joint meta-analysis of all data sets and showed p<0.05 in both discovery and replication analyses. Of these 10 SNPs, 3 SNPs were in introns of a chromosome 16 gene *A2BP1* (rs3095508, meta-p=6.829×10^{-4}; rs42189 meta-p=8.038×10^{-4}; and rs811919 meta-p=9.188×10^{-4}). *A2BP1* is a known and compelling ASD candidate gene and a recent study has shown an evidence of dysregulated splicing of*A2BP1*-dependent alternative exons in the ASD brain using transcriptome analysis (Voineagu et al., *Nature*, 2011).

2509T

A polygenic analysis of autism, depression and schizophrenia risk alleles' effect on reproductive fitness in the general population. *R.A. Power[1], R. Uher[4], S. Ripke[3], C.M. Lewis[1], P. McGuffin[1], P. Lichtenstein[2], H. Wallum[2].* 1) Social Genetic & Developmental Psychiatry Centre, Institute of Psychiatry, London, United Kingdom; 2) Karolinska Instiutet, Sweden; 3) The Analytic and Translational Genetics Unit, Massachusetts General Hospital, U.S.A; 4) Brain Repair Centre, Dalhousie University, Halifax, Canada.

The schizophrenia paradox has long existed in psychiatric research: how does a disorder remain highly prevalent if also highly heritable and damaging to an individual's reproductive fitness? Following from previous work showing increased fecundity in relatives of those with psychiatric illnesses, we looked for a similar molecular signal. With progress in genome-wide association studies, it is now possible to directly examine how risk alleles affect reproductive fitness in the general population. While few causal loci have been replicated for psychiatric disorders, methods using aggregate scores of risk alleles have been reasonably successful in explaining a significant proportion of variance in affection status. In this analysis we tested for an association between polygenic scores for autism, depression and schizophrenia with number of grandchildren in the general population. Using data from the Psychiatric GWAS Consortia analysis of autism, depression and schizophrenia, polygenic scores were constructed within a Swedish dataset of 6,219 individuals born in 1910–43. The Swedish Multi-Generation Register was used to identify their grandchildren. Linear regression was then used to test for association between polygenic scores for the disorders and number of grandchildren. Autism and depression scores weren't associated with reproductive fitness at any threshold for significance. As it has been suggested depression risk variants reflect `plasticity alleles', we tested for a difference in mean variance in fitness across depression scores. The results were in the expected direction with increased variance for those with increased depression score but not significant (p=0.11). Greater schizophrenia scores were associated with fewer grandchildren (p=0.008), explaining 0.11% of the variance. This association was not driven by extreme scores or a sex specific effect. These results exacerbate the schizophrenia paradox by increasing the negative selection against schizophrenia risk alleles, particularly as only a portion of the heritability of schizophrenia is captured by the polygenic score. The results may reflect deleterious mutations that lead both to schizophrenia and also behavioural deviations that are sub-clinical but severe enough to reduce fitness. This might be evidence for the need to better screen controls for sub-clinical features or to view psychosis on a continuum. Overall there is no evidence for balancing selection for these disorders.

2510F

Maternal genetic effects on autism risk: results from the Early Markers for Autism (EMA) study. *J.F. Quinn[1], L.C. Croen[2], L.W. Weiss[1].* 1) Institute for Human Genetics, University of California, San Francisco, San Francisco, CA; 2) Autism Research Program, Kaiser Permanente Division of Research, Oakland, CA.

Purpose: Studies of main genetic effects in autism have failed to explain the high heritability estimates for the disease suggested by classical twin studies and by correlation of broader autism phenotypes within families. Recently, researchers have begun to study the prenatal maternal environment as a risk factor in autism (PMID 18262386; 22012245; 18571628; 21810230). We hypothesize that maternal genetic effects, including an interaction between maternal and filial genetic variation (transgenerational epistasis), may influence susceptibility to autism. Methods: To test our hypothesis of maternal genetic effects in autism, we examined a sample of 400 maternal-filial duos where the child has been diagnosed with autism and 400 matched control duos where the child does not have autism. Using the Affymetrix Axiom EUR array, we genotyped our samples at approximately 600k single nucleotide polymorphisms (SNPs). We performed two novel genome-wide association studies: first, we compared the mothers who had given birth to an autistic child to the control mothers (maternal genetic main effects); second, we considered the maternal-filial genotype combination at each marker and performed several tests to see if the relative proportion of genotype combinations differed between case and control duos (transgenerational epistasis). Specifically, at each locus we compared case and control cohorts using three models: 1) comparing the proportion of duos that have identical genotypes, 2) comparing the proportion where the mother possesses an allele that her child does not possess, and 3) comparing the proportion where the child possesses an allele that the mother does not possess. Results: No results were genome-wide significant (P < 5e-8). For the test of maternal genetic main effects, we found variation within or nearby the following genes to be significant at the $P < 10^{-5}$ level: *MAML2, KRT18P55, and XK*. For the tests of transgenerational epistasis, we found variation within or near the following genes to be significant at the $P < 10^{-5}$ level in at least one of our three models: *C4orf37, INSC, SLC7A8, GABRB3, GOT2, PALM2, KIAA0430, SLC14A2, SEMA3C, LY86, and CDH11*. In addition, we found variation in or near the following genes previously implicated in autism to be significant at the $P < 10^{-4}$ level in at least one of our three models: *MACROD2, NRXN1, RELN, GRIN2B, and PCDH9*. We intend to follow up our results from this discovery cohort through replication.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2511W

Genome-wide pathway analysis in multiple cohorts implicates histone genes and inflammation in hippocampal atrophy. *V.K. Ramanan[1,2,3], L. Shen[2,4], J.D. West[2], S. Kim[2], S.L. Risacher[2], K. Nho[2], Y. Wang[2], J. Wan[2], T.M. Foroud[1,4], L.A. Farrer[5], A.J. Saykin[1,2,4], The Alzheimer's Disease Neuroimaging Initiative (ADNI).* 1) Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN; 2) Center for Neuroimaging, Department of Radiology and Imaging Sciences, Indiana University School of Medicine, Indianapolis, IN; 3) Medical Scientist Training Program, Indiana University School of Medicine, Indianapolis, IN; 4) Center for Computational Biology and Bioinformatics, Indiana University School of Medicine, Indianapolis, IN; 5) Department of Medicine (Biomedical Genetics), Boston University School of Medicine, Boston, MA.

Hippocampal volume loss is a known marker of prodromal and incident Alzheimer's disease (AD) and of other clinical states with impaired memory and cognition. Importantly, the genetic susceptibility architecture underlying hippocampal atrophy is not yet known and may involve multiple loci with differential effects across the structure's distinct subfields. To identify functional gene sets whose variants exhibit collective effects on hippocampal volume, we performed pathway analyses (Ramanan et al. *TIG* 2012) of genome-wide association study (GWAS) data in three independent cohorts with control, AD, and mild cognitive impairment or cognitive complaint participants (combined $N > 1000$). Volumes for whole hippocampus and 5 hippocampal subfields were extracted from MP-RAGE MRI scans using the FreeSurfer software and were adjusted for age, gender, education, handedness, and intracranial volume. Following genotype imputation and quality control, an additive model GWAS was performed for each volumetric phenotype in each cohort using Plink. All analyses employed a 1% minor allele frequency threshold and included diagnosis as a covariate. Gene-based p-values accounting for gene size and linkage disequilibrium were calculated within each GWAS using GATES (Li et al. *AJHG* 2011). Pathway analyses on these p-values were performed with GSA-SNP (Nam et al. *NAR* 2010) using gene sets from multiple pathway databases, and the false discovery rate (FDR) was applied to correct for multiple comparisons. A total of 35 pathways exhibited enrichment of association ($p_{FDR} < 0.05$) to a regional volume. In particular, pathways containing genes related to histones and regulation of DNA structure, repair, and transcription were enriched across several regions in multiple cohorts (smallest $p_{FDR} = 1.34 \times 10^{-7}$). Genetic variation in these pathways may impact susceptibility to epigenetic histone modifications related to learning and memory. Other enriched pathways included genes involved in immune response and cytokine signaling, such as *TNF*, and genes involved in synaptic plasticity for learning and memory, such as *NFKB* and *RXRA*. Subsequent analyses of plasma and cerebrospinal fluid analytes in one cohort revealed significant correlations between hippocampal volumes and candidate markers for stress and inflammation. Further investigation of the interactions between inflammation and DNA integrity in hippocampal development and neurodegeneration appears warranted.

2512T

Dissection of genetic architecture of bipolar disorder and schizophrenia: Results from a combined dataset of nearly 40,000 individuals. *S. Ripke[1], D.M. Ruderfer[2], Psychiatric Genomics Consortium.* 1) Analytic and Translational Genetics Unit, Mass Gen Hosp, Boston, MA; 2) Division of Psychiatric Genomics, Department of Psychiatry, Mount Sinai School of Medicine, NY, NY.

Background Bipolar disorder (BP) and schizophrenia (SCZ) are two chronic debilitating psychiatric illnesses that together affect ~3% of the world's population. They have been dichotomized since the early part of the century and this distinction has been maintained in research and epidemiology until very recently. A major aim of the Psychiatric Genomics Consortium is to quantify the extent of common genetic variation across distinct psychiatric diseases. Here, we present the insights generated by the SCZ / BP working group in better understanding the specific loci shared by both diseases as well as those that discern the two. Methods Genotypes were present for all individuals and were obtained from a variety of Illumina and Affymetrix arrays. Imputation was performed using the Beagle software. All analyses were done using dosage data in Plink. Results We analyzed a dataset of 39,202 independent individuals by combining the PGC SCZ sample (9,369 cases and 8,723 controls) and the PGC BP sample (10,410 cases and 10,700 controls). All subjects were distinct and independent. As reported in previous literature, we find highly significant polygenic overlap across disorders and many shared associated loci. Several follow-up analyses were performed in order to dissect the signal at regions/genes of interest for disease specificity. Furthermore we investigated the presence of genetic loci with significantly different effects by comparing odds ratios as well as creating a disease only dataset and performing GWAS with SCZ as case and BP as control. Finally we are investigating the impact of mood symptoms in SCZ, as well as psychosis in BIP, on overall polygenic overlap between the two disorders. Discussion Our data show that large cross-disorder datasets can be a powerful resource for understanding the shared genetic architecture of these disorders and ultimately will help in defining the differences in molecular function and treatment.

2513F

The Kaiser Permanente/UCSF Genetic Epidemiology Research Study on Adult Health and Aging: Genome-Wide Association Study of Anxiety Disorders. *L. Shen[1], T. Hoffmann[2], M. Kvale[2], L. Sakoda[1], D. Ranatunga[1], L. Walter[1], S. Sciortino[1], D. Ludwig[1], Y. Banda[2], S. Hesselson[2], P. Kwok[2], R. Risch[2], C. Schaefer[1].* 1) Division of Research, Kaiser Permanente Northern California, Oakland, CA; 2) Institute for Human Genetics, UCSF, San Francisco, CA.

The Kaiser Permanente/UCSF Genetic Epidemiology Research Study on Adult Health and Aging (GERA) cohort includes over 100,000 individuals with comprehensive longitudinal medical records along with genome-wide genotype data. As such, it provides an unprecedented opportunity to conduct a number of genome-wide association studies (GWAS) of psychiatric disorders. Acute stress disorder (ASD), post-traumatic stress disorder (PTSD), and panic disorder are prevalent, persistent and disabling anxiety disorders. Numerous candidate gene studies have been published with inconsistent findings. As yet there have been no large GWA studies of these anxiety disorders. We conducted the first GWAS of stress disorders, combining ASD and PTSD in one GWAS, and the first large scale GWAS of panic disorder as another. The studies were limited to non-Hispanic whites; both cases and controls were genotyped on the custom Affymetrix Axiom EUR arrays with 674,518 SNPs. We identified a total of 6177 combined ASD and PTSD cases with two or more psychiatric diagnoses of ASD and/or PTSD; 23.6% of the cases had five or more ASD or PTSD diagnoses; 77.6% of cases were female. The GERA cohort supplied 63,613 controls. In an analysis using an additive model that controlled for age, gender, and ancestry, no SNP associations reached genome-wide level of significance (5 × 10-8). However, a suggestive association was found with a SNP in an intergenic region on chromosome 10 near the ANKRD20A gene (OR=1.11, P=1.42×10-7), as well as neighboring SNPs. A total of 1483 persons with two or more diagnoses of panic disorder and 75,379 controls were included in the GWAS of panic disorder. Over 50% of the cases had five or more diagnoses of panic disorder; 72.3% of the cases were female. Cases and controls were unscreened for other major axis I psychiatric disorders. In an analysis using an additive model that controlled for age, gender, and ancestry, no SNP associations reached genome-wide level of significance (5 × 10-8). Additional studies with larger samples may help in identifying important loci for these disorders. GWAS of psychiatric disorders remain challenging.

2514W

Genome-wide association of structural MRI data in a large, normally developing, pediatric population. *P.M.A. Sleiman[1,2,3], T. Satterthwaite[4], K. Ruparel[4], C. Kim[1], R. Chiavacci[1], M.E. Calkins[4], R.C. Gur[4], R.E. Gur[4], H. Hakonarson[1,2,3].* 1) Center Applied Genomics, CHOP, Philadelphia, PA; 2) Division of Human Genetics, Children's Hosp Philadelphia, Philadelphia, PA; 3) Dept Pediatrics, Perelman School of Medicine, Univ of Pennsylvania, Philadelphia PA; 4) Brain Behavior Laboratory, Perelman School of Medicine, Univ of Pennsylvania, Philadelphia PA.

We carried out a genome-wide association study of subcortical brain volumes derived from the structural MRI data of ~1390 normally developing children and young adults aged between 8 and 21. All subjects were recruited into the study from the greater Philadelphia area, imaged on the same scanner at the University of Pennsylvania and the resulting data processed through the same pipeline to ensure homogeneity. Volumetric data for five subcortical regions, amygdala, caudate, hippocampus, putamen and thalamus were scaled to account for differences in intracranial volume and tested for association as quantitative traits in linear models. All samples were genotyped on either Illumina HH550, HH610 or OmniExpress arrays at the Center for Applied Genomics and imputed to ~37M SNPs against the 1KGP data. The first 1000 samples were used as the discovery cohort and the following 390 subjects, increasing to 500, as the replication cohort. Genetic ancestry was determined by principal component analysis and the first 10 principal components included as covariates in the regression to adjust for population stratification. Copy number variants were called for each individual from the gene chip data and tested for association against the quantitative traits using the Nexus software package. The analysis revealed several loci that show genome wide association with specific subcortical volumes in our pediatric dataset. These data suggest that individual brain regions may be under individual genetic control. Copy number variation analysis is underway and the results will be presented together with the association data.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2515T

Gene-Based Analyses of Attention Deficit Hyperactivity Disorder (ADHD) GWAS Data. *I.D. Waldman[1], S.V. Faraone[2], Psychiatric Genomics Consortium: ADHD Subgroup.* 1) Dept of Psychology, Emory University, Atlanta, GA; 2) Depts of Psychiatry and of Neuroscience & Physiology, SUNY Upstate Medical University, Syracuse, NY.

Attention Deficit Hyperactivity disorder (ADHD) is a highly heritable disorder, which has led researchers to search for underlying genetic risk factors via both candidate gene studies and, more recently, large-scale Genome Wide Association Scans (GWAS's). Many extant candidate gene studies have yielded mixed results, while recent GWAS's have yielded some suggestive findings but no genome-wide significant results. Two possible reasons for this are the relatively small sample sizes of the first generation of ADHD GWAS's and the small effect sizes of individual SNPs that show suggestive association. These small effects are consistent with those for other psychiatric disorders and medical conditions, and have led geneticists to utilize several alternative analytic methods, such as analyses of haplotypes and gene pathway or network analyses. Alternative analytic techniques that are relatively underutilized are omnibus gene-based analyses that simultaneously evaluate the statistical significance of the association between a trait or disorder and multiple SNPs within particular genes and their flanking regions. These methods have the advantage that they greatly reduce the number of statistical tests performed and hence involve a gene-based rather than SNP-based threshold for genome-wide significance. In this study, we apply such gene-based analyses of association to the second generation of ADHD GWAS data. Specifically, we included data from 6669 parent-offspring trios and 3744 cases and 11552 controls from 9 ADHD GWAS samples and performed gene-based analyses of ADHD diagnosis and symptom dimensions using the program VEGAS, which combines SNP-based p-values with information on LD among the SNPs within each gene and its 50 kb flanking region to obtain an association p-value for each gene. We demonstrate that this gene-based analytic method yields evidence for association that is at least as strong as conventional SNP-based methods and in many cases is much stronger.

2516F

Genome-wide association study of bipolar disorder with a history of binge eating. *S.J. Winham[1], S. McElroy[2], S. Crow[3], M. Frye[4], C. Colby[1], D. Walker[4], D-S. Choi[4], M. Chauhan[5], J.M. Biernacka[1,4].* 1) Department of Health Sciences Research, Mayo Clinic, Rochester, MN; 2) Lindner Center of HOPE, Mason, OH; 3) Department of Psychiatry, University of Minnesota; 4) Department of Psychiatry and Psychology, Mayo Clinic, Rochester, MN; 5) Department of Psychiatry and Psychology, Mayo Clinic Health Systems, Austin, MN.

Bipolar disorder (BD) has been reported to be associated with obesity-related measures including eating disorders and body mass index (Wildes et al, 2007). In particular, a positive association between BD and binge eating disorder has been identified (McElroy et al, 2011). To date, multiple genome wide association (GWA) studies of BD have been performed. However, potential gene-environment interactions between SNPs and obesity-related factors have not been investigated for BD, and factors such as body mass index have not been considered as covariates or potential confounding factors in GWA studies of BD. We utilized publically available data from the Bipolar Disorder Genome Study Consortium (BiGS), part of the Genetic Association Information Network (GAIN), to investigate if genetic variants influencing BD are dependent on presence of a binge eating (BE) history. After quality control, we performed GWA analyses of 729,454 SNP markers using 1001 European American bipolar cases and 1034 mentally healthy controls. We first performed a GWA analysis of gene-environment interaction using a case-only approach; assuming independence between SNPs and BE, this analysis identifies SNPs associated differentially with BD depending on presence of BE. We also performed GWA analyses stratified by bipolar subtype: mentally healthy controls were separately compared to only cases with the presence (n=206) or absence (n=723) of BE. The rate of BE in bipolar cases in the analyzed GAIN data was 22.2%. In the genome-wide gene-environment interaction analysis, the top ranking SNP was rs6006893, in an intron of the PRR5 gene (OR=2.25, p=6.5E-7). PRR5 is a subunit of mTORC2, which may be implicated in obesity. When comparing BD cases with BE to controls, the strongest signal was rs13233490, which is 603kbp upstream of PER4, a pseudogene which may be related to circadian rhythm; the minor allele at this SNP was associated with higher odds of developing the bipolar + BE subtype (OR=6.37, p=3.9E-7). SNP rs2367911 in CACNA2D1, related to CACNA1C, was also highly ranked (OR=3.29, p=1.6E-6). This study demonstrates the importance of phenotypic sub-classification in genetic studies of complex psychiatric traits, and lays the groundwork for further genome-wide investigations of obesity-related measures in bipolar patients. Samples from the Mayo Clinic Bipolar Biobank will be utilized in future studies to examine potential genetic effects underlying both BP and obesity-related measures.

2517W

Variation in Genes Related to Cochlear Biology is Strongly Associated with Adult-onset Deafness in Border Collies. *J.S. Yokoyama[1], E.T. Lam[2], A.L. Ruhe[3,4], C.A. Erdman[1], K.R. Robertson[4], A.A. Webb[5,6], D.C. Williams[7], M.L. Chang[8], M.K. Hytonen[9], H. Lohi[9], S.P. Hamilton[1], M.W. Neff[3,10].* 1) Department of Psychiatry and Institute for Human Genetics, University of California, San Francisco, CA, USA; 2) Institute for Human Genetics, University of California, San Francisco, CA, USA; 3) Neurogenomics Division, The Translational Genomics Research Institute, Phoenix, AZ, USA; 4) Veterinary Genetics Laboratory, University of California Davis, Davis, CA, USA; 5) CullenWebb Animal Neurology and Ophthalmology Centre, New Brunswick, Canada; 6) Department of Clinical Neuroscience, Faculty of Medicine, University of Calgary, Calgary, AB, Canada; 7) School of Veterinary Medicine, University of California Davis, Davis, CA, USA; 8) Department of Anthropology, University of Oregon, Eugene, OR, USA; 9) Department of Veterinary Biosciences and Research Programs Unit, Molecular Medicine, University of Helsinki and Folkhälsan Research Center, Helsinki, Finland; 10) The Van Andel Research Institute, Grand Rapids, MI, USA.

Purpose: The domestic dog suffers from hearing loss that can have profound impacts on working ability and quality of life. We have identified a type of adult-onset hearing loss in Border Collies that appears to have a genetic cause, with an earlier age of onset (3–5 years) than typically expected for aging dogs (8–10 years). Studying this complex trait within pure breeds of dog may greatly increase our ability to identify genomic regions associated with risk of hearing impairment in dogs and in humans.

Methods: We performed a genome-wide association study (GWAS) to detect loci underlying adult-onset deafness in a sample of 20 affected and 28 control Border Collies. To further localize disease-associated variants, targeted next-generation sequencing (NGS) of one affected and two unaffected dogs was performed. We then performed targeted genotyping of additional cases (n=23 total) and controls (n=101 total) and for an independent replication cohort of 16 cases and 265 controls to validate NGS findings.

Results: We identified a region on canine chromosome 6 that demonstrates extended support for association surrounding SNP Chr6.25819273 (p-value = 1.09×10^{-13}). Through targeted NGS and genotyping of additional dogs, we identified variants in *USP31* that were strongly associated with adult-onset deafness in Border Collies, suggesting the involvement of the NF-κB pathway. We found additional support for involvement of *RBBP6*, which is critical for cochlear development. These findings highlight the utility of GWAS-guided fine-mapping of genetic loci using targeted NGS to study hereditary disorders of the domestic dog that may be analogous to human disorders.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2518T

Genome-wide association study of Korean patients with autism spectrum disorder. *H. Yoo[1,2], S. Kim[7], M. Park[8], I. Cho[9], S. Cho[1,2], B. Kim[1,2], J. Kim[1,2], S. Park[1,2], D.H Geschwind[3,4,5,6], Y. Kwak[10], J.K. Lowe[3,4,5], Korean Autism Genetic Study Consortium.* 1) Department of Psychiatry, Seoul National University Bundang Hospital, Seongnam, Gyeongi, Korea; 2) Deparment of Psychiatry, Seoul National University College of Medicine, Seoul, Korea; 3) Neurogenetics Program and Department of Neurology, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA 90095, USA; 4) Center for Autism Research and Treatment, Semel Institute, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA 90095, USA; 5) Center for Neurobehavioral Genetics, Semel Institute, University of California, Los Angeles, Los Angeles, CA 90095, USA; 6) Department of Human Genetics, University of California, Los Angeles, Los Angeles, CA 90095, USA; 7) Department of Pharmacology, College of Medicne, Eulji University, Daejon, Korea; 8) Department of Preventive Medicine, College of Medicine, Eulji University, Daejon, Korea; 9) Department of Psychiatry, Gacheon University of Medicine and Science, Incheon, Korea; 10) Department of Psychiatry, Jeju University, Jeju, Korea.

Objectives: To identify genetic factors underlying autism spectrum disorder (ASD) in a Korean cohort through 1) a genome-wide association study (GWAS) for qualitative ASD, 2) a GWAS for related quantitative behavioral traits, and 3) analysis of population structure between Koreans, Chinese and Japanese. Methods: Subjects with ASD were diagnosed based on DSM-IV-TR, using the Korean versions of the ADOS and ADI-R (Lord et al. 1994; Lord et al. 1997). A total of 146 patients were genotyped on the Affymetrix 5.0 platform. Genotype data were filtered for <5% missingness/sample or / SNP; MAF>0.01; HWE p>0.0001. Cases were verified as unrelated using identity-by-descent estimation in PLINK v1.07. Genotype data for n=858 sex-matched controls were drawn from the Korea Association Resource database. For quantitative traits, we explored three domains and five factors (Social interaction, Preoccupation, Aggression, Visuospatial Ability/Memory, Reading/Drawing) from the Korean ADI-R and analyzed quantitative association with EMMAX. Population structure was examined using multi-dimensional scaling (MDS; PLINK v1.07). Five SNPs with the lowest p-values in the exploratory GWAS and 61 tag SNPs within 100kb of the best-associated SNP on chr2q35 (MAF>0.1; $r^2 \geq 0.8$ in HapMap3 Asians) were genotyped using a GoldenGate™ Assay and tested for family-based association (FBAT v2.0.2) in an expanded cohort of 250 Korean families with ASD. Results: 1) GWAS of n=142 cases (87.4% autism, 79.55±33.74 months, 14.7% female) and n=858 controls revealed strong association between ASD and three correlated SNPs on chr2q35 (rs10932677, $p=6.91 \times 10^{-8}$; rs17773965, $p=1.60 \times 10^{-7}$; rs4674113, $p=1.06 \times 10^{-6}$). 2) The quantitative association test shows significant association of Visuospatial Ability/Memory with rs1278259 at chr10q26.2 (nominal $p=8.15 \times 10^{-7}$. Bonferroni-corrected p= 0.004). 3) MDS analysis confirms that Asians are clearly separated from other eight ethnic groups in HapMap3. Analysis of Asians alone shows that the Korean, Chinese, and Japanese populations fall into distinct groups. 4) In the family-based validation, rs2241196 was associated with ASD in an additive model (nominal p<0.01; adjusted p<0.05 after FDR). Conclusion: We observed strong associations between SNPs on chr2q35 and ASD in a Korean cohort.

2519F

Combined Genome-wide Association Study of Tourette Syndrome and Obsessive-Compulsive Disorder. *D. Yu[1,2,3,4], TSAICG, IOCFGC, TS GWAS Consortium.* 1) Psychiatric and Neurodevelopmental Genetics Unit, Center for Human Genetics Research, Massachusetts General Hospital, Boston, MA; 2) Tourette Syndrome Association International Consortium for Genetics (TSAICG); 3) International Obsessive Compulsive Disorder Foundation Genetic Consortium (IOCFGC); 4) TS GWAS Consortium.

Tourette syndrome (TS) and obsessive-compulsive disorder (OCD) are developmental neuropsychiatric disorders that are frequently co-morbid and thought to be etiologically related. Both disorders are inherited in a non-Mendelian fashion, and the identification of definitive TS and OCD susceptibility genes remain elusive. Here, we report a combined genome-wide association study (GWAS) of TS and OCD using European ancestry samples previously employed in independent TS and OCD GWAS studies, including 1286 TS cases, 1437 OCD cases, and 6638 ancestry-matched controls. In addition to general European descent individuals, these samples also include three population isolates, Ashkenazi Jews from North America and Israel, French Canadians from Quebec, Canada, and Afrikaners from South Africa. Following imputation of SNPs from the 1000 Genomes Project using IMPUTE2, no SNPs achieved genome-wide significance ($P < 5.0 \times 10^{-8}$), with the top hit being rs76879034 on chromosome 3 within *ITGA9* ($P = 9.94 \times 10^{-7}$). The top distribution of LD-pruned SNPs ($P < 0.001$) were significantly enriched for eQTLs in frontal cortex and methylation QTLs (mQTLs) in cerebellum, whereas no enrichment in lymphoblast cell lines was detected, implicating their potential involvement in gene expression and regulation in the brain and in the etiology of TS and OCD. Strong evidence for a polygenic component was found in 1279 non-isolate European OCD cases with matched controls ($P = 2.0 \times 10^{-4}$), explaining 3.3% of the phenotypic variance. By contrast, only borderline significance for polygenicity was found in 776 non-isolate European TS cases (P = 0.06), which is most likely due to the smaller number of TS samples. Unexpectedly, we did not detect a substantial shared polygenic component between TS and OCD, suggesting that these two phenotypically and pathophysiologically related disorders may arise through divergent causal pathways.

2520W

Genome-wide significant evidence of association of a *NUP153* missense variant with familial late-onset Alzheimer's disease (LOAD). *A.R. Parrado[1], B. Hooli[1], K. Mullin[1], L. Gotta[1], M. Mattheisen[2], C. Lange[2], L. Bertram[3], R.E. Tanzi[1].* 1) Genetics and Aging Research Unit, Mass General Institute for Neurodegenerative Disease (MIND), Department of Neurology, Massachusetts General Hospital, Charlestown, MA, USA; 2) Department of Biostatistics, Harvard School of Public Health, Boston, MA, USA; 3) Max-Planck Institute for Molecular Genetics, Berlin, Germany.

Alzheimer's disease (AD) is the most common form of dementia in the elderly. LOAD cases with increased familial aggregation, are assumed to harbor susceptibility alleles at a higher frequency than unaffected family members, therefore genetic association analysis focused on families is a favorable strategy that can lead to improve power to detect susceptibility loci. We set out to identify single nucleotide polymorphisms (SNPs) in the genome that associate with risk or protection for LOAD. For this purpose, we performed a genome-wide association study on two large well characterized independent cohorts containing families of self-report European descent, the National Institute of Mental Health (NIMH) Genetics Initiative (410 families; 1342 subjects) and the National Repository of Research on Alzheimer's disease (NCRAD, 331 families; 1099 subjects), using the Affymetrix Genome-Wide Human SNP Array 6.0 which includes ~900K genomic SNPs. We performed a genome-wide SNP association analysis using PBAT/FBAT. We report p-values from the Liptak test statistic, which utilizes all available information. The Liptak method attains higher power levels, than the traditional FBAT approach, by combining the Z-statistics that correspond to the p-values of the family-based test (the within family information) with the rank-based p-values for population-based analysis (the between family information). We identified a nucleoporin 153kDa (*NUP153*) missense SNP on chromosome 6 with genome-wide significant evidence for association (Liptak $p = 4 \times 10^{-8}$) with LOAD in the NCRAD sample. This association with *NUP153* has not previously been reported with LOAD. Replication of the *NUP153* signal was achieved (FBAT p = 0.005) in an independent (NIA-LOAD) family sample (503 families; 1764 subjects), but was not confirmed in the NIMH sample. Additionally, meta-analysis, combining p-values from NCRAD, NIMH, NIA-LOAD, GenADA, and TGEN2, identified other genes containing SNPs with suggestive evidence of association ($p < 0.001$) with LOAD: *POU2F3*, *IL1RAP*, and *NOTCH2*. This study represents an extensive genome-wide screen for SNPs associated with susceptibility to LOAD. *NUP153* is involved in nuclear import and export of proteins and RNA. Recently, *NUP153* has been implicated in guarding genome integrity by promoting nuclear import of 53BP1. This novel association of a *NUP153*missense variant with risk for LOAD may provide valuable pathogenetic clues regarding the etiology and pathogenesis of AD.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2521T

Beta-Amyloid Toxicity Modifier Genes and the Risk of Alzheimer's Disease. *S.L. Rosenthal[1], X. Wang[1], M.M. Barmada[1], F.Y. Demirci[1], O.L. Lopez[2, 3], M.I. Kamboh[1,2].* 1) Department of Human Genetics, University of Pittsburgh, Pittsburgh, PA; 2) Alzheimer's Disease Research Center, University of Pittsburgh, Pittsburgh, PA; 3) Department of Neurology, University of Pittsburgh, Pittsburgh, PA.

Late-onset Alzheimer's disease (LOAD) is a neurodegenerative disease, affecting 1 in 8 people over age 65 in the United States. Risk of developing this complex disease has been repeatedly linked to *APOE* polymorphisms, while other genes including *PICALM, CLU, BINI1, CD2AP, CR1, CD33, EPHA1, ABCA7,* and *MS4A4A/MS4A6E* have been implicated in more recent genome-wide association studies (GWAS). Aβ plaques develop in the brains of LOAD patients and are considered to be a pathological hallmark of this disease. Recently 12 new Aβtoxicity modifier genes (*ADSSL1, PICALM, SH3KBP1, XRN1, SNX8, PPP2R5C, FBXL2, MAP2K4, SYNJ1, RABGEF1, POMT2,* and *XPO1*) have been identified (*Science* 2011; 334: 1212-3) that potentially play a role in LOAD risk. In this study, we have examined the association of 222 SNPs in these 12 candidate genes with LOAD risk in 1291 Caucasian American LOAD cases and 938 age-matched, cognitively normal Caucasian American controls. Single site and haplotype analyses were performed using PLINK. Following adjustment for *APOE* genotype, age, sex, and principal components, we found single nucleotide polymorphisms (SNPs) in the *PPP2R5C, PICALM, SH3KBP1, XRN1,* and *SNX8* genes that showed significant associations (P<0.05) with the risk of LOAD. The top SNP was located in intron 3 of *PPP2R5C* (P=0.009017), followed by an intron 19 SNP in *PICALM* (P=0.0102). Haplotype analysis revealed significant associations (P<0.05) in 5 genes (*ADSSL1, PICALM, PPP2R5C, SNX8*, and *SH3KBP1*). Our data indicate that genetic variation in these new candidate genes affects the risk of LOAD. Further investigation of these genes, including additional replication in other case-control samples and functional studies to elucidate the pathways by which they affect Aβ, are necessary to determine the degree of involvement these genes have for LOAD risk.

2522F

WTCCC3 and GCAN: A genomewide association scan of Anorexia nervosa. *V. Boraska[1,2], C.M. Bulik[3], D.A. Collier[4], P.F. Sullivan[3], E. Zeggini[1], Genetic Consortium for Anorexia Nervosa, Wellcome Trust Case Control Consortium 3.* 1) Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, United Kingdom; 2) Medical School University of Split, Split, Croatia; 3) University of North Carolina at Chapel Hill, Chapel Hill, NC, United States; 4) Kings College London, London, United Kingdom.

Anorexia nervosa (AN) is a debilitating and potentially life-threatening mental illness. The Genetic Consortium for Anorexia Nervosa (GCAN) is a 20-country consortium dedicated to identifying genes that influence AN. As part of the Wellcome Trust Case Control Consortium 3 (WTCCC3), we conducted the largest genome-wide association study (GWAS) for any eating disorder ever performed. A GWAS was performed to identify unequivocal evidence of associations between DNA sequence variations and AN. All cases met DSM IV diagnostic criteria for AN (excluding amenorrhea) and were genotyped using the Illumina 660W-Quad array. Genotyping was completed on 2907 AN cases and 14,860 controls originating from 15 different countries of European ancestry. Controls were carefully selected to match for ancestry of each AN site. Individual association analyses were carried out across 15 strata. Eleven association analyses were conducted per stratum: one with no principal components (PC) adjustment and the others adjusted for up to 10 PCs. On the basis of QQ plots, lambda values and 54 population stratification SNPs, we selected the best analysis per each stratum for the fixed-effects meta-analysis. Only one stratum needed an adjustment for PC1. The genome-wide significance threshold was set to 5×10-8. One SNP on chr5 exceeded genome-wide significance. An additional 94 SNPs, derived from 65 independent signals, had p-values < 10-4. Most of these SNPs are in the frequency and effect size classes expected for common diseases. Replication of SNPs, representing each independent signal, is currently underway in 2400 samples of European origin and 900 Japanese samples. For strong/replicated findings, GCAN is well positioned to immediately pursue genomic, transcriptomic, molecular biological, and epidemiological studies in humans. Cross disorder analyses with psychiatric disorders and obesity are planned. Furthermore, functional analysis of human anorexia genes in mice will provide novel insights in the neurobiological mechanisms underlying this perplexing psychiatric disorder.

2523W

Novel approach to genome wide association meta-analysis for bipolar disorder reveals network of calcium channel and calcium signaling genes including epistatic effects. *V.J. Vieland[1], K.A. Walters[1], K. Tobin[2], M.A. Azaro[2], A. Hare[2], Y. Huang[1], S-C. Seok[1], T. Lehner[3], L.M. Brzustowicz[2].* 1) Battelle Ctr Math Med, Res Inst Nationwide Child Hosp, Columbus, OH; 2) Dept of Genetics, Rutgers Univ, Piscataway, NJ; 3) National Institute of Mental Health, Bethesda, MD.

A small number of calcium channel genes have previously been implicated in bipolar disorder (BP) and other psychiatric disorders; in the largest GWAS to date [Sklar et al. Nat Genet 2011], the smallest p-value occurred for a SNP in the voltage-gated calcium channel gene (V-GCCG) *CACNA1C*. We initially analyzed BP GWAS data from GAIN (N=390 cases, 890 controls, European ancestry (EA) only) combined with BP family data (N=306 multiplex families) from the NIMH repository, for combined association-linkage analysis under the PPL framework. Among SNPs with posterior probability of linkage and trait-marker LD (PPLD) > 5%, or the top 188 SNPs (0.02% of total), were three of four genes in the $\alpha 2\delta$ subunit of another V-GCCG (*CACNA2D4*, *CACNA2D3*, *CACNA2D1*, PPLD=95%, 7%, 6% respectively). Because the PPLD measures evidence against as well as in favor of association with a small prior probability of association (0.0004), these signals all stand out above background noise. Pathway analysis in IPA showed calcium signaling to be the highest scoring Canonical Pathway (-log (p-value)=3.8) among the genes implicated by the 188 SNPs. We then undertook genome-wide meta-analysis including an additional 6 data sets from the PGC (total N=3328 cases, 4017 controls, EA only). The PPL framework parameterizes an underlying genetic model allowing for dominant, recessive, or intermediate inheritance, variable allele frequencies and heterogeneity, and integrates over these parameters for a model-free analysis *within* each set of data; results are then accumulated *across* data sets using Bayesian sequential updating. This allows for differences in underlying trait parameters and levels of allelic association between studies, and can yield different rank ordering of SNPs compared to standard “mega”- or meta-analysis. We corroborated the salience of calcium signaling in the full PGC data set (IPA -log(p-value)= 3.2) as well as associations to the 3 $\alpha 2\delta$ V-GCCGs, and also picked up *CACNA1B*, *CACNA1A*, *CACNA1S*,*CACNB2* (PPLDs = 24%, 13%, 14%, 19%); results indicated considerable heterogeneity between data sets. In an initial set of two-locus analyses, we detect evidence of epistastic effects on the BP phenotype between the calcium-dependent activator of secretion gene *CADPS*, a calcium binding protein that regulates exocytosis of synaptic vesicles in neurons, and *CACNA2D1*,*CACNA2D4* and *CACNA1A*, respectively. Analyses to detect additional BP-related interactions within this gene family are ongoing.

2524T

Constructing a prediction model for bipolar disorder using genome-wide association dataset. *L. Chuang, C. Kao, P. Kuo.* Institute of Epidemiology and Preventive Medicine, Taipei, Taiwan.

Genome-wide association (GWA) data have provided large amount of genetic information for exploring pathogenesis of complex traits. Bipolar disorder is one of the common psychiatric disorders with high heritability around 0.6 to 0.8. So far, the determination of bipolar disorder diagnosis still largely depends on subjective report of symptoms, rather than valid biological or laboratory tests. Recently, machine learning approaches are popular in association studies to reveal not only marginal effects but also gene-gene or higher dimensional interactions for the risk of genetic variants on trait of interest. The current study aimed to construct and validate a genetic prediction model for bipolar disorder using machine learning approaches in large-scale GWA datasets. We used the GWA bipolar data from the Genetic Association Information Network. Random Forest algorithm was applied to calculate the importance of each genetic variant for bipolar disorder with two indexes, conditional importance and Gini impurity. We selected the top 150 of union of the two indexes. Markers were then evaluated in logistic regression with step-wise selection to identify the final set of 135 genetic variants to construct the prediction model for bipolar disorder. The area under the receiver operating characteristic curve of our predictive model reached 0.932 (95% confidence interval, 0.922–0.942). The optimal cut-off value for corresponding predictive score was 130, with a high sensitivity (0.836) and specificity (0.865) to distinguish bipolar disorder patients from controls. After conducting leave-one out cross-validation, the error rates of discrimination in healthy controls and bipolar disorder patients were 0.135 and 0.164, respectively. In conclusion, using GWA data with random forest procedures, we constructed a prediction model for bipolar disorder. With further validation and replication, such application may eventually benefit clinical diagnosis to provide a reliable and rapid screening for the risk of developing bipolar disorder.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2525F

Involvement of the PTCHD1 gene in autism and intellectual disability. *K. Mittal[1], K. Sritharan[1], J. Ross[2], A. Vaags[3], S. Scherer[3], A. Noor[4], J. Ellis[2], J. Vincent[1].* 1) Molecular Neuropsychiatry and Development Laboratory, Centre for Addiction and Mental Health, Toronto, Ontario, Canada; 2) Developmental & Stem Cell Biology Program The Hospital for Sick Children, MaRS Centre, TMDT Toronto, Ontario, Canada; 3) The Centre for Applied Genomics The Hospital for Sick Children, MaRS Centre, TMDT Toronto, Ontario, Canada; 4) Department of pediatrics and laboratory medicine, The Hospital for Sick Children, Toronto, Ontario, Canada.

A number of candidate genes for autism spectrum disorder have been identified through recent copy number variant (CNV) studies. Our lab reported CNVs in the X-chromosome PTCHD1 (patched-related) gene, in autism spectrum disorder (ASD) and in intellectual disability (ID) (Noor et al., 2010). Since then, several additional families with exonic deletions have been identified. There are also additional reports with upstream deletions of PTCHD1. These new CNVs will be discussed. Systematic sequencing of PTCHD1 among 900 ASD probands and 225 ID patients showed 5 missense changes in ASD and two missense changes in ID. However, the presence of several of these mutations in the 1000 genomes and NHLBI exomes datasets provides evidence against the involvement of some of them. An additional 500 probands from autism and Intellectual disability are being sequenced for additional mutation identification. Human PTCHD1 has 3 exons, however the UCSC Genome Browser suggests two different splice variants. The large isoform resulting in a 888 amino acid protein (PTCHD1a) and the smaller isoform from just exons 1 and 2, resulting in a 361 amino acid protein (PTCHD1b), however we were unable to find any supporting evidence for PTCHD1b, but instead found evidence of a cerebellar isoform of 542 amino acids resulting from exclusion of exon 2 (PTCHD1c), as well as evidence of an upstream exon. The presence of the 888 and 542 amino acid isoforms appears to be supported by western blot data. We have now developed induced pluripotent stem cells from members of the 167Kb deletion family (Noor et al., 2010)-these splice variants were present in iPS cell-derived neuronal precursor cells from the mother's wild type allele, but not from the proband. Co-localization of PTCHD1 in neuronal cells and in cilia in 10T1/2 cells is also currently under investigation. A full update will be presented.

2526W

Genetic schizophrenia risk variants jointly modulate total brain and white matter volume. *A.F. Terwisscha van Scheltinga[1], S.C. Bakker[1], N.E.M. van Haren[1], E.M. Derks[2], J.E. Buizer-Voskamp[1], H.B.M. Boos[1], W. Cahn[1], H.E. Hulshoff Pol[1], S. Ripke[3], R.A. Ophoff[1,4], R.S. Kahn[1], Psychiatric GWAS Consortium.* 1) University Medical Centre Utrecht, Utrecht, Utrecht, Netherlands; 2) Department of Psychiatry, Academic Medical Centre, University of Amsterdam, Amsterdam, the Netherlands; 3) Center for Human Genetic Research, Massachusetts General Hospital, Boston, United States of America; 4) Center for Neurobehavioral Genetics, University of California, Los Angeles, United States of America.

Background Thousands of common single nucleotide polymorphisms (SNPs) are weakly associated with schizophrenia. It is likely that subsets of disease-associated SNPs are associated with distinct heritable disease-associated phenotypes. Therefore, we examined the shared genetic susceptibility modulating schizophrenia and brain volume.

Methods Odds ratios for genome-wide SNP data were calculated in the sample collected by the Psychiatric GWAS Consortium (8,690 schizophrenia patients and 11,831 controls, excluding subjects from the present study). These were used to calculate individual polygenic schizophrenia (“risk”) scores (PSSs) in an independent sample of 152 schizophrenia patients and 142 healthy controls with available structural MRI scans.

Results In the entire group, the PSS was significantly associated with total brain volume (R2=0.048, p=1.6×10-4) and white matter volume (R2= 0.051, p=8.6×10-5) equally in patients and controls. The number of (independent) SNPs that substantially influenced both disease risk and white matter (n=2,020) was much smaller than the entire set of SNPs that modulated disease status (n=14,751). From the set of 2,020 SNPs, a group of 186 SNPs showed most evidence for association with white matter volume and an explorative functional analysis showed that these SNPs were located in genes with neuronal functions.

Conclusions These results indicate that a relatively small subset of schizophrenia genetic risk variants is related to the (normal) development of white matter. This in turn suggests that disruptions in white matter growth increase the susceptibility to develop schizophrenia.

2527T

PET scans with [^{11}C]Flumazenil demonstrate a downregulation of the $GABA_A$ receptor availability in fragile X syndrome patients. *F. Kooy[1], I. Heulens[1], C. D'Hulst[1], N. Van der Aa[1], K. Van Laere[2].* 1) Dept Medical Genetics, Univ Antwerp, Antwerp, Belgium; 2) Div. Nuclear Medicine, University Hospital and Katholieke Universiteit Leuven, Leuven, Belgium.

Expansion of a CGG repeat at the FRAXA site silences the expression of the FMR1 gene and so causes the fragile X syndrome, a frequent form of inherited mental impairment. In addition to a significantly lower IQ, patients are characterized by specific dysmorphic features and behavioral abnormalities including increased anxiety, hyperactivity and autistic-like features. Epileptic seizures are common. Over the last 2 decades, the function of the causative gene has been studied intensively. The gene appears involved in diverse cellular processes, including RNA transport and stability, protein translation and RNA interference. Of particular interest from a therapeutic perspective is the observation that the imbalance between the inhibitory and excitatory signaling is compromised in the disorder and that this aberrant signaling can explain many, if not all clinical symptoms in patients. Genetic rescue experiments and drug trials with glutamatergic and gabaergic drugs in animal models have convincingly shown that a correction of the imbalance is able to revert some of the symptoms. Trial in patients have shown a potential improvement in condition only in the most severely affected group of patients and, so far, no direct evidence for abnormalities in either the glutamatergic or the gabaergic system in patients has been provided. Using positron emission tomography (PET) with [^{11}C]Flumazenil, a known $GABA_A$ receptor antagonist, as a radioligand we determined the amount of $GABA_A$ receptors in the brain of 10 fully characterized fragile X patients and 10 controls. We found an overall 6–17% decrease in $GABA_A$ receptor availability, with was significant in most brain regions. Examining the data of individual patients, it is striking that a reduction is present in only about half of our patients. Our data confirm that the GABAergic system is involved in the pathophysiology of the fragile X syndrome and provide a possible explanation why some patients respond to therapy while others do not.

2528F

δ-catenin is Genetically and Biologically Associated With Cortical Cataract and Future Alzheimer-Related Structural and Functional Brain Changes. *G. Jun[1,2], J.A. Moncaster[3], C. Koutras[9], S. Seshadri[4,7], J. Buros[1], A.C. McKee[4,5,8,10], G. Levesque[11], P.A. Wolf[4,6,7], P.St. George-Hyslop[9,12], L.E. Goldstein[3,4,5,8], L.A. Farrer[1,2,4,6,8].* 1) Medicine (Biomedical Genetics Section), Boston University, Boston, MA; 2) Ophthalmology, Biostatistics, Boston University, Boston, MA; 3) Psychiatry, Boston University, Boston, MA; 4) Neurology, Boston University, Boston, MA; 5) Pathology & Laboratory Medicine, Boston University, Boston, MA; 6) Epidemiology, Boston University, Boston, MA; 7) Framingham Heart Study, Boston, MA; 8) Alzheimer's Disease Center, Boston University, Boston, MA; 9) Tanz Centre for Research in Neurodegenerative Diseases, University of Toronto, Ontario, Canada; 10) Geriatric Research Education Clinical Center, Bedford Veterans Administration Hospital, Bedford, MA; 11) Neurosciences Research Centre-CHUL, Universite Laval, Quebec, Canada; 12) Cambridge Institute for Medical Research, University of Cambridge, Cambridge, UK.

Multiple lines of evidence suggest that specific subtypes of age-related cataract (ARC) and Alzheimer disease (AD) are related etiologically. To identify shared genetic factors for ARC and AD, we estimated co-heritability of quantitative measures of cataract subtypes with AD-related brain MRI traits among 1,466 participants of the Framingham Study who had a brain MRI scan approximately ten years after the eye exam. Cortical cataract (CC), but not other cataract subtypes, was found to be co-heritable with future development of AD and with several MRI traits, especially temporal horn volume (THV, ρ=0.24, $P<10^{-4}$). Based on the hypothesis that genetic variants with pleiotrophic effects for CC and THV can be identified in preclinical stages of brain dysfunction, we conducted a genome-wide association study using mid-age dementia-free subjects from the Framingham Offspring Study (N=1,249) whose mean ages at the eye exam and MRI exams were 51 and 61 years, respectively. A genome-wide study testing association of 187,657 single nucleotide polymorphisms (SNPs) with the joint outcome of CC and THV identified genome-wide significant association with CTNND2 SNPs rs17183619, rs13155993 and rs13170756 ($P<2.6\times10^{-7}$). These SNPs were also significantly associated with joint outcomes of CC and scores on several highly heritable neuropsychological tests ($5.7\times10^{-9} < P < 3.7\times10^{-6}$). Statistical interaction was demonstrated between rs17183619 and APP SNP rs2096488 on CC (P=0.0015) and CC-THV (P=0.038). A rare CTNND2 missense mutation (G810R) 249 base pairs from rs17183619 altered δ-catenin localization and increased secreted amyloid-β1-42 in neuronal cell culture. Immunohistopathological analysis of lens tissue obtained from two autopsy-confirmed AD subjects and two non-AD controls revealed elevated expression of δ-catenin in epithelial and cortical regions of lenses from AD subjects compared to controls. Our findings suggest that genetic variation in delta catenin may underlie both cortical lens opacities in mid-life and subsequent MRI and cognitive changes that presage the development of AD.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2529W

Genes, Brain and Cognition in Down Syndrome. *J.R. Korenberg[1], L. Dai[1], S. Gouttard[2], J. Edgin[3], L. Nadel[3], E.K. Jeong[4], J. Anderson[5], G. Gerig[2].* 1) Center for Integrated Neuroscience and Human Behavior, and Department of Pediatrics, University of Utah, Salt Lake City, UT 84112, USA; 2) School of Computing, University of Utah, Salt Lake City, UT; 3) Department of Psychiatry, University of Arizona, Tucson, AZ; 4) UCAIR, School of Medicine, University of Utah, Salt Lake City, UT; 5) Department of Neuroradiology, University of Utah, Salt Lake City, UT.

Down syndrome (DS) is the prototype of human aneuploidy and a major concern for the development of novel therapeutics. Despite decades of research, no genes have been linked to specific features, and no treatment for the cognitive deficits has been developed. Treatment of DS requires multidisciplinary studies linking the genetic, neuroanatomical (MRI), neural circuitry (DTI), and neurocognitive features of DS. Data on MRI/DTI for 12 DS and 9 age, gender matched controls has been collected and analyzed. MRI results showed that white and grey matter are proportionately decreased in non-normalized IC volumes in DS. In an anterior-posterior gradient, white is more decreased than gray in all lobes. In addition, DTI tractography shows first evidence for abnormal development of motor tracts and arcuate fasciculus, substrates for visuo-motor cognitive functions and language. A significant part of this appears to be not just numbers of axons but actual diffusion both in the direction of axonal flow and radially, implicating a moving target, myelination. Therefore, even if white matter volumes are decreased early in development, there appear to be in the adult, additional myelination defects that are ongoing and approachable. Finally, neuropsychological and IQ data are collected for 12 DS and correlated with their MRI data. Our pilot results suggest that DS performance on the paired associates learning test (CANTAB; total errors and mean trials to success) is correlated with decreased normalized white matter. Abnormal intrinsic connectivity networks in DS revealed significant increases in local (regional homogeneity) greatest in limbic and deep gray nuclei and in regions less than 4 cm distant. In contrast long range connections showed reduced functional correlation in DS. These findings suggest that white matter is possibly a new therapeutic focus for DS that may in part be reflected in decreased long range connectivity and therefore abnormal interactions between default and attention networks, in part accounting for executive deficits and impaired processing speed in DS. These results provide unprecedented characterization of the neural and genetic systems disturbed in DS, and will result in novel targets and clinical assays to develop therapeutics to improve cognitive function in DS.

2530T

The non-invasive quantification of muscle degeneration in facioscapulohumeral muscular dystrophy with fast bound-pool fraction imaging: a 3.0T MRI study. *H.R. Underhill[1], S.D. Friedman[2], V.L. Yarnykh[3], D.W.W. Shaw[2].* 1) Division of Medical Genetics, University of Washington, Seattle, WA; 2) Department of Radiology, Seattle Children's Hospital, Seattle, WA; 3) Department of Radiology, University of Washington, Seattle, WA.

Facioscapulohumeral muscular dystrophy (FSHD) is an autosomal dominant muscular dystrophy with progressive muscle degeneration, atrophy and fatty replacement. Selective muscle involvement in FSHD complicates clinical evaluation. Magnetic resonance imaging (MRI) studies have non-invasively measured fatty infiltration as a possible surrogate for muscle deterioration; however, fatty infiltration is a late stage and may be beyond intervention. Edema has also been studied, but it is not clear the extent or consistency of edema in the disease process. We hypothesized that fast bound-pool fraction imaging (FBFI), a new quantitative MRI technique that directly measures alterations in tissue structural organization, may be sensitive to an earlier pathophysiologic stage in FSHD. FBFI exploits the effects of magnetization transfer (MT) to specifically measure the fraction of protons bound to macromolecules (i.e. the bound-pool). In addition, lipids do not contribute to MT thereby giving FBFI the potential to isolate alterations in diseased muscle in FSHD independent of the signal from fat. We first developed and optimized FBFI on a 3.0T clinical scanner using *in vivo* leg muscles in a healthy control and an FSHD patient. Subsequently, 6 FSHD and 8 healthy controls were imaged. Sequences used to calculate the pixel-based bound-pool fraction included: MT images (SPGR, TR/TE=45/6 ms) acquired with variable off-resonance saturation pulses (=4, 12, 96 kHz) and a nominal flip angle of 950°; B_0 and B_1 maps to account for magnetic- and RF-field non-uniformity, respectively; and R_1 maps via the variable flip-angle method. In FSHD, muscle groups without fatty infiltration were isolated, which identified 2 subgroups: 1) healthy appearing muscle with a similar bound-pool fraction as seen in controls (10.3±0.5% vs. 10.4±0.3%, p=0.68); and 2) muscle with a significantly reduced and more variable bound-pool fraction compared to controls (6.4±1.3% vs. 10.4±0.3%, p=0.001). In both normal and diseased muscle there was no correlation between the bound-pool fraction and signal intensity on STIR (TR/TE/TI=4200/60/210 ms), a qualitative MRI technique used to identify edema. These findings offer evidence that FBFI is identifying changes in FSHD muscle before fatty infiltration is seen and not directly related to edematous changes. FBFI may allow direct, quantitative monitoring of muscle deterioration in FSHD and provide a new biomarker for disease progression and/or response to therapy.

2531F

Neural Cell Adhesion Gene Variation and Brain Morphometry in Alzheimer's Disease. *K.N. Holohan[1,2], S.L. Risacher[2], S. Swaminathan[1,2], J.D. West[2], M. Inlow[3], S.K. Conroy[2], V. Ramanan[1,2], T. Foroud[1], L. Shen[2], A.J. Saykin[1,2], Alzheimer's Disease Neuroimaging Initiative (ADNI).* 1) Medical and Molecular Genetics, Indiana University-Purdue University Indianapolis, Indianapolis, IN; 2) Center for Neuroimaging, Indiana University-Purdue University Indianapolis, Indianapolis, IN; 3) Rose-Hulman Institute of Technology, Terre Haute, IN.

Two genes coding for neural cell adhesion molecules, *CDH8* and *NXPH1*, have been associated with magnetic resonance imaging (MRI) measures of hippocampal volume or grey matter density change, which are both biomarkers of Alzheimer's disease (AD). Polymorphisms in neural cell adhesion genes could provide protection against volume loss via more effective adhesion, or accelerate loss due to weaker adhesion, as well as influence other cellular processes. Therefore, we hypothesized that neural cell adhesion pathway gene variants influence risk of developing AD and the phenotypic expression of atrophy on structural MRI. Participants included Caucasian individuals from the Alzheimer's Disease Neuroimaging Initiative (ADNI), with available MRI and genome-wide association study data (N= 737). Age, gender, *APOE* e4 allele carrier status, and total intracranial volume were included as covariates. Based on previous results and literature data-mining, we compiled lists of candidate genes, one including interacting cadherins with cerebral expression (16 genes), and one containing other interacting neural cell adhesion genes (13 genes). Single nucleotide polymorphisms (SNPs) within 2,000 base pairs of selected genes with minor allele frequencies greater than ten percent were analyzed in PLINK in gene level sets for association with hippocampal volume. One cadherin gene, *CDH19*, survived Bonferroni correction ($p<0.003125$); this gene has not been previously associated with any disease. The SNP driving this signal was rs508987 (p=0.0003). Whole brain voxel-based analysis demonstrated a significant association in the right putamen ($p<0.005$, 50 voxel extent threshold). Individuals with C alleles had higher gray matter density in this area than individuals with two T alleles, possibly indicating a protective effect. Another ongoing study in our lab using pathway analysis of ADNI individuals with psychometric memory scores implicated neural cell adhesion. Since this analysis did not include *CDH19*, we analyzed rs508987 and found that it was significantly associated (p=0.0167) with memory score change from baseline to six months, though interestingly it was not correlated with the baseline score. This may indicate that hippocampal volume is a more sensitive indicator of neural cell adhesion-related pathology than neuropsychological testing. Currently we are working to replicate these results in another study cohort; preliminary results are promising.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2532W

Genetic interactions in focal adhesion and extracellular matrix pathways are associated with increase in ventricle size over time in the Alzheimer's Disease Neuroimaging Initiative cohort. *M. Koran[1,2], S. Meda[3], M. Sivley[1], T. Thornton-Wells[1].* 1) Center for Human Genetics and Research, Department of Molecular Physiology & Biophysics, Vanderbilt University, Nashville, TN; 2) Medical Scientist Training Program, Ph.D. Program in Human Genetics, Vanderbilt University School of Medicine, Nashville, TN; 3) Olin Neuropsychiatry Research Center, Harford Hospital, Hartford, CT.

Recent genetic studies in late onset Alzheimer disease (LOAD) have led to the discovery of genetic risk factors but the investigation of gene-gene interactions (GxG) has been limited. Genetic studies often use binary disease status as phenotype but studies of complex disease require rich phenotypic information that can be mapped to distinct genetic etiologies, which may involve GxG or gene-environment interactions. For brain-based diseases, neuroimaging data can provide such quantitative traits (QTs). Many QTs in LOAD have been shown to correlate with disease status and to have greater sensitivity in detecting early pathology. In this study, we tested for association of GxG with AD-related longitudinal MRI measures from the Alzheimer's Disease Neuroimaging Initiative (ADNI) cohort. We applied a pathway-driven approach in which only pairs of SNPs from genes belonging to a common biological pathway were tested. We used linear regression in INTERSNP software to test the ability of 151 million SNP-SNP models to predict 12-, 24- and 36-month annual atrophy rates of 15 different brain regions implicated in AD structural pathology. The most significant results were seen in the right and left inferior lateral ventricles (RILV & LILV). 575 significant SNP-SNP interactions were identified in the RILV and 316 were seen in the LILV; 7 SNP-SNP interactions replicated over two of the three time points in the RILV. Four GxG replicated across time and in both the RILV and LILV: SLIT3-EPHB1, and SV2C, FN1, and ITGA1 each interacting with LAMC1. Three of these 4 GxG replicates belonged to the extracellular matrix (ECM) receptor interaction KEGG pathway, and 2 of the 4 belonged to the focal adhesion (FA) KEGG pathway. Using a pathway based approach for GxG associated with AD related QTs, we were able to identify interactions that not only replicated over time but also over location. We then were able to use current knowledge of pathway interactions to deduce a plausible biological context for these significant interactions. There is a theory in AD that neuron loss is mediated through FA signaling. Proteins involved in FA are activated by β-amyloid and communicate with the ECM to induce cell cycle activation, which activates the loss of cell adhesions and leads to subsequent cell death. The ECM and FA pathways were implicated in AD-related neuron loss seen as MRI measurement changes in this study, supporting the relationship between FA and neuron loss in AD etiology.

2533T

Gene Expression of GABA Receptors and Regulators in Peripheral Blood of a Drug Naïve First-Episode Psychosis sample. *V.K. Ota[1, 2], C.S. Noto[2], A. Gadelha[2], M.L. Santoro[1,2], J.J. Mari[2], M.I. Melaragno[1], M.A.C. Smith[1], Q. Cordeiro[3], R.A. Bressan[2], S.I. Belangero[1,2].* 1) Morphology and Genetics, UNIFESP, Sao Paulo, Sao Paulo, Brazil; 2) LiNC - Laboratorio Interdisciplinar de Neurociências Clínicas, UNIFESP, Sao Paulo, Sao Paulo, Brazil; 3) Departamento de Psiquiatria, Irmandade da Santa Casa de Misericórdia de São Paulo, Sao Paulo, Sao Paulo, Brazil.

BACKGROUND: Schizophrenia is a severe mental health disorder with a high heritability. The investigation of patients during their first episode of psychosis (FEP), particularly before taking any antipsychotic medication, opens avenues for understanding this clinically and genetically complex disease and avoiding harmful progression. The study of gene expression levels in peripheral blood may yield biomarkers for mental disorders, which will be of great value because brain tissue is not readily accessible. Since cortical dysfunction in schizophrenia and related disease may be associated with changes in GABAergic circuitry, in this study we aimed to compare the expression levels of GABA receptors and regulatory genes, in whole blood of eight drug-naïve FEP patients and seven healthy controls. METHODS: Whole blood was drawn from patients and controls during clinical evaluations into tubes containing RNA stabilizer solution Expression levels of GABRA1, GABRA2, GABRA3, GABRB1, GABRB2, GABRD, GABRE, GABRG1, GABRG2, GABRP, GABRQ, GABRR1, GABRR2 and GAD1 were assessed with the Human Neurotransmitter Receptors and Regulators RT2 Profiler™ PCR Array System, which is based on SYBR Green detection. For data analysis, we compared 2-Ct values between FEP and control group using Wilcoxon test. RESULTS: Seven of the 14 genes (GABRA2, GABRA3, GABRB2, GABRE, GABRG1, GABRP, and GABRR1) presented undetectable expression levels in whole blood, and, hence were excluded from the analyses. Significant downregulation in FEP of one gene was noted (GABRD: Fold regulation=-2.24; p=0.008). Moreover, GAD1 was also downregulated, but no statistical difference was detected (Fold regulation=-2.02; p=0.132). DISCUSSION: These preliminary results indicated downregulation of two genes: GABRD, which encodes a subunit of GABA-A receptor, and GAD1, the gene encoding a key enzyme for GABA synthesis (GAD67). Alterations in GAD1 mRNA in cerebral and cerebellar cortex of schizophrenia patients were reported previously. To our knowledge, this is the first study to investigate expression levels of these genes in peripheral blood of drug-naïve FEP patients, though in a small sample size. In conclusion, our preliminary findings suggest that dysregulation of expression levels of GABAergic genes may be involved in the pathophysiology of schizophrenia and may be possible biomarkers for its diagnosis.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2534F

Effect of DRD1 rs4532 polymorphism on treatment resistant schizophrenia. *L.MN. Spindola[1], V.K. Ota[1,2], A. Gadelha[2,3], A.F. Santos Filho[3], M.L. Santoro[1,2], D.M. Christofolini[1], F.T.S. Bellucco[1], J.J. Mari[3], M.I. Melaragno[1], M.A.C. Smith[1], R.A. Bressan[2,3], S.I. Belangero[1,2].* 1) Disciplina de Genética, Departamento de Morfologia e Genetica, Universidade Federal de Sao Paulo, Sao Paulo, SP., Brazil; 2) Laboratório Interdisciplinar de Neurociências Clínicas (LiNC), Departamento de Psiquiatria, Universidade Federal de Sao Paulo, Sao Paulo, Brazil; 3) Departamento de Psiquiatria, Universidade Federal de Sao Paulo, Sao Paulo, Brazil.

Schizophrenia is a complex disease with genetic and environmental factors interacting to develop the disease. Dysregulation of dopaminergic neurotransmission is involved in the pathophysiological processes of schizophrenia and, hence, some studies have investigated the role of genes encoding dopamine receptors. Positron Emission Tomography (PET) studies have been demonstrated that dopamine receptor D1 dysfunction was associated to cognitive impairments and negative symptoms in schizophrenia. Regarding therapeutics aspects, the overall treatment response is still limited and it is estimated that 20–30% of patients who have schizophrenia do not respond to treatment with conventional antipsychotics. The aim of this study was to investigate the association between DRD1 rs4532 (A/G) polymorphism and schizophrenia diagnosis and treatment resistant (TR) schizophrenia. We have analyzed 178 patients with schizophrenia, consisting of 59 TR and 65 non-TR, and 266 healthy controls. The subjects were genotyped for rs4532 polymorphism by TaqMan probe based real time PCR assay. Although we have not find association between rs4532 polymorphism and schizophrenia (p=0.690), we found an association between G-allele and TR schizophrenia (p= 0.001; OR= 2.71; 95% CI=1.48–4.96). Reinforcing this finding, for genotypes, we found a dose-response gradient with risk increasing with the number of G-allele copies and with GG-homozygous presenting a five-fold risk compared to AA-homozygous (p=0.014; OR=5.037; 95% CI= 1.39–18.26) with an intermediate result for AG-genotype (p=0.037; OR= 2.491; 95% CI=1.06–5.88), setting the common AA-genotype as reference. Our result suggests a possible role of DRD1 rs4532 polymorphism on TR schizophrenia, with G-allele being the genetic risk factor. Although the association of this polymorphism and schizophrenia is controversial, its effect on some phenotypes has been demonstrated, such as tardive diskenesia, negative symptoms and response to clozapine treatment. Therefore, rs4532 polymorphism may be a potential pharmacogenetic marker for treatment response to antipsychotic drugs, influencing directly on clinical treatment outcome.

2535W

Genetic Risk Factors (CHRNA5-CHRNA3-CHRNB4) and Smoking Cessation Among Hospitalized Patients with Acute Myocardial Infarction. *LS. Chen[1], PA. Lenzini[2], RG. Bach[3], JA. Spertus[4], LJ. Bierut[1], S. Cresci[3].* 1) Department of Psychiatry, Washington University School of Medicine, St. Louis, MO; 2) Department of Genetics, Washington University School of Medicine, St. Louis, MO; 3) Department of Medicine, Washington University School of Medicine, St. Louis, MO; 4) Mid America Heart Institute, University of Missouri - Kansas City, Kansas City, MO.

Objective: Smoking is an important, modifiable risk factor for acute myocardial infarction (MI). However, active smoking and the inability to quit are both common among patients hospitalized with acute MI. Existing research suggests that nicotinic receptor genes are associated with smoking quantity and cessation success. This study tests whether variants in these genes are associated with smoking cessation among patients hospitalized with an acute MI. Methods: Using data from the Translational Research Investigating Underlying disparities in acute Myocardial infarction Patients' Health Status study (TRIUMPH), a prospective 24-center observational registry of acute MI treatment and outcomes, we ascertained current vs. former (quit ≥30 days before presentation) smoking status during the index hospitalization. After confirming consistent genetic effects across European and African Americans (EA and AA), pooled analysis using logistic regression to model the association between smoking cessation and CHRNA5-CHRNA3-CHRNB4 variants (rs680244, rs16969968), adjusting for age, sex, race, and history of alcohol abuse. Results: In the genetic sample of TRIUMPH, 595 EA and 194 AA subjects were current smokers vs. 631 and 110, respectively, who were former smokers. Genetic variants in the CHRNA5-CHRNA3-CHRNB4 region, rs680244 (A) and rs16969968 (A), were associated with smoking cessation, but in opposite directions (OR=1.198, 95% CI=1.017–1.412, p=0.0078 and OR=0.780, 95% CI=0.649–0.937, p=0.0303, respectively). There was a borderline significant association between rs680244 (A) and smoking quantity among current smokers (β[SE]=0.083[0.05], p-value=0.0982). Conclusions: Among patients hospitalized with acute MI, those carrying high-risk genetic variants in CHRNA5-CHRNA3-CHRNB4 were less likely to have successfully quit before hospitalization. By identifying a genetic group with heightened risk for cessation failure, this work may support efforts to identify patients for enhanced, personalized smoking cessation treatment.

2536T

STXBP1 and KCNQ2 are mutated in the majority of patients having Ohtahara syndrome that does not evolve into West syndrome. *M. MILH[1,2,3], C. LACOSTE[4], J. SUTERA-SARDO[1,2,3], S. AUVIN[8], M.A. BARTHEZ[6], A. KAMINSKA[9], A. ROUBERTIE[5], D. VILLE[7], C. BADENS[2,3,4], B. CHABROL[1,2,3], L. VILLARD[1,2].* 1) Neurologie Pédiatrique, Hôpital d'Enfants La Timone, Marseille, France; 2) Aix Marseille Université, Faculté de Médecine, Marseille, France; 3) Inserm, U910, Marseille, France; 4) Dpt de Génétique Médicale et de Biologie Cellulaire, Hôpital d'Enfants La Timone, Marseille, France; 5) Neuropédiatrie, Hôpital Gui de Chauliac, Montpellier, France; 6) Neuropédiatrie, CHU de Clocheville, Tours, France; 7) Neuropédiatrie, Hôpital Femme Mère Enfant, Bron, France; 8) Neuropédiatrie, Hôpital Robert Debré, Paris, France; 9) Neuropédiatrie, Hôpital Necker, Paris, France.

Ohtahara syndrome (OS) is a rare early onset epileptic encephalopathy characterized by the association of frequent seizures (mostly epileptic spasms and partial seizures) with an EEG showing a so-called {{oq}} suppression-burst {{ea}} pattern, i.e bursts of paroxysmal activity alternating with periods of flatness of the traces. The pronostic is poor, with drug resistant epilepsy and evolution into West syndrome (Ohathara et al. 2003). Most OS cases are in relation with a brain structural abnormality (lissencephaly, hemimegalencephaly, Aicardi syndrome...), but OS also occurs in an apparently normal brain. Among this group of OS patients without any brain structural abnormality, several genetic causes have been identified : ARX mutations were the first mutations described for OS; SLC25A22 homozygous mutations have been described in two consanguineous families; de novo heterozygous mutations of STXBP1 have been found in about 10% of OS patients (Saitsu et al.); de novo mutations of KCNQ2 have been found in about 10% of patients having an early onset epileptic encephalopathy, from whom more than a half had an Ohtahara syndrome (Weckhuysen et al. 2012). We prospectively collected a cohort of 100 patients with an early onset epileptic encephalopathy, among which 38 had OS. In order to identify key clinical features and perform genotype-phenotype correlations, we screened the known genes in our cohort. We did not find any mutation in ARX or SLC25A22 while 6 patients were mutated for STXBP1 and 8 patients were mutated for KCNQ2. We have already shown that the course of the epilepsy was unexpected in patients having a mutation of STXBP1, with a remission of the epilepsy in the majority of the patients, despite the persistence of non epileptic movements disorders and severe cognitive impairment (Milh et al. 2011). Interestingly, the course of the epilepsy was similar in the patients having a mutation of KCNQ2, with a dramatic decrease in seizure frequency occurring during the first year, together with an amelioration of the EEG. In both groups, no patient evolved into West syndrome, although this has classically been described for OS patients. Selecting OS patients that do not evolve into West syndrome we found a mutation of STXBP1 or KCN2 in 14 patients. These findings indicate that STXBP1 and KCNQ2 mutations account for the majority of OS patients that do not evolve into West syndrome.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2537F

Identification of LRRK2 p.G2019S disease modifiers. *J. Trinh[1], D. Evans[1], C. Thompson[1], A. Donald[1], B. Shah[1], C. Szu Tu[1], F. Pishotta[1], C. Vilarino-Guell[1], R. Amouri[2], F. Hentati[2], R. Gibson[3], M. Farrer[1].* 1) Medical Genetics, University of British Columbia, Vancouver, BC, Canada; 2) Neurologie, Institut National de Neurologie, La Rabta, Tunis, Tunisia; 3) GlaxoSmithKline Pharmaceuticals, Research and Development, Harlow, Greenford, Hammersmith, UK.

Parkinson's disease (PD) is an insidious and progressive neurodegenerative disease. It is a global health problem that currently affects 1% of individuals over age 65. Current therapies treat motoric symptoms and not the physiological mechanisms of disease. Mutations in the gene known as leucine-rich repeat kinase 2 (LRRK2) in PD confer the highest genotypic and population attributable risk; LRRK2 p.G2019S is most frequent and found in greater than 30% of Arab-Berber patients in Tunisia. Disease penetrance among carriers is incomplete, resulting in elder asymptomatic carriers and patients with early onset of disease. We hypothesize that there are genetic factors that can modulate the effect of LRRK2 p.G2019S on disease penetrance.

We have recruited clinical data and DNA samples for 602 PD patients from Tunisia. This includes 400 patients with idiopathic PD, 60 affected probands with 70 family members, and 142 sporadic patients with LRRK2 p.G2019S. We estimated the case-control penetrance and familial penetrance of p.G2019S with Kaplan-Meier and kin-cohort analyses respectively. We also compared the clinical motor phenotypes across homozygote, heterozygote and non-LRRK2 p.G2019S carriers. We performed whole-genome-sequencing in 7 early-onset and 7 elder asymptomatic carriers to identify potential genetic modifiers of age of onset.

Disease penetrance for LRRK2 p.G2019S carriers was estimated to be 30% at age 50 and 80% at age 70. Interestingly, a significant difference was observed between genders with the average age at onset for females being 5 years younger than that observed in males in cases and families. This gender bias appears LRRK2 p.G2019S specific. Differences in age at onset between genders were not observed for idiopathic patients. WGS in 14 LRRK2 p.G2019S patients identified ~9×106 variants of which 287 coding changes discriminate between early-onset and asymptomatic groups; a subset have been selected for validation in the larger series to identify polymorphisms that may modulate the phenotypic expression of p.G2019S-linked disease.The age at onset of PD of p.G2019S carriers in this study spans 50 years despite their ethnic, genetic and environmental homogeneity. Some validated targets of WGS appear to modulate age at onset and may highlight novel therapeutic strategies to delay the onset of disease.

2538W

***APOE* and *SNCA* predict cognitive performance in Parkinson's disease.** *I. Mata[1,2], J. Leverenz[1,2], J. Trojanowski[3], A. Siderowf[3], B. Ritz[4], S. Rhodes[4], S. Factor[5], C. Wood-Siverio[5], J. Quinn[6], K. Chung[6], A. Espay[7], F. Revilla[7], K. Edwards[8], T. Montine[9], C. Zabetian[1,2].* 1) Dept Neuroscience, VA Puget Sound Healthcare Sys, Seattle, WA; 2) Dept Neurology, University of Washington, Seattle, WA; 3) University of Pennsylvania, PA; 4) UCLA, Los Angeles, CA; 5) Emory University, Atlanta, GA; 6) Portland VA Medical Center, Portland, OR; 7) University of Cincinnati, Cincinnati, OH; 8) Dept Epidemiology, University of Washington, Seattle, WA; 9) Dept. Pathology, University of Washington, Seattle, WA.

Objective: To determine if common variation in the *MAPT, SNCA,* and *APOE* genes is associated with cognitive performance in patients with Parkinson's disease (PD) Background: Cognitive impairment (CI) is an increasingly recognized problem in PD and approximately 80% of patients will develop dementia during the course of the disease. Though there are several well-established susceptibility genes for PD, whether common genetic variation influences risk for CI in PD has not yet been determined. Design and Methods: We recruited 937 PD patients from six academic centers in the U.S. All study participants underwent psychometric testing that assessed performance in four major cognitive domains: memory (HVLT), language (phonemic and semantic verbal fluency), visuospatial skills (Judgment of Line Orientation), and executive function (Letter-Number Sequencing and Trail Making Test Parts A and B). Subjects were genotyped for single nucleotide polymorphisms (SNPs) in three genes; two (*MAPT* and *SNCA*) known to alter risk for PD and one (*APOE*) reported to associate with cognitive function in PD. Linear regression was used to test for association between genotype and cognitive performance (at baseline) adjusting for sex, years of education, disease duration, and age at testing. Results: The*APOE* ε4 allele was associated with lower performance on HVLT immediate (p= 0.0002) and delayed recall (p=0.02), semantic verbal fluency (p=0.02), Letter-Number sequencing (p=0.006), and Trail Making B-A (p=0.007). The *SNCA* SNP rs356219 was correlated with performance on HVLT immediate (p=0.02) and delayed recall (p=0.02). The *MAPT* H1/H2 haplotype was not associated with scores on any tests. Conclusion: Our data indicate that *APOE* ε4 is an important predictor of cognitive performance in PD across multiple domains, and *SNCA* exhibits an effect on memory-related tasks. The latter observation suggests that *SNCA* is both a susceptibility gene for PD and modifier of cognitive function in patients with an established diagnosis of PD. These results serve as proof of principle that common genetic variation influences cognition in PD and provide a rationale for expanding future studies to a genome-wide scale.

2539T

Abnormal Responses to Visual Cortex Activation in Early Stage Huntington Disease Patients using 31P-NMR Spectroscopy. *F. Mochel[1,2,3], TM. Nguyen[1], D. Deelchand[4], D. Rinaldi[1], R. Valabregue[5], C. Wary[3,6], A. Durr[1,2], PG. Henry[4].* 1) Inserm UMR S975, Institut du Cerveau et de Moelle, Hôpital Pitié-Salpêtrière, Paris, France; 2) AP-HP, Département de Génétique, Hôpital La Salpêtrière, Paris, France; 3) Université Pierre et Marie Curie, Paris, France; 4) University of Minnesota, Center for Magnetic Resonance Research, 2021 6th Street SE, Minneapolis MN, USA; 5) CEntre de NeuroImagerie de Recherche, Institut du Cerveau et de la Moelle, Hôpital La Salpêtrière, Paris, France; 6) Laboratoire de RMN, Institut de Myologie, Hôpital La Salpêtrière, Paris, France.

Objective: To test the hypothesis that brain energy metabolism is abnormal in patients at an early stage of Huntington disease (HD). Background: Energy metabolism has been a major focus of HD research for many years due to several observations in both patients and models of the disease. However, there are currently no in vivo biomarkers of brain energy metabolism in HD. Methods: We coupled noninvasive 31P-NMR spectroscopy with activation of the occipital cortex in order to measure the levels of ATP, phosphocreatine (PCr) and inorganic phosphate (Pi) before, during and after a visual stimulus. We studied 15 HD patients at an early stage of the disease (mean motor UHDRS= 18 ±9) and 15 age- and sex-matched controls. Results: In controls, we observed an 11% increase in Pi/PCr ratio (p=0.024) and a 13% increase in Pi/ATP ratio (p=0.016) during brain activation, reflecting increased ATP synthesis and ADP levels. Subsequently, controls had a return to baseline levels during recovery (p=0.012 et 0.022 respectively). In HD patients, both Pi/PCr and Pi/ATP ratios were unchanged during and after visual stimulation, reflecting altered mitochondrial bioenergetics. In addition, in HD patients the ratio of Pi/ATP correlated with the UHDRS score during the activation (p=0.014) and recovery periods (p=0.009), while Pi/PCr ratio correlated with the UHDRS score during recovery (p=0.016), reflecting a correlation between brain energy metabolism and disease severity in HD. Conclusions: 31Phosphorus nuclear magnetic resonance spectroscopy could provide functional biomarkers of brain energy deficit to monitor therapeutic efficacy in Huntington disease.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2540F

Sympathovagal balance in Major Depressive Disorder: influence of the BDNF and 5-HTTLPR (serotonin transporter) polymorphisms. *A. Perez[1], A.R. Brunoni[2], B. Coprerski[1], D. Brunoni[1].* 1) Centro de Genetica Medica, Unifesp - EPM, Sao Paulo, Brazil; 2) Centro de Pesquisas Clínicas, 3o andar Hospital Universitário, USP.

Major depressive disorder (MDD) is a clinical condition whose core symptoms are depressive mood and anhedonia. In recent years, several biological markers have been identified for MDD such as decreased heart rate variability (HRV). Other biomarkers such as the polymorphisms of the Brain Derived Neurotrophic Factor (BDNF) and serotonin transporter (5-HTTLPR) genes have also been associated with depression. Decreased heart rate variability (HRV), the Val66Met polymorphism of the brain-derived neurotrophic factor (BDNF) gene and the short variant of the repeat length polymorphism, serotonin transported-linked polymorphic region (5-HTTLPR) have been associated in the pathophysiology of major depressive disorder (MDD) and anxiety. Nevertheless, the relationship of BDNF, 5-HTTLPR and HRV in depressed patients have not yet been explored. Therefore, we explored such relationship in 120 drug-free, moderate-to-severe patients with acute MDD. This study was approved by the Local and National Ethics Committee (protocols 873/08 and 991/10) and it is registered in the clinicaltrials.gov database (NCT01033084). We genotyped BDNF and 5-HTTLPR polymorphisms and assessed HRV in a cross-sectional study. General linear models were performed, using 4 HRV indexes (SDNN, RMSSD, LF and HF) as dependent variables and BDNF, 5-HTTLPR, age, gender, depression severity and anxiety as independent variables. Although we did not observe main effects of BDNF and 5-HTTLPR on HRV, we found specific gene × gene interactions: in fact, the Val/Val l/l and Val/Met s-carrier genotypes exhibited higher HRV whereas the Val/Met l/l and Val/Val s-carrier genotypes presented lower HRV. These effects were even more marked when controlling for co-morbid anxiety. Our results suggest that there is a genetic 5-HTTLPR × BDNF interaction that modulates sympathovagal balance in depressed patients; therefore, decreased HRV observed in MDD patients could also be explained via genetic mechanisms. Such genetic interaction can also play a role in modulating cardiac autonomic functions, and also on sympathovagal impairment in individuals on antidepressant treatment. Supported by FAPESP: 11/08960-8.

2541W

Complex genetic scores can predict smoking cessation success and uptake of addictive substance use. *G. Uhl[1], D. Sisto[1,2], D. Walther[1], W. Eaton[2], N. Ialongo[2], J. Rose[3].* 1) Dept Molec Neurobiology, NIDA/NIH, Baltimore, MD; 2) Dept of Mental Health, Johns Hopkins Bloomberg School of Public Health, Baltimore MD; 3) Center for Nicotine and Smoking Cessation Reserach, Duke University, Durham NC.

Validated genotype scores that predict relevant clinical outcomes represent major goals of complex genetics. Addiction vulnerability and abilities to quit smoking are phenotypes with substantial heritabilities, complex genetic underpinnings, few major gene influences and several modest sized GWAS datasets. We now report validation of an established, 12063SNP v1.0 smoking cessation success genotype score, developed using data from 3 smoking cessation clinical trials, in a new smoking cessation clinical trial and describe striking differences in the score in individuals who display differing developmental trajectories of use of common addictive substances. In a cessation study, v1.0 genotype scores predicted ability to quit with p = 0.00056 and area under ROC curve 0.66. About 43 vs 13% quit in the upper vs lower genotype score terciles. Latent class growth analyses of a developmentally-assessed sample identified three latent classes based on substance use. Higher v1.0 scores were associated with a) higher probabilities of participant membership in a latent class that displayed low use of common addictive substances during adolescence (p = 0.0004) and b) lower probabilities of membership in a class that reported escalating use (p = 0.001). These results suggest that: a) we have identified genetic predictors of smoking cessation success, b) genetic influences on quit success overlap with those that influence the rate at which addictive substance use is taken up during adolescence and c) individuals at genetic risk for both escalating use of addictive substances and poor abilities to quit may provide especially urgent focus for prevention efforts. (Support NIDA IRP, DA009897 (WE) and 4R37DA011796-11 (NI), private funding through JHBSPH and grant to Duke University (PI, Dr. Rose) from Philip Morris, USA for work performed prior to January, 2012).

2542T

Integrated Copy Number and Gene Expression Analysis Detects a CREB1 Association with Alzheimer's Disease. *N. Sule[1], Y. Li[2], I. Shefer[2], C. Shaw[2], S. Powell[3], B. Dawson[2], S. Zaidi[4], K. Bucasas[2], J. Lupski[2], K. Wilhelmsen[5], R. Doody[4], K. Szigeti[6].* 1) Pathology, Roswell Park Cancer Institute, Buffalo, NY; 2) Dept Molecular and Human Genetics, Baylor College of Medicine, Houston, TX; 3) Dept Pathology, Methodist Hospital, Houston, TX; 4) Dept Neurology, Baylor College of Medicine, Houston, TX; 5) University North Carolina, Chapel Hill, NC; 6) Dept Neurology, University at Buffalo, Buffalo, NY.

Background: A study to test the hypothesis that in a genetically heterogeneous disorder such as sporadic Alzheimer disease, utilizing gene expression as a quantitative trait (eQTL) and CNVs as a genetic marker map within the same individuals increases the power to detect relevant loci. Methods: Discovery cohort comprised of 24 AD and 15 neuropathologically normal control temporal lobes simultaneously typed for gene expression and CNV state. These data were integrated to perform a CNV-gene expression case-control association analysis. A replication cohort comprised of 377 AD and 193 normal control subjects from the Texas Alzheimer Research Consortium was analyzed to verify candidate findings by a genotype only case-control analysis accounting for both APOE and gender as additional covariates. Subsequently we performed differential gene expression between the deletion carrier and diploid AD cases and annotated the differentially expressed gene. We also assessed whether there is a cell type specific difference in the expression of CREB1 between the deletion carrier and diploid AD brain specimens using three CREB-1 specific antibodies. Results: Using this approach an 8 kb deletion was identified that contains a PAX6 binding site on chr2q33.3 upstream of CREB1 encoding the cAMP responsive element binding protein1 transcription factor. The association of the CNV to AD was confirmed by a case-control replication study. We identified 350 differentially expressed genes including CREB1 targets. Conclusions: Using disease state information together with gene expression as an eQTL and cis-genotype (CNV state) within the same individuals is a powerful method to detect association with disease in a genetically heterogeneous disorder. This method could serve as an alternative approach to genome-wide association studies to identify risk alleles with modest effect sizes.

2543F

Genetic modifiers of EEG brain activity during sleep in humans. *S. Warby[1], H.E. Moore IV[1], O. Carrillo[1], J. Faraco[1], L. Lin[1], P.E. Peppard[2], T. Young[2], E. Mignot[1].* 1) Center for Sleep Sciences and Medicine, Stanford University, Palo Alto, CA. 94025; 2) School of Medicine and Public Health, University of Wisconsin-Madison, Madison, WI 53726.

The brain generates complex patterns of electrical activity that can be measured on an electroencephalogram (EEG) and thought of as an individual EEG fingerprint. Sleep EEG measures the dramatic and predictable changes in basal brain activity associated with sleep stages. These progressive changes in activity over the course of the night are inherent in the neuronal circuitry and are the result of a complex dialogue between the cortex and subcortical structures. We are developing tools to measure distinct features with the EEG signal, such as sleep spindles, in order to better understand their genetic basis. Sleep spindles are a 0.5–2 second burst of cortical activity that occurs predominately in stage 2 sleep. They are most often identified by visual inspection of the EEG but can be measured by analyzing power in the sigma range (12–16Hz). Recently, spindles have been implicated in sleep arousal thresholds, brain plasticity, learning and memory consolidation. Significant alterations in spindle frequency have been observed in many disorders such as schizophrenia, epilepsy, autism, and neurodegenerative disorders. Although very little is known about sleep spindles at the genetic level, twin studies suggest that sigma power and spindle density are highly heritable traits. The purpose of this study is to perform a genome wide association study for sleep spindle characteristics in the normal human population. Sleep spindles were measured as sigma power in stage 2 sleep from 1,876 EEG studies collected longitudinally from 1,300 adult individuals from the Wisconsin Sleep Cohort. Longitudinal correlation of sigma power over 4 years within and individual is high (r2 >0.75), consistent with the idea that spindles are stable and trait-like. Genotyping was performed genome-wide using Affymetrix 6.0 genechip and direct Taqman genotyping. Genetic associations will be determined using an additive genetic model and linear regression with subject age, sex, BMI, and medications as potential covariates. We find significant effects of age and interactions with several drug categories including benzodiazepines, sedatives and blood pressure medications. Genome-wide gene association may identify several loci implicated in the regulation spindles and reveal support for the genetic control of sleep spindles in humans. Further studies are needed to replicate and extend on genetic findings and to examine gene-environment interactions.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2544W

The eQTLs in language-related genes reveal a gene × gene interaction between BDNF and PCDH17 in the human brain. *N. Li[1], S.Y. Cheong[1], M. Goode[1], S.A. Petrill[2], A.S. Bassett[3], L.M. Brzustowicz[4], C.W. Bartlett[1].* 1) The Battelle Center for Mathematical Medicine, The Research Institute at Nationwide Children's Hospital and Department of Pediatrics, The Ohio State University, Columbus, OH; 2) Department of Human Development and Family Science, The Ohio State University, Columbus, OH; 3) Department of Psychiatry, University of Toronto, Toronto CA; 4) 4Center for Molecular and Behavioral Neuroscience, Rutgers University, Newark, NJ.

Background The genetics of language, a uniquely human trait, is challenged by limited options for research in animal models. Recent studies on language impairment have replicated several genes in language related traits but finding the functional alleles that underlie the genetic signals still challenging though postmortem human brain samples may be a powerful tool for finding those genetically functional alleles. We propose that finding expression quantitative trait loci (eQTLs) in the human brain may provide strong candidate SNPs for future mechanistic studies. The eQTL data from five-replicated language related genes (ATP2C2, CMIP, CNTNAP2, FOXP2, PCDH17) were used to search expression SNPs. Additionally, brain-derived neurotrophic factor (BDNF) plays an important role in neuronal plasticity and is associated with working memory, which is crucial for language processing. A gene × gene (GxG) interaction with BDNF functional SNP (rs6265) was shown previously with PCDH17 to substantially increase risk for specific language impairment (Simmons et al 2010). We sought to determine if this interaction could be demonstrated at the gene expression level. Methods 1177 individuals from seven publicly available brain eQTL datasets were employed for analysis. SNPs from 1000 genome Phase I Integrated Release Version 2 Haplotypes were imputed to create a common marker map across studies. The posterior probability of linkage disequilibrium was used for analysis of the five language-related genes. Results An eQTL in PCDH17 was found to be consistent across by multiple probes/datasets (PPLD=49–67%), and in CNTNAP2 an eQTL was consistent across by multiple probes/datasets (PPLD=11–14%). Several additional possible eQTLs were suggested in CNTNAP2 and CMIP. PCDH17 and BDNF were found to be negatively correlated across multiple probes in both genes (-0.42 to -0.12) with p-values ranging from 0.04 to 2.2×10^-13. GxG interaction analysis of SNPs in the extended promoter region of PCDH17 (up to 100kb upstream) yielded a maximized LOD of 2.6 suggesting an interaction between BDNF genotypes and PCDH17 expression. Discussion This is the largest human brain eQTL study, derived from the existing publicly available datasets, and it has identified novel eQTLs in language related genes. SNPs discovered in this dataset represent logical targets for future studies of language related traits. Key Words: expression quantitative trait loci; language; gene-gene interaction.

2545T

The *KIAA0319* gene is associated with language progression more than its level. *S. Tezenas du Montcel[1,2], B. Feng[3], C. Barr[4], B. Falissard[2], M. Boivin[3], G. Dionne[3].* 1) ER4 Modeling in Clinical Research, University Pierre et Marie Curie-Paris 6, Paris, France; 2) INSERM U669, France; 3) Université Laval, Quebec City, Canada; 4) Toronto Western Research Institute, Canada.

Language is a complex behavior with a genetic basis in line with caregivers and stimulation from the child environment. Based on twin studies, the heritability of language skills in the preschool years is modest, around 20%, but appears more substantial at the lower end of the continuum. Over the course of development, the heritability of language skills increases to around 50%. Although we know heritability changes over time, the heritability of the slope of progression has never been tested until now. A chromosome 6 region including the *DCDC2*, *VMP*and *KIAA0319* genes has been associated with language-related skills and reading disability. In this study, we estimated the heritability of language development and the influence of chromosome 6 candidate genes on language progression with data from the Quebec Newborn Twin Study (QNTS). Participants were 1086 twins assessed on language skills at 18 and 30 months, and 5, 7 and 8 years of age. SNPs for the 3 candidate genes on chromosome 6 (*DCDC2, VMP, KIAA0319*) were genotyped for both twins. Autoregressive Latent Trajectory model adapted to twin data and candidate genes were used. Non-parametric trajectory models were also fitted and the influence of the candidate genes on the trajectories were tested. We found that, although the mean level of language was mainly due to shared environment, the heritability of language development (the slope from 18 months to age 8 years) was high (56%), with a moderate involvement of common environment (28%). In addition we identified three trajectories of language development: 1) typical language skills, stable across ages (40.3% of the sample), 2) high language skills, stable across ages (51.3%), and 3) poor language skills with a mild recovery upon school entry (8.4%). The poor language skill trajectory (versus the other two) was associated with 3 SNPs of the *KIAA0319* gene. In conclusion, the developmental slope of language progression from 18 months to age 8 years appears more heritable than differences in language skills. In addition, children with language skills below the normal range over this period recover partially upon school entry, but belonging to this group is associated with the*KIAA0319* gene which has already be shown to be associated with reading disorders.

2546F

Genome-wide association of executive function in a normally developing pediatric cohort. *J.J. Connolly[1], P.M. Sleiman[1,2,3], R.C. Gur[4], M.E. Calkins[4], D.J. Abrams[1], R.M. Chiavacci[1], H. Qiu[1], F.D. Mentch[1], R.E. Gur[4], H. Hakonarson[1,2,3].* 1) Ctr Applied Genomics, Children's Hosp Philadelphia, Philadelphia, PA; 2) Division of Human Genetics, Children's Hosp Philadelphia, Philadelphia, PA; 3) Dept Pediatrics, Perelman School of Medicine, Univ of Pennsylvania, Philadelphia PA; 4) Brain Behavior Laboratory, Perelman School of Medicine, Univ of Pennsylvania, Philadelphia PA.

We conducted a genome-wide association study (GWAS) of scores on the Penn Conditional Exclusion Test (PCET), a computerized measure of executive function that examines abstraction and concept formation. The PCET is an `odd-one-out' paradigm, requiring participants to decide which of four objects does not belong to a specific category (e.g. size, shape, and thickness of a line). A cohort of 8,500 participants aged 8–21 years were divided into discovery (the first 5000 participants) and replication sets (~3,500 participants). Scores were corrected for age and sex, and linear regression used to examine accuracy and speed. Variants at one locus on chromosome 5 survived multiple testing correction for PCET speed measures, and were also associated with speed in the replication set. These mapped to the ERRBB2IP gene. The most significantly associated SNP was a missense mutation—S1112L—at a highly conserved base. The ERBB2IP protein, commonly referred to as Erbin, has been implicated in a variety of neuronal functions, including regulation of L-type voltage-gated Ca2+ channels at the synapse, and regulation of NRG1 and signaling/myelination. Replication efforts are underway and results will be presented together with further characterization of the protein.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2547W

Lipidomic profiling in Lewy body disease brain autopsies carrying Lysosomal Storage Disorder gene variants. *N. Parmalee[1], R.B. Chan[1], N. Park[2], E.P. Cortes[3], J.P. Vonsattel[1,3], K. Marder[1,2,4], L. Honig[1,2,4], J.H. Lee[1,2], G. Di Paolo[1], L. Clark[1].* 1) Department of Pathology and Cell Biology & Taub Institute for Research on Alzheimer's Disease, Columbia Univ, New York, NY; 2) Gertrude H. Sergievsky Center, Columbia Univ, New York, NY; 3) New York Brain Bank, Columbia University, New York, NY; 4) Department of Neurology, Columbia University, New York, NY.

Objective: To determine the relationship of Glucocerebrosidase (GBA) and other lysosomal storage disease (LSD) gene variants to lipid profiles and Lewy Body (LB) and Alzheimer disease (AD) pathology. Background: Mutations in the LSD gene, glucocerebrosidase (GBA), are associated with LB disorders (LBD). We hypothesized that other LSD genes in the same biochemical pathway may also represent susceptibility genes for LBD. Four genes Glucocerebrosidase (GBA), Hexosaminidase A (HEXA), Sphingomyelin phosphodiesterase 1(SMPD1) and mucolipin 1 (MCOLN1) and their lipid substrates and products were analyzed. Methods: Brain tissue from 243 autopsies from the NYBB at CU included: 128 brains with primary neuropathological diagnoses of LB disorders with or without AD changes, 79 brains with AD without significant LB pathology, and 36 control brains with neither LB or AD pathologies. To capture rare variants (MAF<1%), all four LSD genes were sequenced in 243 autopsies. Rare variant analysis was performed using the CCRaVAT program. Liquid chromatography-mass spectrometry (LC/MS) was used to profile grey matter lipid extracts from the primary motor cortex (BA4) of brain autopsy material from a subset of samples (n=46: 18 controls, 8 AD, 8 LBD and 12 LBD with LSD variants). Quantification of a total of approximately 200 lipid species, representing 14 major lipid classes was performed. Results: Several variants that we identified in LBD subjects including, N370S (GBA), L444P (GBA), c.1278insTATC, (HEXA), c.672+30T>G (HEXA), P330R (SMPD1) and E515V (SMPD1) have been reported previously as mutations in patients with LSD. Multiple variants in GBA, SMPD1 and MCOLN1 were significantly associated with LB and LBVAD pathology (p range: 0.01 - 6.45×10-5) but not with AD. Several classes of lipids were significantly altered in LBD LSD mutation carriers compared to LBD (wildtype), controls and AD autopsy samples ($p<0.05$). In LBD autopsy samples carrying GBA mutations we observed significant changes in a number of glycosphingolipids, including GM3, monohexosylceramide and ceramide, compared to LBD without GBA mutations, controls and AD autopsy samples. Conclusion: Lipid profiling in LBD autopsy samples carrying GBA mutations (N370S, 84insGG, D409H and R463C) revealed accumulation of several lipid classes. Since α-synuclein can be degraded via the lysosome (CMA), our study suggests that lysosomal impairment may lead to accumulation of α-synuclein and LB formation.

2548T

Study of plasma metabolites levels in discordant sibling pairs with autism. *G. Malerba[1], R. Wang-Sattler[2], E. Trabetti[1], J. Adamski[3,4], P. Prandini[1], L. Xumerle[1], C. Zusi[1], A. Pasquali[1], R. Galavotti[1], P.F. Pignatti[1], T. Illig[2,5], The Italian Autism Network (ITAN).* 1) Life and Reproduction Sciences, University of Verona, Verona, Verona, Italy; 2) Research unit of Molecular Epidemiology, Helmholtz Zentrum München, 85764 Neuherberg, Germany; 3) Institute of Experimental Genetics, Genome Analysis Center, Helmholtz Zentrum München, 85764 Neuherberg, Germany; 4) Chair of Experimental Genetics, Technische Universität München, Munich, Germany; 5) Hannover Unified Biobank, Hannover Medical School, 30625 Hannover, Germany.

Autism spectrum disorders (ADS) are a group of pervasive neurodevelopmental disorders whose etiology is not known. Metabolomics might help in giving new insights about the pathogenesis of the disease. In this study, within the frame of the Italian Autism Network (ITAN, 1), plasma samples of 39 discordant ASD sibpairs, composed by an autistic male child and his unaffected brother, were measured for 188 metabolites concentration based on targeted metabolomics approach with the *AbsoluteIDQ™* p180 kit (BIOCRATES Life Sciences AG, Innsbruck, Austria). We found four metabolites concentration differed significantly between affected and non affected children (paired t-test, age-adjusted): carnitine (p=0.035), ornithine (p=0.012), hexoses (p=0.026) and PC ae C40:4 (p=0.041). Although after correction for multiple testing none of the 4 associations remained statistically significant it is noteworthy that pathways involving two of the metabolites here reported (carnitine and ornithine), whose association was significant before the adjustment, have been described associated with the autistic phenotype (2,3). In particular previous studies reported a lower plasma levels of carnitine in autistic patients as in the present study and mutations in genes involved in the synthesis of carnitine, and urea cyclic disorders associated with mental retardation and cognitive impairment. This exploratory study suggests that dysregulation of carnitine and urea cyclic disorders might be associated with ASD and that future genetic investigations should focus on these pathways in detail. Moreover the four identified metabolites might serve as candidate biormarkers for autism phenotypes. References 1) Prandini et al., 2012. The association of rs4307059 and rs35678 markers with autism spectrum disorders is replicated in Italian families. Psychiatrics Genetics doi: 10.1097/YPG.0b013e32835185c9 2) Celestino-Soper PB et al., 2012. A common X-linked inborn error of carnitine biosynthesis may be a risk factor for non-dysmorphic autism. PNAS 109:7974-81 3) Görker I, Tüzün U, 2005. Autistic-like findings associated with a urea cycle disorder in a 4-year-old girl. J Psychiatry Neurosci. 30:133-5.

2549F

Genome scan for cognitive trait loci of dyslexia: Rapid Naming and Rapid Switching of Colors, Letters and Numbers. *W.H. Raskind[1,2], K. Kevin Rubenstein[3], M. Matsushita[1], V.W. Berninger[4], E.M. Wijsman[1,3].* 1) Depts of Med; 2) Psychiatry and Behavioral Sciences; 3) Biostatistics; 4) Educational Psychology, Univ Washington, Seattle, WA.

Dyslexia is a complex learning disability with strong evidence for a genetic basis. Deficits can persist into adulthood. Strategies that may be useful for dissecting the genetic basis of dyslexia include study of component phenotypes, which may simplify the underlying genetic complexity, and an analytic approach that accounts for the multilocus nature of the trait to guide the investigation and increase power to detect loci. Here we present results of a genome scan of several measures of Rapid Automatized Naming (RAN) and Switching (RAS) as component phenotypes in a sample of 260 2–4 generation pedigrees (~2040 individuals, ~1775 with phenotypes, ~1120 genotyped) selected through a dyslexic proband. Scores from the Wolf RAN/RAS tests were used (Letters, Numbers, and Colors naming; and Letters/Numbers and Letters/Numbers/Colors Switching), after adjustment for age, age^2, cohort, and gender. We used an MCMC-based oligogenic trait model for both initial segregation analysis and for subsequent joint segregation and linkage analysis in a genome scan. Bayes' factors (BF) for linkage were used for evaluation of evidence for linkage.

Covariates explained 57–65% of the phenotypic variance, with speed for test completion slower in males than females, increasing with age in children, and decreasing again with age in adults. Correlations between adjusted trait values were high, ranging from a low of 0.55 (e.g., RAN Colors and RAN Letters) to 0.95 (e.g., RAS Letters/Numbers and RAN Letters). Oligogenic segregation analysis identified a small number of quantitative trait locus models (QTLs) that are sufficient to explain the inheritance of some of the adjusted traits. In the genome scans, the strongest linkage signals were obtained on chromosome (chr) 2 (RAN Letters, BF=59, 81 cM), chr 10 (RAN Colors, 47 cM) and chr 1 (RAS Letters/Numbers, BF=30, 63 cM). Notable signals were also obtained at two close positions on chr 12 (RAS Colors/Letters/Numbers, BF=26, 95 cM; and RAN Numbers, BF=24, 115 cM). Based on modeling studies of individuals with and without dyslexia, RAN letters is thought to be a phenotype for phonological loop and RAS is thought to be a phenotype for supervisory switching attention of working memory. Identification of the genes involved in RAN and RAS performance is therefore of great interest. Fine mapping of these 5 regions is ongoing.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2550W

Identification of genes and pathways regulated by epigenetic mechanisms in determining general intelligence (*g*) of inbred mice. *P. Cha[1], K. Kobayashi[1], Y. Ando[1], C. Yu[1], S. Yamagata[2], K. Okada[3], J. Sese[4], K. Takao[5], T. Miyakawa[6], T. Toda[1].* 1) Division of Neurology/Molecular Brain Science, Kobe University Graduate School of Medicine, Kobe, Hyogo, Japan; 2) Admission Research Unit, National Center for University Entrance Examinations; 3) Department of Psychology, Senshu University, Kanagawa, Japan; 4) Department of Computer Science, Tokyo Institute of Technology, Tokyo, Japan; 5) Center for Genetic Analysis of Behavior, Section of Behavior Patterns, National Institute for Physiological Sciences, Aichi, Japan; 6) Division of Systems Medical Science, Institute for Comprehensive Medical Science, Fujita Health University, Aichi, Japan.

To study epigenetic regulation of general intelligence (*g*), 39 inbred mice raised under similar environmental conditions were subjected to five behavioral tests. 11 parameters measuring performance of mice that reflect their cognition were analyzed by using principal component analysis. The first principal component (PC1) explains 24.4% of variance in cognition and is highly correlated with average score of mice in behavioral tests (IQ score) (R^2=0.83). Hippocampal RNAs obtained from the 5, 12, and 8 mice consistently classified as having high, medium, and low cognition, respectively, by PC1 score and IQ score were subjected to microarray gene expression analysis. 17 genes showed at least 1.3-fold difference in expression between mice with high and low cognition (FDR<0.30), including two genes (*Ldlr* and *Stxbp2*) that were down-regulated and 15 genes (*e.g. Rab39, Rapgapl1, Hist1h3d, Vcl, Osbpl6,* etc) that were up-regulated in high cognitive mice. Pathway-based analysis of the expression data revealed that most genesets (GS) enriched in high cognitive mice play important roles in protein synthesis, indicating that memory consolidation or long-term potentiation might occur actively in hippocampus of mice with high cognition. By contrast, most GS enriched in low cognitive mice play important roles in transmembrane transport, extracellular matrix and regeneration of vesicles. A number of genes in GS implicated in our study have been reported to be associated with cognitive impairments. Our findings not only provide insights into genes and pathways regulated by epigenetic mechanisms in determining "*g*", but may also deepen our understanding on various disorders associated with cognitive impairments.

2551T

Quantitative linkage analysis with the endophenotype of social responsiveness identifies genome-wide significant linkage to a narrow region at 8p21.3. *J.K. Lowe[1,2,3,7], D.M. Werling[4,7], J.N. Constantino[5,7], R.M. Cantor[3,6,7], D.H. Geschwind[1,2,3,6,7].* 1) Neurogenetics program and Department of Neurology, University of California, Los Angeles, Los Angeles, CA; 2) Neurogenetics Program and Center for Autism Research and Treatment, Semel Institute, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA; 3) Center for Neurobehavioral Genetics, Semel Institute, University of California, Los Angeles, Los Angeles, CA; 4) Interdepartmental PhD Program in Neuroscience, Brain Research Institute, University of California, Los Angeles, Los Angeles, CA; 5) Departments of Psychiatry and Pediatrics, Washington University School of Medicine, St. Louis, MO; 6) Department of Human Genetics, University of California, Los Angeles, Los Angeles, CA; 7) ACE Genetics Consortium.

Diminished social responsiveness (SR) is one of the three core deficits in autism spectrum disorder (ASD), a heterogeneous neurodevelopmental trait with a strong genetic component. We performed quantitative linkage and association analyses for SR as measured by the Social Responsiveness Scale (SRS). SRS is a 65-item questionnaire completed by a teacher or parent, assessing interpersonal behavior, communication, and repetitive/stereotypic behavior in children ages 4–18 years. Using nearly 500 families from the Autism Genetics Research Exchange (AGRE) repository, we performed a genome-wide linkage scan and identified genome-wide significant multi-point quantitative linkage to SRS on chr8p21.3 (LOD 5.16). The peak SNP (rs6587004) lies in an intronic region of *GFRA2* (glial cell line-derived neurotrophic factor (GDNF) family receptor α-2), which encodes a cell-surface receptor for neurotrophic factors involved in neuron survival and differentiation. The one-LOD interval encompassing the linkage peak spans 966 kb and contains 16 genes. We report single-marker and haplotype association findings under the linkage peak and replication studies of the linkage signal in an independent data set. Interestingly, this chr8 peak co-localizes with suggestive linkage signals for ASD as a qualitative trait in a larger set of AGRE families (n>800) as well as in a subset of families with at least one female affected offspring (n=315). These data support the use of quantitative intermediate traits such as SR to increase power by reducing heterogeneity in a complex trait such as ASD and identifies a highly significant, narrow linked region on chr8p21.3.

2552F

Nonpathogenic traits: a study of susceptibility to sympathetic yawning. *E.T. Cirulli, A.J. Bartholomew, D.B. Goldstein.* CHGV, Duke Univ, Durham, NC.

While extensive efforts continue to be made to understand the genetics of human diseases, both common and rare, there has been much less serious effort dedicated to understanding the genetic bases of traits that comprise everyday differences among healthy individuals, many of which are of great interest biologically. We propose to systematically measure everyday, non-disease phenotypes in a large cohort of healthy volunteers and to search for the genetic correlates of those phenotypes via whole-exome sequencing. Here, we have chosen to study sympathetic yawning as an example nonpathogenic trait that could offer great insight into the genetics of empathy. In controlled studies, approximately 40% of healthy volunteers will yawn when presented with a visual stimulus of someone yawning or a written account of yawning. No other human trait exhibits a mimicry response that is so difficult to control, stressing the importance of characterizing the specific biological processes that contribute to this unusual effect. These sympathetic yawns do not occur in humans until ~5 y of age and have been shown to be impaired in diseases such as schizophrenia and autism, indicating that differences in this trait between humans may be genetic and relevant to disease. We showed a 3 minute video clip of yawning stimuli to over 100 healthy volunteers for whom we have already collected extensive demographic and intelligence data; the volunteers were not observed but were asked them to click a button every time they yawned. We found that approximately 50% of the healthy volunteers yawned in response to the video, with the number of yawns ranging from 1 to 10. Surprisingly, we found a strong association for subjects with higher levels of education to be less likely to yawn; when correcting for this effect, we found a suggestive trend for more intelligent individuals to be more likely to have sympathetic yawns in response to the video. Interestingly, this effect was education only influenced whether or not the subjects yawned and did not influence how many times they yawned. Our future work will focus on expanding this study to additional individuals and to comparing the measurements to genotypes from whole-exome sequence data. We will also begin to measure additional nonpathogenic traits, such as night visual acuity and face recognition, with the eventual goal of identifying genetic variants that contribute to the control of biologically important, nonpathogenic traits.

2553W

Combined genetic analysis of DNA methylation and gene expression in schizophrenia identifies disease susceptibility loci. *K.R. van Eijk[1,2], M.P.M. Boks[2], S. de Jong[3], E. Strengman[2], R.S. Kahn[1], S. Horvath[4], R.A. Ophoff[1,3,4].* 1) Department of Psychiatry, Rudolf Magnus Institute of Neuroscience, University Medical Center Utrecht, The Netherlands; 2) Department of Medical Genetics, University Medical Center Utrecht, The Netherlands; 3) Center for Neurobehavioral Genetics, University of California Los Angeles, Los Angeles, CA; 4) Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, CA.

Emerging evidence suggests that epigenetic variation such as DNA methylation plays a role in schizophrenia susceptibility. DNA methylation is known to affect gene expression and may thereby contribute to the development of schizophrenia. We performed a study in whole blood with the goal to gain a better understanding of relationships between genetic variation and DNA methylation profiles in schizophrenia patients and healthy controls. To examine the association between DNA methylation and genetic variants, genome-wide genotype and DNA methylation (27K Illumina array) data was obtained from whole blood of 260 schizophrenia patients and 250 healthy controls. For a subset of these subjects array-based gene expression data was also available. By combining epigenetic and gene expression data we identified 661 CpG sites in the genome that are differentially methylated between schizophrenia patients and unaffected controls, and of which methylation status is associated with gene expression levels; moreover, these genes are also differentially expressed between cases and controls. Genetic analysis revealed that 7% of these CpGs are under local (cis) genetic control. We hypothesized that identification of biological relevant links between DNA methylation and gene expression in the context of disease may be an efficient way to identify disease susceptibility loci that are under genetic control (methylation QTLs, mQTLs). We used the available results of the large Psychiatric Genomics Consortium (PGC) schizophrenia genome-wide association (GWA) study to examine whether the observed mQTLs indeed represent susceptibility loci. Overall, we observed significant enrichment of schizophrenia association signal of mQTLs with most pronounced effects for those loci that are also associated with differential gene expression in cases and controls. While this enrichment was already visible when using nominal significant thresholds ($p<0.05$ in PGC) the effects were amplified at more stringent values (e.g. p<1e-4). We identified five disease-associated loci that control DNA methylation in cis, which in turn affect gene expression in a case/control study. Interestingly, one of these loci was previously implicated in the PGC schizophrenia GWA study. Our results suggest that enrichment of biological signal by combining genetic, epigenetic and gene expression profiles from whole blood may be an efficient approach to identify disease susceptibility loci for neuropsychiatric traits.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2554T

Infantile-onset dystonia and juvenile parkinsonism associated with spastic paraplegia caused by mutations in SPG11. *G. Yoon[1,2], B. Baskin[3], M. Tarnopolsky[4], P. Rebeiro[3], B. Banwell[2], P. N. Ray[3].* 1) Div Clinical & Met Gen, Hosp Sick Children, Toronto, ON, Canada; 2) Division of Neurology, The Hospital for Sick Children, Toronto, Canada; 3) Division of Molecular Genetics, The Hospital for Sick Children, Toronto, Canada; 4) Neuromuscular and Neurometabolic Clinic, McMaster University, Hamilton, Canada.

The hereditary spastic paraplegias (HSPs) comprise a genetically and clinically heterogeneous group of neurodegenerative disorders characterized by progressive spasticity, weakness and hyperreflexia of the lower limbs. Mutations in SPG 11 are estimated to account for up to 50% of all recessively inherited HSP, and cause a complex phenotype. We describe a 17 year old girl, born to non-consanguineous parents of mixed Dutch, Polish and Ukrainian descent, who presented with episodic dystonia at 18 months but otherwise normal motor and language development in early childhood. Prenatal and perinatal course was normal, and the family history was non-contributory. She was formally diagnosed with a learning disability at 7 years. She developed a mildly spastic gait at 10 years, which progressively worsened by age 13. At that time she became stiff, developed bradykinesia, and her cognitive function became noticeably worse. At 17 years of age she had developed significant dysarthria, swallowing difficulties and required assistance for most activities of daily living. Physical exam revealed masked facies, bradykinesia, cogwheel rigidity, lower extremities weakness, spasticity and extensor plantar response. She initially responded well to a trial of Levodopa at age 14 years, with improvement of the Parkinsonian symptoms, however by age 17 years, her symptoms had progressively worsened, with increasing rigidity, spasticity and cognitive decline. An MRI of the brain revealed mild diffuse cerebral volume loss, periventricular white matter lesions and thinning of the anterior corpus callosum. Array-CGH, metabolic studies and CSF neurotransmitter analysis were normal. Molecular genetic testing on genomic DNA from blood for the common spinocerebellar ataxias, PINK1, PARK2, and DYT1 was negative. Further molecular analysis of SPG11 revealed a heterozygous c.3664_3665insT (p.Lys1222IlefsX15) mutation. Subsequent mRNA analysis using RT-PCR and sequencing of muscle tissue revealed a novel r.4667_4774del mutation, which caused deletion of exon 27 in the SPG11 transcript, and deletion of 36 amino acids (p.Tyr1556_Val1592delinsPhe) in the SPG11 protein. Levodopa-responsive juvenile Parkinsonism is a rare but increasingly recognized feature of hereditary spastic paraplegia caused by SPG11 mutations. Molecular analysis of muscle tissue may be indicated in patients with a heterozygous mutation but typical features of this condition.

2555F

CNV detection from affymetrix axiom SNP chips in the early markers for autism study. *G. Desachy[1], L. A. Croen[2], L. A. Weiss[1].* 1) Psychiatry, UCSF, San Francisco, CA., USA; 2) Autism Research Program, Kaiser Permanente Division of Research, Oakland, CA.

Purpose:

Autism Spectrum Disorder (ASD) is a highly heritable neurodevelopmental disorder. Although previous research has shown that both single nucleotide polymorphisms (SNPs) and copy number variants (CNVs) contribute to psychiatric disorders including autism, few genetic causes of autism have been identified. Because the new Affymetrix Axiom EUR array (Hoffmann et al., 2011) has not been designed for CNV detection, our primary aim was to develop a method to make copy number calls from this SNP microarray data, and then to follow up deletions or duplications of potential interest for autism risk.

Methodology:

Normalization of the raw intensity data is based on two steps. First, there is a quantile normalization step to make individuals' intensity data comparable (Bolstad, 2001). The second step consists of transforming the probe intensities into Copy Number (CN) calls by subtracting the estimated background signal for each allele of each SNP and then by scaling to estimated copy number (i.e. heterozygotes have median CN=1 for each allele, homozygotes have median CN=2 for one allele and CN=0 for the other allele). A Hidden Markov Model (HMM) is used to make the CNV calls from the normalized data. Due to the variability in the data and the fact that consecutive SNPs are not independent, we implement the HMM on 5-SNP-moving-median data and instead of single SNP data. We developed the methodology through testing on HapMap samples with established CNV calls from other methods.

Results:

We discovered several deletions in regions that might be of great interest because of the genes they include: one case with a deletion on chr8p23.1 (*CSMD1*, previously found to be associated with speech delay, autism and learning difficulties - Glancy et al., 2008) and one case with a deletion on chr9p24.1 (*PTPRD*, previously showing suggestive association in a genome-wide SNP study - Weiss et al., 2009).

2556W

Prioritization of whole exome sequencing in Parkinson Disease. *J. Farlow[1], H. Lin[1], K. Hetrick[2], H. Ling[2], E. Pugh[2], K. Bowling[3], P. Jain[3], Y. Liu[1], K. Doheny[2], R.M. Myers[3], T. Foroud[1].*1) Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN; 2) Center for Inherited Disease Research, Johns Hopkins University, Baltimore, MD; 3) HudsonAlpha Institute for Biotechnology, Huntsville, AL.

Parkinson Disease (PD) is the second most common incurable progressive neurodegenerative disease behind Alzheimer's Disease. Despite being relatively common, the genetic etiologies of PD remain largely unclear, with causative mutations identified in fewer than 5% of patients and risk alleles accounting for only a fraction of the population attributable risk. Identification of additional genes important in disease risk could provide insights about the pathogenesis of PD, as well as options for more accurate ante mortem diagnosis, presymptomatic testing and therapeutic interventions. Whole exome sequencing (WES) was performed in 33 families, which included 92 individuals affected with PD based on UK Brain Bank Criteria and 3 possible PD cases. All samples were sequenced (Agilent© SureSelect™ 50Mb Human All Exon Kit) at one of two centers, with one family sequenced at both centers for quality assurance. Application of Genome Analysis Tool Kit (GATK) quality filters resulted in 149,055 single nucleotide variants (SNVs). Nonsynonymous SNVs within an exon having an allele frequency of <3% in at least one of three gene databases (RefSeq, UCSC, Ensembl) were retained. After removing variants that were not predicted to be deleterious by at least one of three protein prediction programs (SIFT, Polyphen2, MutPred) or that were not in a highly conserved region (Gerp), 20,276 variants remained. Further filtering was performed based on Mendelian inheritance (autosomal dominant or recessive), leaving 10,419 SNVs. A prioritization strategy involving Gene Ontology structural and functional annotations, normal brain expression data (Allen Brain Atlas), and the number of families in which variants in the gene of interest was found generated a list of 59 SNVs in 18 genes. Identified variants included a novel SNV located within LRRK2, a gene already implicated in PD pathogenesis. While several WES studies have been successful at identifying genes important to rare diseases, few have examined how to produce a list of candidate genes contributing to a complex disease from WES data. We show that using a rigorous family-based approach with a prioritization strategy combining rich bioinformatics resources is a powerful method to prioritize candidate genes for a complex disease like PD. Sequencing of these candidate genes in additional familial PD patients is required to validate the utility of this approach.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2557T

High yield of massively parallel sequencing in autosomal recessive intellectual disability in 23 families: identification of the causative mutation in 12 families. *R. Abou Jamra[1], R. Buchert[1], S. Uebe[1], H. Tawamie[1], S.H. Eck[2], E. Graf[2], K. Eberlein[1], J. Bauer[1], EM. Dill[1], C. Blumentritt[1], F.F. Brockschmidt[3], M.M. Nöthen[3,4], J. Schumacher[4], A. Ekici[1], T. Strom[2,5], A. Reis[1].* 1) Institute of Human Genetics, University of Erlangen-Nuremberg, Erlangen, Bavaria, Germany; 2) Institute of Human Genetics, Helmholtz Center Munich, German Research Center for Environmental Health, Neuherberg, Germany; 3) Department of Genomics, Life & Brain Center, University of Bonn, Bonn, Germany; 4) Institute of Human Genetics, University of Bonn, Bonn, Germany; 5) Institute of Human Genetics, Klinikum rechts der Isar, Technische Universität München, Munich, Germany.

To elucidate the genetics of autosomal recessive intellectual disability (ARID), we undertook systematic autozygosity mapping in 69 large, consanguineous families from Syria. We focused on 23 families with smaller candidate regions of sizes between 10 and 45 Mb. After enriching the exomes with the Agilent SureSelect Kit (6 exomes using the 38Mb and 17 using the 50Mb versions), we undertook massively parallel sequencing of the index patients on Solexa (six index patients) and on SOLiD platforms. Subsequently we validated variants by Sanger sequencing and tested them for segregation. We clarified the etiology in four families by identifying pathogenic mutations in the known intellectual disability genes AIH1, ALDH5A1, GPR56, and HGSNAT. Additionally, we described one new ARID gene, AP4S1, and in this context characterized the AP4 deficiency syndrome (Abou Jamra et al., Am J Hum Genet 2011). Furthermore, in seven families we identified convincing mutations (two in-frame deletions and four missense mutations) in EDC3, c9orf4, FAR1, HMG20A, NCAPD2, PPFIA1, and SPATA5. Those mutations were excluded in 280 ethnically matched controls and in silico analysis predicted pathogenic effects in all cases. The function of SPATA5 is unknown, while all other genes are highly expressed in brain and are reported to be relevant for neuronal function. In 11 families the underlying mutation remained undetected. However, as a consequence of the enriching method, part of the targeted nucleotides was not or not properly covered. To identify the mutations in the rest of the families, ensuring complete and high coverage of all exons in the linkage regions, we therefore designed targeted sequencing enriching sets (Agilent, SureSelect) and analysed the products on a SOLID5500xl platform. Those sequences yielded a significantly better variation call (average of 40% more exonic variation). Validation using Sanger sequencing and extensive bioinformatic analysis of all new variants and genes are currently ongoing. We showed that linkage analysis followed by targeted enriching and massively parallel sequencing is an effective method to identify mutations of autosomal recessive intellectual disability.

2558F

Identification of Rare Variants in Childhood Onset Schizophrenia using Exome Sequencing. *A. AMBALAVANAN[1], S.L. Girard[2], J. Gauthier[2], L. Xiong[2], A. Dionne-Laporte[2], D. Spiegelman[2], E. Henrion[2], O. Diallo[2], A. Levert[2], P.A. Dion[2], J. Rapoport[3], R. Joober[1], G.A. Rouleau[2].* 1) Douglas Mental Health University Institute, McGill University, Montreal, Quebec, Canada; 2) Centre de Recherche du Centre Hospitalier de l'Université de Montréal (CRCHUM), and Department of Medicine, University of Montreal, Montreal, Quebec, Canada; 3) Child Psychiatry Branch, National Institute of Mental Health, National Institutes of Health, Bethesda, Maryland, USA.

Background: Childhood Onset Schizophrenia (COS) is a rare severe form of schizophrenia for which, though highly heritable, definitive genetic causes remain elusive. The linkage and association studies of COS have been ineffective in attempting to identify explicit risk variants. We hypothesized that rare de novo mutation in different genes account for a portion of COS individuals. The approach is to identify the causative gene with an enrichment of rare variants from a set of twenty trios (proband and parents) by targeted capture of the exome and massively parallel sequencing of each individual in the family. **Methods:** We capture the whole exome using the Agilent SureSelect enrichment kit (50Mb). We used massively-parallel paired-end sequencing to resequence the exome of each individual on a Solid4 system. We prioritized coding variants which are found only in the affected children within the family. Our best candidate identified variants are validated by Sanger sequencing in the proband and both parents. **Results:** We found rare nonsynomymous, synonymous, splicing and frame shifting denovo variants identified in the coding regions. These potential variants are being validated with Sanger sequencing and further to explore the genes leading to affect the major biochemical pathway.**Conclusions:** We believe that insights from the identification of COS genes may help to device diagnostic tools which in turn may help to intervene earlier before the onset of symptoms which ultimately lead to the identification of therapeutic avenues for better treatment.

2559W

The Asp620Asn mutation of VPS35 in Japanese patients with typical Parkinson disease. *M. Ando[1], M. Funayama[1,2], Y. Li[2], K. Kashihara[3], Y. Murakami[4], N. Ishizu[5], C. Toyoda[6], K. Noguchi[7], T. Hashimoto[8], N. Nakano[9], R. Sasaki[10], Y. Kokubo[10], S. Kuzuhara[11], K. Ogaki[1], C. Yamashita[1], H. Yoshino[2], H. Tomiyama[1,12], N. Hattori[1,2,12].* 1) Neurology, Juntendo University, Tokyo, Tokyo, Japan; 2) Research Institute for Diseases of Old Age, Graduate School of Medicine, Juntendo University, Tokyo, Japan; 3) Department of Neurology, Okayama Kyokuto Hospital, Okayama, Japan; 4) Department of Neurology, Saiseikai Kurihashi Hospital, Saitama, Japan; 5) Department of Neurology, Saitama National Hospital, Saitama, Japan; 6) Department of Neurology, Jikei Daisan Hospital, Tokyo, Japan; 7) Department of Neurology, Kakio Kinen Hospital, Tokyo, Japan; 8) Hashimoto Clinic, Osaka, Japan; 9) Department of Neurosurgery, Kinki University Hospital, Osaka, Japan; 10) Department of Neurology, Mie University Graduate School of Medicine, Tsu, Mie, Japan; 11) Department of Medical Welfare, Faculty of Health Science, Suzuka University of Medical Science, Suzuka, Mie, Japan; 12) Department of Neuroscience for Neurodegenerative Disorders, Juntendo University School of Medicine, Tokyo, Japan.

Background Vacuolar protein sorting 35 (VPS35) was recently reported as a causal gene for late-onset autosomal dominant Parkinson disease (ADPD) using exome sequencing. To date, VPS35 mutations have been detected only in Caucasians with Parkinson disease (PD). The aim of the present study was to determine the incidence and clinical features of Japanese PD patients with p.Asp620Asn (D620N) substitution of VPS35 known as a potentially disease-associated mutation of PD. Methods We investigated the p.D620N (c.1858G>A) mutation in 300 Japanese index patients with ADPD by direct sequencing. We also examined 433 patients with sporadic PD (SPD) and 579 controls by High Resolution Melting analysis. Results The p.D620N mutation was detected in three patients with ADPD (1.0%), one patient with SPD (0.23%), and none of the control subjects. Haplotype analysis suggested at least three independent founders for Japanese patients with the p.D620N mutation. Patients with VPS35 mutation showed typical tremor-predominant PD. Conclusion We reported Asian PD patients with VPS35 mutation. The frequency of the p.D620N mutation of VPS35 is almost the same level or slightly higher in Japan compared to the previous populations studies from Western countries. Still VPS35 mutations are infrequent, p.D620N substitution may be a mutational hotspot across different ethnic groups. Based on the clinical features, VPS35 should be analyzed in patients with PD, especially ADPD or tremor-predominant PD.

2560T

The Autism Sequencing Consortium: Autism gene discovery in >20,000 exomes. *J.D. Buxbaum[1], M.J. Daly[2], B. Devlin[3], T. Lehner[4], K. Roeder[5], M.W. State[6] for the Autism Sequencing Consortium.* 1) Mount Sinai Sch Med, New York, NY; 2) Broad Institute and Harvard Medical School, Boston, MA; 3) University of Pittsburgh, Pittsburgh, PA; 4) National Institute of Mental Health, National Institutes of Health, Bethesda, MD; 5) Carnegie-Mellon University, Pittsburgh, PA; 6) Yale University, New Haven, CT.

Background: There have been tremendous advances in the elucidation of the genetic contribution to autism spectrum disorders (ASD), highlighting the contribution of rare genetic variation, placing limits on the risk conferred by common variation, underscoring the role of de novo mutation, confirming a high degree of genetic heterogeneity, demonstrating pleiotropic effects of syndrome-associated mutations, and identifying an increasing number of specific genes and loci conferring risk. The Autism Sequencing Consortium (ASC) was formed in 2010, in anticipation both of the tremendous impact next generation technologies would have on ASD genomics as well as the many challenges the field would face as a consequence. **Methods:** The ASC sites will be jointly sequencing 9,000 ASD individuals and a total of almost 30,000 samples. Some 25% of the samples are complete, with 65% of the samples to be completed in 12 months. The ASC is managed by a Coordinating Committee and includes multiple working groups to achieve the ASC aims. **Results:** A recent open meeting of investigators, funders and other stakeholders crystallized the plans around molecular approaches in ASD (see video cast at http://videocast.nih.gov/pastevents.asp). Prospective data sharing has led to the simultaneous publication of three papers that benefited significantly from prepublication data-sharing. These studies identified several new ASD genes based on recurrent deleterious de novo mutations while developing frameworks for evaluating statistical significance for such variation. In addition, they estimate the number of ASD genes at ~500 that can be discovered using de novo variation as a filter and highlight the role of advanced paternal age in de novo mutation. The ASC Bioinformatics Hub now hosts several thousand exomes for ASC analyses. We will present further analyses of published and unpublished data, tackling both theoretical and empirical issues involving gene- and site-specific mutation rates and their interpretation.**Discussion:** The current focus of the ASC is gene discovery because it represents a first research step towards understanding the pathogenesis of ASD and developing pharmacological strategies for treating the core symptoms of ASD and etiologically related neurodevelopmental disorders.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2561F

Host genetic contribution to development of progressive multifocal leukoencephalopathy in the presence of immune suppression or immune-modulating drugs. *J. P. Carulli[1], C. Sun[1], S. Szak[1], C. Organ[1], A. G. Day-Williams[1], C. Liu[2], H. H. Ackerman[3], T. Olsson[4], F. Piehl[4], P. Cinque[5], C. S. Tan[6], I. J. Koralnik[6], R. P. Viscidi[7], L. Gorelik[1], K. Simon[1], P. Duda[8], T. Harris[1], S. Bushnell[1], T. Compton[1].* 1) Translational Medicine, BiogenIdec, Cambridge, MA; 2) Framingham Heart Study, Framingham, MA; 3) Department of Statistics, University of California, Berkeley; 4) Department of Neurology, Karolinska Institute, Stockholm, Sweden; 5) Division of Infectious Diseases, San Raffaele Scientific Institute, Milan, Italy; 6) Division of Neurovirology, Beth Israel Deaconess Medical Center and Harvard University Medical School, Boston, MA; 7) Johns Hopkins University, Baltimore, MD; 8) Clinical Development, BiogenIdec, Cambridge, MA.

Progressive Multifocal Leukoencephalopathy (PML) is a rare and often fatal demyelinating disease caused by the JC polyomavirus. JC virus has a reported seroprevalance ranging from 30% to 80% in adult populations worldwide, but PML is very rare and occurs primarily under immune suppression or on treatment with immune-modulating drugs. Prior to the development of anti-HIV drugs, roughly 5% of AIDS patients developed PML. More recently PML has been observed, albeit rarely, in patients taking a number of immune-modulating drugs, including natalizumab (anti-VLA4) for multiple sclerosis (MS). PML can also occur in patients with primary immune deficiencies including Hyper IgE, Hyper IgM, and Wiskott-Aldrich Syndrome, or even with occult or minimal immunosuppression. To test the hypothesis that host genetic variation contributes to development of PML in immune suppressed and/or drug-treated subjects without clinically diagnosed primary immune deficiencies, we have performed a genome-wide association scan using Illumina 610K chips of 109 AIDS-PML cases with 1200 healthy volunteers as controls. We have also completed whole genome sequencing of 15 subjects with MS who developed PML while being treated with natalizumab and 30 MS controls treated with natalizumab for two years or more without developing PML. Both studies were limited to subjects of European descent. The genome-wide association scan indicates that common genetic variation, including copy number changes inferred from the SNP data, does not make a large contribution to development of PML: no SNPs or CNVs reach genome-wide significance. However, genome sequence data reveal rare single nucleotide variants and structural variants that may contribute to PML susceptibility. Putative PML associated variants identified in whole genome sequences are being evaluated for population frequency relative to 1,000 genomes data, validated by genotyping or Sanger sequencing, and further screened at the molecular level in a panel of healthy volunteer and MS DNA samples. Our analysis suggests a complex and genetically heterogeneous contribution to this rare adverse event.

2562W

Disruption of a large intergenic non-coding RNA in subjects with neurodevelopmental disabilities. *C. Ernst[1], M. Talkowski[2], J. Gusella[2], C. Morton[3].* 1) McGill University, Montreal, Quebec, Canada; 2) Center for Human Genetic Research, Harvard Medical School, Boston, MA; 3) Brigham and Women's Hospital, Harvard Medical School, Boston, MA.

Large intergenic non-coding (linc) RNAs represent a novel class of ribonucleic acid whose importance in human disease remains undefined. We identified a severely developmentally delayed 16 year-old female with karyotype 46,XX,t(2;11)(p25.1;p15.1)dn in the absence of clinically significant copy number variants (CNVs). DNA capture followed by next-generation sequencing of the translocation breakpoints revealed de novo disruption of a single non-coding gene on chromosome 2, LINC00299, whose RNA product is expressed in all tissues measured, but most abundantly in brain. Among a series of additional, unrelated subjects referred for clinical diagnostic testing who showed CNV affecting this locus, we identified one with an intragenic, exon-crossing deletion in association with neurodevelopmental abnormalities. No deletion of LINC00299 was seen in any of 5,969 normal adults. Together, these subjects with disruption of LINC00299 implicate this particular non-coding RNA in brain development and raise the possibility that, as a class, abnormalities of lincRNAs may play a significant role in human developmental disorders.

2563T

A rare haplotype containing two AKAP9 missense mutations identified by whole-exome sequencing is associated with risk of Alzheimer Disease in African Americans. *M.W. Logue[1,6], M. Schu[1], B.N. Vardarajan[1], J. Farrell[1], C.T. Baldwin[1,7], M.D. Fallin[8], L.A. Farrer[1,2,3,4,5,6].* 1) Biomedical Genetics, Boston Univ Sch Med, Boston, MA; 2) Neurology, Boston Univ Sch Med, Boston, MA; 3) Opthalmology, Boston Univ Sch Med, Boston, MA; 4) Genetics and Genomics, Boston Univ Sch Med, Boston, MA; 5) Epidemiology, Boston University Sch of Public Health, Boston, MA; 6) Biostatistics, Boston University Sch of Public Health, Boston, MA; 7) Center for Human Genetics, Boston University Sch of Medicine, Boston, MA; 8) Epidemiology, Johns Hopkins University School of Public Health, Baltimore, Maryland.

Previously, we demonstrated that, apart from variants in *APOE* and *ABCA7*, common gene-based SNPs associated with Alzheimer disease (AD) in Caucasians are not risk factors for the disorder in African Americans (AA). To identify functional AD risk variants that are enriched in AAs, we performed whole-exome sequencing (WES) of 7 AA cases from the Multi-Institutional Research on Alzheimer Genetic Epidemiology (MIRAGE) Study, which consists of primarily discordant sibling pairs. WES data were obtained using an Illumina GA IIx sequencer and subsequently mapped to the human genome using BWA. Variant calling with GATK revealed a total of 431 novel (according to dbSNP build 131) non-synonymous SNPs that were observed in more than one subject. Of these, 64 SNPs from 58 genes were prioritized for genotyping based on predicted functionality, potential relevance to AD, and gene-network analysis. Forty-four of these variants were successfully genotyped in 435 AA cases and 422 cognitively normal age-matched AA controls from the MIRAGE and Genetic and Environmental Risk Factors for Alzheimer Disease among African Americans (GenerAAtions) studies. Nominally significant associations were observed with several variants including two non-synonymous variants (rs144662445 and rs149979685) in *AKAP9* (p=0.004 and 0.038, respectively) These SNPs have not been observed in white non-Hispanics, occur at < 1% frequency in AAs according to the ESP database, and are in very high LD. Seven subjects were doubly heterozygous (DH) for these two SNPs, all of them cases. All DH mutations were observed on the same background haplotype. Six other subjects (5 cases and 1 control) were identified which have the same background haplotype and one of the mutations. We are currently performing Sanger sequencing to determine whether or not these individuals are in fact DH. *AKAP9* encodes a kinase anchor protein which is expressed in the hippocampus, cerebellum, and the cerebral cortex. We are currently performing functional studies to determine the mechanisms by which these SNPs might influence AD risk. This study demonstrates how small-scale sequencing projects in diverse ethnic populations can lead to new hypotheses about disease pathogenesis.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2564F

Transposable elements (TE) in schizophrenia: evolutionary analysis and hypothesis for a functional role. *F. Macciardi[1, 4], M. Vawter[1], T. VanErp[1], F. Torri[1], B. Lerer[3], E. Osimo[4], J.A. Knowles[5], S.G. Potkin[1], C. Pato[5], M. Pato[5], S. Gaudi[2].* 1) Dept of Psychiatry & Human Behavior, University of California, Irvine, Irvine, CA; 2) Istituto Superiore di Sanità, Roma, Italy; 3) Hadassah Medical Organization-Hebrew University, Jerusalem, Israel; 4) Universita' degli Studi di Milano, Milan, Italy; 5) Dept of Psychiatry, University of Southern California, Los Angeles, CA.

Introduction: Protein-coding genes in the human genome account for 2% of the total sequence, the noncoding genome accounting for the remaining 98%. Of this large part, Transposable Elements (TEs) represent about 45 to 48%, with an increased architectural complexity within and around brain-related genes. TEs have been implicated in playing a key role in dynamically cis-regulating both genome development and function in recent studies. Since the majority of genome wide association studies (GWAS) in Schizophrenia (SZ) have mapped disease-associated variants to noncoding regions, with an alleged regulatory role, and most variants fall within well-conserved TEs, we are interested to understand how TEs may be implicated in the etiology of SZ. Methods: Our aim is to investigate a possible functional relationship between TEs and genes identified as “best” candidates for Schizophrenia, via SNP mapping or CNV analyses. After in-silico analysis of the genomic architecture of TEs and TE-derived Transcriptional Start Sites (TSS) within and around these SZ-associated genes, we have performed a target resequencing of these genomic regions, and initial whole genome and tissue-specific RNA-sequencing in a sample of SZ and CTRL subjects, to look for differences between who do or do not present the associated risk variants. We then made an evolutionary analysis of TEs and their putative regulated genes. Finally, using DNA variants that may work as eQTL proxies for differential expression of TEs and genes, we looked for differences in fMRI cognitive-based performances in an independent sample of patients and controls. Results and conclusion: We found preliminary evidence of alternative regulation of tissue-specific gene expression in schizophrenic patients vs controls. Using AHI1 (Torri et al, 2010) as an example, we found a peculiar architecture of Alu, Line1 and other TE elements across the gene region, with a brain-specific expression pattern that looks regulated by TE-derived TSS. Only evolutionary recent TEs show a potential regulatory effect for SZ-related genes, as well as marked sequence differences for both DNA and tissue-specific RNA transcriptome between cases and controls. Our results strongly suggest that any investigation of SZ risk genes must include a detailed analysis of TE distribution and expression, and that it is perhaps only through inclusion of these elements that psychiatric diseases will be fully understood.

2565W

Whole-genome sequencing of 50 LRRK2 G2019S carriers discordant for Parkinson's disease. *C.Y. McLean, E. Drabant, E. Harrington, C.B. Do, N. Eriksson.* 23andMe, Inc., Mountain View, CA.

Parkinson's disease is a neurodegenerative disorder that affects over 6 million people worldwide. Over 20 SNPs have been associated with Parkinson's disease susceptibility. The most common known association is the LRRK2 G2019S mutation, with a population frequency ranging from under 1% in Asians to 40% in North African Arabs. G2019S is present in both familial autosomal dominant and sporadic Parkinson's disease and has an estimated penetrance of 30–100%. Higher penetrance estimates in familial cases suggest the presence of additional genetic or environmental penetrance modifiers. We performed whole-genome sequencing on 50 unrelated European individuals concordant for the presence of the LRRK2 G2019S mutation, the absence of the Parkinson's-associated GBA N370S and L444P mutations, and the absence of family history of Parkinson's disease. The cohort, drawn from consenting 23andMe customers, comprises 50 individuals: 37 affected by Parkinson's disease and 13 healthy controls. The median age of onset for affected individuals is 56 (range 38–85). The median age of healthy controls is 66 (range 58–95). Whole-genome sequencing was performed on DNA extracted from bio-banked saliva samples. Individuals were sequenced to a median mapped depth of 45-fold coverage (range 34–52), spanning a median of 97.9% of the genome (range 97.8–98.2%). Mutational burden was similar in cases and controls: a median of 3.59M SNPs and 599k indels arise per genome, in agreement with previous estimates of individual human variation. Concordance between mutations identified by whole-genome sequencing and chip genotyping exceeds 99.7% for all samples. Premature stop codon mutations occur in a median of 96 genes (range 81–114). This study demonstrates the power of the innovative 23andMe customer database for generating research cohorts of interest, confirms the feasibility of generating accurate whole-genome sequencing data from saliva, and provides a rich data set that may help identify genomic differences between LRRK2 G2019S carriers discordant for Parkinson's disease.

2566T

Cumulative Mutation Load in PDZ Domains 4, 5, and 6 of Glutamate Receptor Interacting Protein 2 in Autism. *T. Niranjan[1,2], A. Adamczyk[1], M. Han[1], R. Mejias[1], R. Rose[1], H. Bravo[3,4], M. Taub[4], C. Schwartz[5], D. Valle[1,6], R. Huganir[7], T. Wang[1,6].* 1) McKusick-Nathans Institute of Genetic Medicine, The Johns Hopkins University, Baltimore, MD; 2) Predoctoral Training Program in Human Genetics, The Johns Hopkins University School of Medicine, Baltimore, MD; 3) Center for Bioinformatics and Computational Biology, Department of Computer Science, University of Maryland, College Park, MD; 4) Department of Biostatistics, Bloomberg School of Public Health, The Johns Hopkins University, Baltimore, MD; 5) Greenwood Genetic Center, Greenwood, SC; 6) Department of Pediatrics, The Johns Hopkins University School of Medicine, Baltimore, MD; 7) Department of Neuroscience, The Johns Hopkins University School of Medicine, Baltimore, MD.

Glutamate Receptor Interacting Proteins 1 and 2 (GRIP1/2) are homologous, neuronal scaffolding proteins that interact with multiple binding partners via seven PDZ domains. We had previously identified multiple gain-of-function variants at PDZ domains 4, 5, and 6 (PDZ4-6) of GRIP1 that are associated with altered GRIP1 function in Yeast-2-Hybrid and neuronal recycling assays. To further assess the role of GRIP1/2 in autism, we sequenced exon-containing regions of *GRIP2* in large cohorts of autism patients (n=480) and ethnically matched normal controls (n=480). The study identified a moderate but statistically significant increase in the cumulative burden of non-synonymous mutations ($p < 0.05$, Fisher's Exact) at PDZ4-6 of GRIP2. Mutation burdens of two other functional domain regions, PDZ1-3 and linker-PDZ7, did not show significant difference between cases and controls. PDZ4-6 forms a conserved functional unit that interacts with several neuronal proteins including AMPA glutamate receptors 2 and 3 (GluA2/3), Liprin-α, and ephrinB1/2. Direct binding of PDZ4-6 of GRIP1/2 with the c-terminal domains of GluA2/3 play important roles in GluA2/3 recycling and synaptic organization. Four of the eight non-synonymous coding variants at PDZ4-6 in our autism cohort were found in families with two or more affected sibs of discordant genotype (heterozygous versus wild type) for the respective variants. Intriguingly, sibs who carry GRIP2 non-synonymous variants show more severe deficits in reciprocal social interactions and verbal communication based on their ADI-R test scores as compared to sibs who do not carry these non-synonymous variants. Functional analysis using Yeast-2-Hybrid indicates that all eight non-synonymous variants of GRIP2 show altered binding capacity to GluA2/3. Results of this study support a role of functional GRIP2 variants at PDZ4-6 in the genetic susceptibility to autism. Further *in vitro* and neuron-based assays should help elucidate the mechanisms of GRIP1/2-mediated glutamate signaling disturbance in autism spectrum disorders.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2567F

Candidate genes for sleepwalking from exome sequencing in an autosomal-dominant family. *B. Schormair[1,2,3], H. Slawik[4], B. Frauscher[5], T. Mitterling[5], E. Graf[1,2], T. Wieland[1,2], B. Högl[5], T.M. Strom[1,2], M. Wiegand[4], T. Meitinger[1,2], J. Winkelmann[1,2,3].* 1) Institute of Human Genetics, Helmholtz Zentrum München—German Research Center for Environmental Health, Neuherberg, Germany; 2) Institute of Human Genetics, Technische Universität München, Munich, Germany; 3) Department of Neurology, Technische Universität München, Munich, Germany; 4) Department of Psychiatry and Psychotherapy, Technische Universität München, Munich, Germany; 5) Department of Neurology, Innsbruck Medical University, Innsbruck, Austria.

Sleepwalking (OMIM %613938) is a common childhood parasomnia (prevalence of 26%) that seldom persists into adulthood (prevalence of 3%). There is clear evidence for a genetic basis of the disease. Twin studies have shown a higher concordance rate in monozygotic compared to dizygotic twins and a linkage study has identified a candidate region on chromosome 20, but no responsible gene has been identified yet. We performed exome sequencing in a sleepwalking family of 9 individuals (4 affected, 5 unaffected) with an autosomal-dominant inheritance pattern of the disease. All affected subjects had childhood onset of symptoms which persisted into adulthood. We sequenced two affected siblings and the unaffected mother using Agilent Sure Select All Exome Kit for enrichment and Illumina 100bp paired-end reads for DNA sequencing. We generated on average 14.9Gb of sequence per individual with 89% of the target sequence covered at least 20x and only 0.94% uncovered. For candidate variant identification, we used an autosomal-dominant disease model and looked for heterozygous variants shared by the affected siblings. We focused on missense, nonsense, splice-site, stoploss and frameshift variants and excluded all variants present in dbSNP135, the 1000genomes data, and in an in-house control exome database of unrelated phenotypes. Subsequently, we excluded those variants shared by the two siblings but also found in the unaffected mother. This approach left 47 variants for co-segregation analysis in the pedigree. Of these 47 variants, 5 were compatible with the disease segregation pattern observed in the pedigree. The respective candidate genes include components of the complement system and transcription factors important for the expression of immunoglobins. Based on the often discussed association of sleep and immune system function, these are interesting candidates. In order to identify the true causal gene, we have sequenced the variant-containing region of the candidate genes (200bp of sequence up- and downstream of the variant) in a cohort of 30 sleepwalking cases, but did not find further mutations in these genes. Larger samples of sleepwalking cases are needed to test the potentially causal variants.

2568W

Rare Variants in GWAS loci of Restless Legs Syndrome. *E.C. Schulte[1,2], F. Knauf[2], B. Schormair[2,3], P. Lichtner[2,3], C. Trenkwalder[4,5], B. Högl[6], B. Frauscher[6], K. Berger[7], I. Fietze[8], N. Gross[1], K. Stiasny-Kolster[9], M. Hornyak[10], W. Oertel[9], C.G. Bachmann[5], W. Paulus[5], A. Zimprich[11], A. Peters[12], C. Gieger[12], B. Müller-Myshok[13], T. Meitinger[2,3], J. Winkelmann[1,2,3].* 1) Institut für Humangenetik, Helmholtz Zentrum München, Munich, Germany; 2) Neurologische Klinik und Poliklinik, Klinikum rechts der Isar, Technische Universität München, Munich, Germany; 3) Institut für Humangenetik, Technische Universität München, Munich, Germany; 4) Paracelsus Elena Klinik, Kassel, Germany; 5) Neurologische Klinik, Georg August Universität, Göttingen, Germany; 6) Neurologische Klinik, Medizinische Universität, Innsbruck, Austria; 7) Institut für Epidemiologie, Westfälische Wilhelms Universität, Münster, Germany; 8) Neurologische Klinik, Charité Universitätsmedizin, Berlin, Germany; 9) Neurologische Klinik, Phillips Universität, Marburg, Germany; 10) Klinik für Psychiatrie und Psychotherapie, Albert Ludwigs Universität, Freiburg, Germany; 11) Klinik für Neurologie, Medizinische Universität Wien, Vienna, Austria; 12) Institut für Epidemiologie, Helmholtz Zentrum München, Munich, Germany; 13) Max-Planck Institut für Psychiatrie, Munich, Germany.

Background: Restless legs syndrome (RLS) is a common and genetically complex neurologic disorder characterized by nightly dysesthesias in the legs leading to sleep disturbances. Although GWAS have identified genetic risk factors, a large portion of the heritability remains unsolved and might be attributed to rare and very rare variants (MAF<5%) of strong effect. To date, no such variant is known for RLS. Methods: To assess the role of rare variants at the known GWAS loci, we screened the coding regions and exon-intron boundaries+-10bp of *MEIS1*,*PTPRD*, *BTBD9*, *MAP2K5*, *SKOR1*, *TOX3* and *BC034767* in 188 RLS cases and 188 KORA population controls by high-resolution melting curve analysis. DNAs with aberrant melting patterns were Sanger sequenced. Identified variants with a MAF<5% were genotyped in 3265 cases (65.2+-11.3yrs; 29.3% male) and 2944 controls (56.1+-13.3yrs; 48.7% male). Results: We identified 51 non-synonymous, synonymous, and (near-)splice variants (MAF<5%). Overall, 69 patients harbored a rare variant in one of the “RLS genes” compared to 35 controls (p<5×10E-4, McNemar Test). A similar trend was seen for most individual loci and reached statistical significance for MEIS1 (9 cases and 1 control with a rare variant; p<0.02, McNemar Test). Genotyping of the identified variants in 3265 cases and 2944 controls further underscored this finding. Both synonymous and non-synonymous variants with a MAF<1% were more common in individuals with RLS (p<0.01, chi-squared test). However, this result was even more pronounced for non-synonymous variants with a MAF<0.1% (p<1×10E-5, chi-squared test; OR=6.36; 95% CI:2.49–16.27) and for non-synonymous variants at solely the *MEIS1* locus (p<0.005, chi-squared test; OR=2.39; 95% CI: 1.29–4.44). Two single rare MEIS1 variants (p. R272H (p=0.04, allelic test in PLINK; OR=6.44; 95% CI: 0.79–52.39) and p. M453T (p=0.02, OR=2.06; 95% CI: 1.06–3.97)) were significantly associated with RLS prior to correction for multiple testing. Conclusions: Coding variants with a MAF<5% at RLS-GWAS loci are more common in RLS cases than in controls. This finding becomes more pronounced with increasing variant rarity. At the *MEIS1* locus, a spectrum of common and rare, non-coding and coding variants appears to contribute to disease development, supporting the concept of allelic series in complex diseases. Several rare variants emerge as candidates for rare causal RLS variants but await replication and functional assessment.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2569T

Exome analysis for early-onset amyotrophic lateral sclerosis with autosomal recessive mode of inheritance. *Y. Takahashi[1], K. Higasa[2], S. Takagi[3], T. Kurita[4], H. Ishiura[1], J. Mitsui[1], Y. Fukuda[1], J. Yoshimura[2], T.L. Saito[2], S. Morishita[2], J. Goto[1], S. Tsuji[1].* 1) Department of Neurology, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan; 2) Department of Computational Biology, Graduate School of Frontier Sciences, The University of Tokyo; 3) Department of Neurology, The Jikei University Kashiwa Hospital,Chiba, Japan; 4) Department of Neurology, The Jikei University, Tokyo, Japan.

[Background] Amyotrophic lateral sclerosis (ALS) is a devastating neurodegenerative disorder characterized by progressive and systemic motor neuron loss. Approximately 10 % of ALS is familial ALS (FALS) and the remaining 90% is sporadic ALS (SALS). Numerous causative genes have been identified in autosomal recessive FALS (AR-FALS) including *ALS2*, *OPTN*,*SPG11*, *SIGMAR1*. Recently, exome sequencing has been shown to be a highly efficient approach for identifying causative genes for Mendelian diseases. [Purpose]To identify a causative gene for AR-FALS. [Subjects and Methods] A Mongolian pedigree with two affected siblings presenting with early-onset and slowly-progressive ALS was enrolled in this study. Mutations in previously known causative genes for ALS were excluded. Genotyping was performed using Genome-wide Human SNP array 6.0 (Affymetrix). Parametric linkage analysis was conducted employing SNP HiTLink system on the assumption of autosomal recessive mode of inheritance. Exome sequencing was conducted employing SureSelect Human All Exon 50Mb Kit (Agilent) and TruSeq (Illumina) for exon capture, and Genome Analyzer IIx (Illumina) for massively parallel sequencing. [Results]Genome-wide linkage analysis revealed candidate regions with the total length of 470.1 Mb including 4Mb of homozygous regions. Exome analysis revealed 338 non-synonymous single nucleotide variants (SNVs) and 40 indels which were not registered in dbSNP 135 or not present in in-house variant database. Of these variants, only one homozygous variant was consistent with autosomal recessive mode of inheritance, shared only by the two affected individuals. The variant was not found in 276 controls, conserved among species, and predicted as probably damaging with PolyPhen-2. Mutational analysis of the gene in other FALS (n=19) and SALS (n=203) patients did not reveal any homozygous or compound heterozygous mutations. Since 641 exons in the candidate regions were covered by no more than 5x of read depth, these exons need to be further analyzed. [Discussion and Conclusion]The variant identified in this study is likely a pathogenic mutation for a rare form of AR-FALS, although further studies including identification of mutations in other AR-FALS families and functional analyses should be conducted. Large-scale collaborative framework will be needed to further validate the pathogenicity of the mutation. Large-scale collaborative frameworks will be needed to further validate the pathogenicity.

2570F

Identification of Candidate Genes Through Analysis of Copy Number Variation in Autism Families from Austria. *J.B. Vincent[1], G. Egger[2], A. Lionel[3], H. Mahmood[1], A. Mikhailov[1], K. Roetzer[2], C. Windpassinger[2], W. Kaschnitz[4], S.W. Scherer[2], E. Petek[3].* 1) Neurogenetics Sect, R30, Ctr Add/Mental Hlth,Clarke Div, Toronto, ON, Canada; 2) Institute of Human Genetics, Medical University of Graz, Graz, Austria; 3) Program in Genetics and Genomic Biology, Hospital for Sick Children, Toronto, ON, Canada; 4) Division of General Pediatrics, Medical University of Graz, Graz, Austria.

Autism or Autism-Spectrum-disorders (ASD) is a range of neurodevelopment disorders starting in early childhood, characterized by impairments in communication and reciprocal social interaction, and the presence of restricted and repetitive patterns of behaviour. The contribution of genetic factors to autism is well established by twin studies (Le Couteur et al., 1996), but the kind of genetic transmission is unclear. However, the factor that ASD is a complex non-Mendelian disorder is apparent in most cases. For the current work we recruited 245 family members from 79 ASD families from Styria, Austria. The DNA from probands was run on Affymetrix SNP 6.0 microrarrays to screen for Copy Number Variations (CNVs) in their genomes. Analysis of data was perfomed using Birdsuite, i-Pattern and Genotyping Console, and a list of stringent calls was compared to existing CNV data from over 10,000 Caucasian controls. For stringent calls not present in controls quantitative Real Time-PCR (qRT-PCR) was used to validate the CNVs in the probands and in family members. 15 CNVs from this validated set appear likely to disrupt genes that may be considered good candidates for neuropsychiatric disorders and autism in particular, including AUTS2, DPP6, DLG2, S100B, ARX, DIP2A, GPHN and SH2B1. Several of these genes have also been implicated previously in autism, such as AUTS2, DPP6 and SH2B1, and evidence of their involvement in ASD is growing.

2571W

Genetic and functional study of the disks large-associated protein 2 (DLGAP2) gene as a candidate gene of schizophrenia. *M. Cheng[1], J. Li[2], S. Luu[2], S. Hsu[1], T. Hu[1], H. Tsai[1], C. Chen[3,4,5].* 1) Department of Psychiatry, Yuli Veterans Hospital, Hualien, Taiwan; 2) Department of Psychiatry, Taoyuan Armed Forces General Hospital, Taoyuan, Taiwan; 3) Division of Mental Health and Addiction Medicine, National Health Research Institutes, Zhunan Town, Miaoli County, Taiwan; 4) Department of Psychiatry, Chang Gung Memorial Hospital at Linkou and Chang Gung University School of Medicine, Taoyuan, Taiwan; 5) Institute of Medical Sciences, Tzu-Chi University, Hualien, Taiwan.

Background: A growing evidence indicates that aberrant synaptic dysfunction plays a role in the pathogenesis of schizophrenia. The DLGAP2 gene that encodes the discs, large homolog-associated protein 2 located at the post-synaptic density of neuronal cells involves the neuron synaptic function. This study aimed to investigate whether the DLGAP2 gene is associated with schizophrenia. Methods: We re-sequenced all the exons of the DLGAP2 gene in 523 patients with schizophrenia and 484 non-psychotic controls from Taiwan and conducted a case-control association study and gene functional assay. Results: We identified 19 known single nucleotide polymorphisms (SNPs) in this sample. However, SNP- and haplotype-based analyses showed no association of these SNPs with schizophrenia. We also detected 14 missense mutations and1 amino acid-insertion in this sample, but no increasing burden of these mutations was found in the patient group. Furthermore, we identified 5 private rare variants (c.-69+9C>T, c.-69+13C>T, c.-69+47C>T, c.-69+55C>T and c.-32A>G) at the intron 1 and untranslated exon 2 of the DLGAP2 gene in 5 out of 523 patients but not in 484 control subjects (p=0.03). Computer program predicts that the c.-69+13C>T may change transcription factor binding sites of ATF, HBP, CREB, and the c.-32A>G may change transcription factor binding sites of c-Myc, EcR. Reporter gene assay of the c.-32A>G demonstrated that the c.-32G mutant had a significantly elevated promoter activity than that of wild type. The promoter activity assays of the other 4 rare variants are underway. Conclusion: Our results suggest that multiple private rare variants might occur in the DLGAP2 gene and contribute to the pathogenesis of schizophrenia in some patients.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2572T

Targeted sequencing reveals association of sets of rare variants in novel candidate genes as well as cell adhesion and GTPase regulatory pathways with autism spectrum disorder. *A.J. Griswold[1], J.M. Jaworski[1], S.H. Slifer[1], D.J. Hedges[1,2], R.H. Chung[1], W.F. Hulme[1], I. Konidari[1], P.L. Whitehead[1], J.A. Rantus[1], A.R. Diaz[1], Y.M. Pasco[1], R.M. Tursi[1], S.M. Williams[3], R. Menon[4], M.L. Cuccaro[1,2], E.R. Martin[1,2], J.L. Haines[5], J.R. Gilbert[1,2], J.P. Hussman[6], M.A. Pericak-Vance[1,2].* 1) John P. Hussman Institute for Human Genomics, University of Miami, Miami, FL, USA; 2) Dr. John T. Macdonald Department of Human Genetics, University of Miami, Miami, FL, USA; 3) Vanderbilt University School of Medicine, Vanderbilt University, Nashville, TN, USA; 4) Rollins School of Public Health, Emory University, Atlanta, GA, USA; 5) Center for Human Genetics Research, Vanderbilt University, Nashville, TN, USA; 6) Hussman Foundation, Ellicott City, MD, USA.

There is a strong genetic component to autism spectrum disorder (ASD), yet genome-wide association studies (GWAS) have failed to account for a significant portion of ASD heritability. This has led to the hypothesis that rare variants, each with a small effect size, likely explain the genetic complexity of ASD. To identify rare variants contributing to ASD, we have sequenced ASD associated genes and regions in 919 unrelated cases and 854 controls. Targeted regions were chosen as described in Hussman et. al., 2011, and selected using a custom Agilent SureSelect panel covering 17Mb including 1) exons of 681 genes, 2) evolutionarily conserved regions in introns of the genes plus 5kb from their 5 and 3' ends, 3) evolutionarily conserved regions in associated non-genic haplotype blocks, and 4) entire haplotype blocks with association p-values<0.01. Samples were sequenced on the Illumina HiSeq2000, data aligned with BWA, variants called with the GATK Universal Genotype Caller, and annotated using SeattleSeq, PolyPhen2, and SIFT. Analysis focused on 48,575 single nucleotide variants (SNVs) identified in exons or splice junctions including 9,012 SNVs predicted to be “possibly” or “probably” damaging. To assess whether rare exonic SNVs in each targeted gene are associated with ASD, we used the sequence kernel association test (SKAT) that tests for association between sets of SNVs and a phenotype adjusting for population stratification. This analysis showed association with the transcription factor *ZNF24* (p=3.37×10-5) and the NAD synthesis enzyme *NMNAT3* (p=2.52×10-4). Both are excellent ASD candidates. *ZNF24* has been previously implicated in mouse brain development and *NMNAT3* protein is localized to neuronal mitochondria and acts as a neuroprotective agent. Further, performing SKAT across 30 molecular pathways overrepresented in our targeted gene list, detected association with sets of rare exonic SNVs in cadherin binding (p=0.007) and cell adhesion (p=0.009) pathways. When examining only predicted damaging SNVs, we detect nominal association with the GTPase regulator activity pathway (p= 0.04). These results lend further strong support to the role of cadherins and cell adhesion molecules in ASD and support emerging evidence of GTPase regulation as a novel pathway involved in the disorder. Overall, this study supports the hypothesis that sequencing for rare variants is a successful approach to recognize genes and pathways contributing to the etiology of autism.

2573F

Multiple hits in the Neuregulin signaling pathway in Schizophrenia. *A. Hatzimanolis[1,2], J. McGrath[1], P. Wolyniec[1], V.K. Lasseter[1], G. Nestadt[1], R. Wang[1], T. Li[3], P. Wong[3], D. Valle[4], A.E. Pulver[1], D. Avramopoulos[1,4].* 1) Department of Psychiatry, Johns Hopkins University, Baltimore, MD; 2) National and Kapodestrian University of Athens, Athens, Greece; 3) Department of Neuropathology, Johns Hopkins University, Baltimore, MD; 4) Institute of Genetic Medicine, Johns Hopkins University, Baltimore, MD.

Schizophrenia (SZ) is a highly heritable mental disorder affecting 1% of the population but only recently have linkage and association results reached statistical robustness through collaborative efforts. These difficulties may be explained by the contribution of rare variants and gene-gene interactions. *NRG1* and *NRG3*, encoding Neuregulin 1 and 3 and *ERBB4*, encoding their receptor, have been associated with SZ by many groups including ours. Moreover a mouse model lacking BACE1 whose product cleaves NRG1 releasing the extracellular EGF domain shows phenotypes reminiscent of SZ models. These observations led us to hypothesize that a subset of families might show accumulation of damaging functional variants in core NRG signaling components. To test our hypothesis we focused on 10 genes, the 4 above and 6 encoding the components of gamma secretase thought to cleave the cytoplasmic tails of ERBB4 and NRG1 and play a role in cell autonomous signaling. In our linkage data on 123 multiplex families we found increased scores over the 10 genes, marginally higher than would be expected by random selection (p=0.07). We further observed many significant (p<0.05) pair-wise positive correlations of signals consistent with our multiple hit hypothesis. In 24 pairs of distantly affected relatives, we sequenced all exons of the genes and found potentially damaging variants in *APH1A, APH1B, NRG1* and *NRG3*. Interestingly these clustered in just a fraction of the families, with significant correlations between the number of variants present in *APH1B, NRG1* and *APH1A* (best $p<10^{-4}$). The NRG3 variant did not correlate with those in other genes, but it was shared by distant relatives more often than expected. We explored whether the symptoms of individuals carrying variants in the NRG pathway genes differed as reflected in the 9 factors previously described in a principal components phenotypic analysis in our sample. We found that NRG pathway mutations lead to significantly higher impairment factor scores (r=0.42, p=0.0027). Collectively, our data strongly support that in a subset of patients damaging variation in more than one NRG signaling pathway genes leads disease and more severe impairment. Functional characterization of the variants, investigation of more genes in the pathway possible inclusion of regulatory variation is necessary to strengthen this genetic link and lead to a better mechanistic understanding of the contribution of deranged Neuregulin signaling in SZ.

2574W

Discovery of Rare Variants of Candidate Genes Related to Nicotine Dependence through Deep Sequencing. *S. Wang[1], W.-Y. Cui[1,2], Z. Yang[1], J. Yang[1], J.Z. Ma[3], T.J. Payne[4], C.A. Hodgkinson[5], D. Goldman[5], M.D. LI[1].* 1) Psychiatry and NB Sciences, University of Virginia, Charlottesville, VA; 2) The State Key Laboratory for Diagnosis and Treatment of Infectious Diseases, The First Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, Zhejiang, China; 3) Public Health Science, University of Virginia, Charlottesville VA; 4) Otolaryngology and Communicative Sciences, University of Mississippi Medical Center, Jackson, MS; 5) Laboratory of Neurogenetics, NIAAA, NIH, Bethesda, MD.

Rare variants either within or across genes are believed to act together to influence complex diseases such as nicotine dependence (ND). The increasing availability of massively parallel sequencing offers the opportunity to study the effect of rare variants on complex traits. To identify susceptibility variants for ND, 32 candidate genes implicated in the etiology of ND were selected for target exon sequencing with a pooling approach. A total of 200 sib pairs from the Mid-South Tobacco Family (MSTF) study were divided to 8 pools (50 samples/per pool) based on the ethnic group (100 sib-pairs of either European Americans and African Americans), smoking status (smokers and non-smokers), and Fagerström Test for ND score (light smokers and heavy smokers). Equal amounts of DNA from each sample were pooled, target-captured using custom Agilent arrays, sequenced on a Solexa GA2X at high coverage, and then followed by data analysis, including base quality recalibration and alignment with the BWA program using hg19 assembly build as a reference. Variant callings were conducted using Syzygy designed specifically for a pooling approach. After filtering and mapping, approximately 147 million reads were mapped to the target regions with a median coverage of 106X for each sample included in each pool. Together, we detected about 430 putatively functional variants from the 8 pools. A minimum minor allele frequency of 0.75% and a minimum sequencing reads count of 500 were used as criteria for variant selection. The detected rare variants were ranked according to Polyphen and SIFT scores and allele frequency. Following identification of these variants, we selected 117 putative functional variants for validation using 2000 MSTF and 5000 MST Case-Control (MSTCC) study samples. Of these genotyped variants, 83 were confirmed as rare variants. Among these variants, 71 had an MAF of less than 1%, and 12 had an MAF between 1% and 5%. Case-control-based association analysis was conducted for these variants under different genetic models, revealing that several variants showed significant association with ND (Supported by NIH grant DA-012844).

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2575T

Rare deletions at the Gephyrin (*GPHN*) locus in Autism Spectrum Disorder and Schizophrenia. *A.C. Lionel[1,2], A.K. Vaags[1], D. Sato[1], M.J. Gazzellone[1,2], H.Y. Chen[1], G. Costain[3], G. Egger[4], S. Walker[1], A. Prasad[1], C.R. Marshall[1,2], E.B. Mitchell[5,6], C. Windpassinger[4], J.B. Vincent[7], B.A. Fernandez[8], S. Kirmani[5], J.C. Hodge[5,6], A.S. Bassett[3,9], S.W. Scherer[1,2].* 1) The Centre for Applied Genomics & Program in Genetics & Genome Biology, The Hospital for Sick Children, Toronto, Ontario, Canada; 2) Department of Molecular Genetics and McLaughlin Centre, University of Toronto, Toronto, Ontario, Canada; 3) Clinical Genetics Research Program, Centre for Addiction and Mental Health, Toronto, Ontario, Canada; 4) Institute of Human Genetics, Medical University of Graz, Graz, Austria; 5) Department of Medical Genetics, Mayo Clinic, Rochester, Minnesota, USA; 6) Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota, USA; 7) Molecular Neuropsychiatry & Development Lab, Neurogenetics Section, Centre for Addiction and Mental Health, Toronto, Ontario, Canada; 8) Disciplines of Genetics and Medicine, Memorial University of Newfoundland, St. John's, Newfoundland and Labrador, Canada; 9) Department of Psychiatry, University of Toronto, Toronto, Ontario, Canada.

Several studies have highlighted the involvement of *de novo* and rare inherited copy number variations (CNVs) and point mutations in the genetic etiology of autism spectrum disorder (ASD), particularly affecting genes encoding proteins involved in the neuronal synaptic complex including those from the neurexin, neuroligin and *SHANK* gene families. Parallel findings of rare mutations affecting some of the same genes (e.g. *NRXN1*, *SHANK3*, *CNTNAP2* etc.) in schizophrenia have suggested the existence of shared genetic risk factors for the two neurodevelopmental disorders. Here, we present clinical and genomic characterization of five unrelated subjects with diagnoses of ASD or schizophrenia, who possess rare *de novo* or inherited heterozygous microdeletions overlapping the *GPHN* gene at 14q23.3. Exonic deletions at the *GPHN* locus are extremely rare in the general population (2 in over 15,000 control individuals). Following up on the CNV findings, we identified rare missense variants in additional ASD cases after sequencing all exons and splice sites of *GPHN* using exome and Sanger sequencing approaches. *GPHN* encodes gephyrin, a key scaffolding protein in the neuronal postsynaptic membrane that is responsible for the clustering and localization of glycine and GABA receptors at inhibitory synapses. Gephyrin complexes directly with the neuroligin proteins NLGN2 & NLGN4, and its function in postsynaptic clustering is mediated by the neurexins. Members of the neurexin and neuroligin gene families have been observed to be mutated in schizophrenia and ASD patients. While homozygous *GPHN* mutations have been reported in connection with autosomal recessive molybdenum cofactor deficiency, this is the first report of rare *GPHN* mutations in ASD and schizophrenia. Our findings add to the evidence linking neuronal synaptic genes as etiologic risk factors across a range of neurodevelopmental conditions.

2576F

Reciprocal duplication of the Williams-Beuren syndrome deletion on chromosome 7q11.23 is associated with schizophrenia. *J.G. Mulle[1,2], A.E. Pulver[3,4], J.M. McGrath[3], P. Wolyniec[3], A.F. Dodd[1], D.J. Cutler[2], J. Sebat[5,6,7,8], D. Malhotra[5,6], G. Nestadt[3], D.F. Conrad[9], C.P. Barnes[25], M. Hurles[9], M. Ikeda[10], N. Iwata[10], D. Levinson[11], P.V. Gejman[12,13], A.R. Sanders[12,13], J. Duan[12,13], A.A. Mitchell[14], I. Peter[15], P. Sklar[16,17,18], C.T. O'Dushlaine[16,17,19], D. Grozeva[20], M.C. O'Donovan[20], M.J. Owen[20], C.M. Hultman[21], A.K. Kähler[21,22,23], P.F. Sullivan[22], G. Kirov[20], S.T. Warren[2,24], The Molecular Genetics of Schizophrenia Consortium.* 1) Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, Georgia, USA; 2) Department of Human Genetics, Emory University School of Medicine, Atlanta Georgia, USA; 3) Department of Psychiatry and Behavioral Sciences, Johns Hopkins School of Medicine, Baltimore, Maryland, USA; 4) Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA; 5) Beyster Center for Genomics of Psychiatric Diseases, University of California, San Diego, La Jolla, California, USA; 6) Department of Psychiatry, University of California, San Diego, La Jolla, California, USA; 7) Department of Cellular Molecular and Molecular Medicine, University of California, San Diego, La Jolla, California, USA; 8) Institute for Genomic Medicine, University of California, San Diego, La Jolla, California, USA; 9) The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, UK; 10) Fujita Health University School of Medicine, Toyake, Aichi, Japan; 11) Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, California, USA; 12) Department of Psychiatry and Behavioral Sciences, NorthShore University HealthSystem, Evanston, Illinois, USA; 13) Department of Psychiatry and Behavioral Sciences, University of Chicago, Chicago, Illinois, USA; 14) Department of Forensic Biology, Office of Chief Medical Examiner of the City of New York, 412 E 26th Street, New York, New York, USA; 15) Department of Genetics and Genomic Sciences, Mount Sinai School of Medicine, New York, New York, USA; 16) Psychiatric and Neurodevelopmental Genetics Unit, Massachusetts General Hospital, Boston, MA; 17) Stanley Center for Psychiatric Research, Broad Institute of Harvard and MIT, Cambridge, MA; 18) Division of Psychiatric Genomics, Department of Psychiatry, Mount Sinai School of Medicine, NY, NY; 19) Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA; 20) Department of Psychological Medicine, Cardiff University, Henry Wellcome Building, Heath Park, Cardiff CF14 4XN, UK; 21) Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden; 22) Departments of Genetics and Psychiatry, University of North Carolina, Chapel Hill, North Carolina, USA; 23) Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway; 24) Departments of Biochemistry and Pediatrics, Emory University School of Medicine, Atlanta, Georgia, USA; 25) Department of Cell and Developmental Biology, University College London, London, UK.

We have assessed a population of 583 individuals with schizophrenia (SZ) and 1014 controls all of Ashkenazi Jewish descent for copy number variation. Consistent with other studies, we find an excess of large rare variants in SZ cases, with carriers of rare variants larger than 500 kb having a 2.44-fold increase in risk (95% CI: 1.37–4.4, p-value 0.001, Fisher's exact test). For individuals with deletions larger than 1 Mb, the odds ratio increases to 5.1 (95% CI: 1.51–22.1, p-value 0.004, Fisher's exact test). We also find a significantly higher rate of de novo events larger than 500kb compared to control trios (0.74% in controls vs. 2.05% in SZ trios, p-value 0.04). Among our de novo CNV we find a ~1.3 Mb duplication on chromosome 7q11.23, the reciprocal event of the Williams-Bueren deletion. As de novo duplications at this locus were recently reported in autism populations,we were inspired to investigate the frequency in SZ populations. After formal meta-analysis, we find the 7q11.23 duplication in 11 of 14,387 SZ cases (0.076%) and one of 28,139 controls (0.0035%), for a Mantel-Haenszel corrected odds ratio of 10.8 (95% CI: 1.46–79.62, p-value 0.007). In three SZ duplication carriers with detailed retrospective data, there is evidence of severe social anxiety, social withdrawal, or autism-like symptoms prior to SZ onset. Given that WBS cases are highly social with well-developed language skills, it appears there may be a direct quantitative relationship between dosage of the 7q11.23 region and verbal and social abilities.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2577W

Assessing the impact of different types of *de novo* rare variation in whole-exome sequenced autism spectrum disorders trio families. *D. Pinto[1,2], M. Fromer[1,2], A. Goldberg[1], S.M. Purcell[1], P. Sklar[1], E. Boerwinkle[3], M.J. Daly[3,4], B. Devlin[3,4], E.H. Cook[3,4], K. Roeder[3,4], R.A. Gibbs[3], G.D. Schellenberg[3], J.S. Sutcliffe[3], M.W. State[4], J.D. Buxbaum[1,3,4] on behalf of the ARRA Autism Sequencing Consortium and the Autism Sequencing Consortium.* 1) Psychiatry, Mount Sinai School of Medicine, New York, NY; 2) Genetics and Genome Sciences, Mount Sinai School of Medicine, New York, NY; 3) Lead Investigators of the ARRA Autism Sequencing Consortium; 4) Lead Investigators of the Autism Sequencing Consortium.

Autism spectrum disorders (ASD) are common complex diseases with a strong genetic component. Cumulative evidence from array-based studies (large CNVs: 5–7% of trios) and pilot whole-exome resequencing studies (SNVs: ~10% of trios) suggests that a substantial proportion of risk for ASD resides in rare de novo variants affecting genes. They include variation across a broad range, from single nucleotide variants (SNVs), insertions/deletions, copy number variation (CNV) and larger chromosome imbalances. Thus, there is the need to develop a platform that can maximize the discovery efforts. We are developing a data storage facility to efficiently process large datasets and an analysis pipeline to accurately call multiple types of variants from exome sequencing data, ranging from SNVs to indels and CNVs, and identify *de novo* coding mutations. Specifically for CNV detection, we are currently analysing exomes from ~350 autism trio families using multiple calling algorithms including XHMM, CoNIFER, Splitread and ExomeCNV. For example, preliminary results from XHMM find 1.5 *de novo* events/ 10 trios. Because each of these methods has different abilities, the use of multiple algorithms should maximize discovery rates and sensitivity, especially for variants smaller than 100kb. Specificity and concordance between callers are obtained from additional experimental validation of rare variants in different size ranges, combined with information from available array CNV data. We will present results from inheritance analyses, *de novo* rates and pinpoint novel autism candidate genes in these families.

2578T

Etiologic rare variants in autism multiplex families: exome and CNVs analyses. *C. Toma[1, 2], A. Tristán[1], A. Hervás[3], B. Torrico[1,2], R. Valdés-Mas[4], N. Balmaña[3], M. Maristany[5], X.S. Puente[4], M. Bayés[6], B. Cormand[1,2].* 1) Department of Genetics, University of Barcelona, Barcelona, Spain; 2) Biomedical Network Research Centre on Rare Diseases (CIBERER), Spain; 3) Child and Adolescent Mental Health Unit, Hospital Universitari Mútua de Terrassa, Spain; 4) Department of Biochemistry and Molecular Biology, University of Oviedo-IUOPA, Spain; 5) Developmental Disorders Unit (UETD), Hospital Sant Joan de Déu, Esplugues de Llobregat, Barcelona, Spain; 6) National Center for Genomic Analysis (CNAG), Barcelona, Spain.

Autism is a severe neurodevelopmental disorder. Twin and family studies suggest high heritability. To date a few autism risk genes have been identified, most of them found on the basis of overlap with other syndromic neurodevelopmental disorders, or because they are involved in chromosomal rearrangements or copy number variants (CNVs). Whole exome sequencing (WES) represents a powerful technology to identify rare single nucleotide variants (SNVs), not studied systematically until recently, that may help to depict the complex genetic architecture of autism. In the present study we performed exome sequencing of 10 autism multiplex families with the aim to identify the inherited pool of rare SNVs and uncover new candidate genes for autism susceptibility. In addition, we performed CNV analysis to exclude families with reported fully penetrant structural variants, but also to consider the whole autism genetic landscape. The 10 multiplex families under study include 41 individuals, with 20 parents and 21 probands that fulfil diagnostic criteria for autism spectrum disorder (ASD). To capture the exome subsets of these individuals we used the NimbleGen SeqCap EZ Exome Library SR kit. The coding exons targeted corresponded to approximately 34 Mb. On average, individuals had 86% of target covered at >20X and 50% at >75X. Nonsense, frameshift and splice-site mutations were included in the Sanger validation phase, whereas missense mutations were previously filtered based on pathogenicity predictions (http://ensembl.org/tools.html). Our selection of inherited variants included mutations present in both affected siblings. Structural variants analysis was performed only in the affected individuals with the CytoScan HD array (Affymetrix). Results from the WES showed the absence of de novo pathogenic mutations in the autistic individuals under study, suggesting, as expected, that inherited mutations have a major role in multiplex families. Our preliminary data confirm that the genetic mechanism at the basis of autism may be the cumulative effect of oligogenic heterozygous variants in several genes. In our study we found variants in genes already associated with syndromic autism such as NF1 and TSC1, but also genes like SCN1A or ANK2, which emerged recently as autism genes from exome sequencing studies. The genes found in WES and in the CNV study will be functionally clusterized to identify pathways involved in the disease.

2579F

Utility of a comprehensive multi-gene testing panel for epilepsy and evaluation of the expectations for whole exome sequencing in these disorders. *G. Richard, E. Butler, D. McKnight, K. Retterer, A. Shanmugham, R. Rubenstein, K. Hruska, E. Haverfield, S. Aradhya.* GeneDx, Gaithersburg, MD.

To address the significant genetic heterogeneity in epilepsy, we developed and validated a targeted diagnostic test combining Next-Generation sequencing and exon-level array CGH for 53 genes associated with major primary and syndromic forms of epilepsy. In 176 unrelated, consecutive clinical cases we identified 190 sequence variants and two exonic deletions. Twenty-five individuals had a variant of definite or possible clinical relevance in more than one gene. Twenty-eight (16%) received a conclusive molecular diagnosis, seven (4%) had a likely mutation, and eight (5%) were heterozygous for a single mutation in an autosomal recessive gene. In 55 individuals (31%), variants of unclear significance were detected and parental testing was recommended. Only 19% of known or predicted mutations were in sodium channel genes. In eight cases results had immediate treatment implications. Therefore, using a comprehensive sequencing and exon array approach yielded known/predicted mutations in 20% of individuals with different epilepsy phenotypes, including unexpected rare disorders. Identifying the genetic cause of epilepsy in negative cases remains challenging. To explore the utility of whole exome sequencing (WES), we analyzed exomes obtained by solution-based capture of 32 unrelated control individuals with non-epilepsy phenotypes, focusing on >230 ion channel genes and >290 epilepsy-related genes. Mean coverage across over 8000 exons was 90X, and on average 1.7 exons/gene showed areas of shallow coverage (<20X). Using a cut-off of 20X, ~92% of published mutations in these genes would be identifiable by WES. However, each of the 32 exomes also contained more than 200 non-synonymous SNPs with allele frequency <1% in 1000 Genomes data, creating challenges for data analysis and interpretation. Given the current challenges of WES, our data suggest the most suitable testing strategy for epilepsy is a targeted, comprehensive multi-gene panel followed by WES if negative.

2580W

Rare variants within *PAX6* enhancer elements are enriched in Rolandic Epilepsy. *L. Addis[1], T. Chiang[2], A. Derkach[2], S. Newhouse[3], S. Dobbins[4], I. Tomlinson[5], R. Houlston[4], L.J. Strug[2, 6], D.K. Pal[1].* 1) Department of Clinical Neuroscience, Institute of Psychiatry, King's College London, London, United Kingdom; 2) Program in Child Health Evaluative Sciences, The Hospital for Sick Children, Toronto, ON, Canada; 3) NIHR Biomedical Research Centre for Mental Health, Institute of Psychiatry, Kings College London, London, United Kingdom; 4) Section of Cancer Genetics, The Institute of Cancer Research, Surrey, United Kingdom; 5) Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom; 6) University of Toronto, Toronto, ON, Canada.

Developmental control genes display complex spatiotemporal and quantitative expression patterns making use of cis-regulatory elements such as enhancers. Errors in expression of these genes leads to developmental disorders. One such gene is the transcription factor *PAX6*, whose control elements are spaced upstream, within introns, and primarily downstream of the gene in the neighboring reverse-oriented transcriptional regulator *ELP4*. This is the ultra-conserved downstream regulatory region, DRR. Using linkage and association in Rolandic Epilepsy (RE) families we mapped Centrotemporal spikes (CTS), the autosomal dominant EEG hallmark of RE (also present in 2–4% of the general population) to the *ELP4-PAX6* region. This region is pleiotropic for Speech-Sound disorder (SSD) in RE. We now sought to uncover the causal variants for CTS and SSD using Illumina NGS (196x read-depth) from long-range PCR of the 650Kb linkage region in 27 RE probands of European origin. We compared the frequency of variants identified in our cases to those from 200 European-ancestry controls sequenced by Complete Genomics (~35x read depth). Quality control filters were the same across the epilepsy and CG datasets. We developed a novel statistic to assess association with variants seen at least once in both the case and control groups, and we determined the need to adjust for platform differences. Using rare variant binning analysis (simple sum test and C-alpha) on SNPs present in both case and control samples with MAF<10%, fourteen variants within *PAX6* enhancer elements were enriched in cases (p=0.022). Five SNPs with potential regulatory function were significant in the *PAX6* and *ELP4* 3'UTR binned region (p=0.015). Rare variants binned by entire gene other than *PAX6* were not significantly different between case and control groups (*DCDC1, DNAJC24, IMMP1L* and *ELP4*). Highest ranked SNPs from single SNP analysis clustered within introns of the neighboring genes *DCDC1* and *DNAJC24*. These variants may be part of an extreme *PAX6* DRR or have an important function in their own right. It will be essential to simultaneously follow-up this alternative hypothesis. Many novel SNPs frequent in cases were found within known DRR regions for *PAX6*; intron 9 and 3 of *ELP4*. All of these variants, as well as other novel SNPs of potential function, are now prioritized for genotyping follow-up in a larger set of 150 RE cases and 300 matched controls with functional work to follow.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2581T

Mutations in the novel protein PRRT2 cause paroxysmal kinesigenic dyskinesia with infantile convulsions. *H.-Y Lee[1], Y. Huang[1, 2], N. Bruneau[3], M. Hermann[1], E. Quinn[1, 2], J. Maas[1], R. Edwards[1], K. Bhatia[4], M.K. Bruno[1, 5], H. Houlden[4], J. Jankovic[6], W. Lee[7], U. Müller[8], B.-W Soong[9], K.J. Swoboda[10], N. Wood[4], M. Hanna[4, 12], A. Bowcock[11, 12], P. Szepetowski[3, 12], Y.-H Fu[1, 12], L.J. Ptác{{ak}}ek[1, 2, 12].* 1) Dept Neurology, Univ California, San Francisco, San Francisco, CA; 2) Howard Hughes Medical Institute, San Francisco, California; 3) Institut de Neurobiologie de la Méditerranée, INMED. Inserm U901. Université de la Méditerranée. Marseille, France; 4) Institute of Neurology, University College London, London, UK; 5) Department of Neurology, The Queen's Medical Center, Honolulu, Hawaii; 6) Parkinson's Disease Center and Movement Disorders Clinic, Department of Neurology, Baylor College of Medicine, Houston, Texas; 7) National Neuroscience Institute, Singapore; 8) Institut für Humangenetik, Justus-Liebig-Universität, Giessen, Germany; 9) Department of Neurology, National Yang-Ming University School of Medicine and the Neurological Institute, Taipei Veterans General Hospital, Taipei, Taiwan; 10) Department of Neurology, University of Utah, Salt Lake City, Utah; 11) Division of Human Genetics, Department of Genetics, Washington University School of Medicine, Saint Louis, MO; 12) Senior Investigator of the International Paroxysmal Kinesigenic Dyskinesia/Infantile Convulsions Collaborative Working Group.

Paroxysmal Kinesigenic Dyskinesia with Infantile Convulsions (PKD/IC) is an episodic movement disorder with autosomal dominant inheritance and high penetrance, but the causative gene is unknown. We have now identified four truncating mutations involving the PRRT2 gene in the vast majority (24/25) of well characterized families with PKD/IC. PRRT2 truncating mutations were also detected in 28 of 78 additional families. The PRRT2 gene encodes a proline-rich transmembrane protein of unknown function that has been reported to interact with the t-SNARE, SNAP25. PRRT2 localizes to axons but not to dendritic processes in primary neuronal culture and mutants associated with PKD/IC lead to dramatically reduced PRRT2 protein levels leading ultimately to neuronal hyperexcitability that manifests in vivo as PKD/IC.

2582F

Identification of rare variants under an alcoholism susceptibility linkage peak in a COGA family using whole exome sequencing. *M. Kapoor[1], J. Wang[1], A. Hinrichs[1], S. Bertelsen[1], J. Budde[1], A. Agrawal[1], J. Tischfield[2], L. Almasy[3], M. Schuckit[4], L. Bierut[1], A. Goate[1].* 1) Washington University School of Medicine, St Louis, MO; 2) Department of Genetics/Human Genetics Institute, Rutgers University, Piscataway, New Jersey; 3) Department of Genetics, Texas Biomedical Research Institute, San Antonio, TX; 4) Department of Psychiatry, University of California, San Diego, CA.

Genetic factors may explain as much as 60% of the variance in risk for alcoholism. Genomewide association studies of alcoholism only accounted for small proportion of this variance. We undertook whole exome sequencing to investigate the possible contribution of rare coding variants to susceptibility to alcohol dependence. Finding missing heritability through rare variants is conceptually easy, as it can involve closer scrutiny of genes and regions of interest. Combining the power of traditional linkage studies with next-gen sequencing can help find missing heritability in the form of rare and novel variants. Multiple chromosomal regions have been linked to alcohol dependence and other key phenotypes in our previous studies. In the present study, we have sequenced a trio (two alcohol dependent and one non-dependent control) selected from a large family that has many alcoholic subjects, and showed a very high NPL score on chromosome 2 with DSM-IV diagnosis. We used the SureSelect Human All Exon Kit and protocols provided by Agilent to capture exons, and performed paired-end sequencing on the Illumina GA2 platform. More than 95% of the exome was captured with >8X read depth. A total of 722 variants on chromosome 2 were detected in the affected parent and offspring, but were not present in the unaffected parent. Among these, 83 variants lie under linkage peak and were coding variants. Out of 83 variants, 22 variants had minor allele frequency (MAF) less than 5% in dbSNP and 7 were novel. We validated these 29 variants by genotyping in rest of family members, and in about 4000 cases and controls. Preliminary association analysis using these rare and novel variants helped us to identify 6 variants in 6 genes which may be associated with alcohol dependence. We further sequenced these 6 genes using the targeted capture approach in case-controls and large families to determine enrichment of functional rare variants among alcohol dependents. Our results showed that linkage information in single family can be useful in identifying the functional rare variants enriched in hetrogenous and complex phenotype like alcohol dependence and related traits.

2583W

Deep Resequencing of 9 Confirmed Late-Onset Alzheimer Disease (LOAD) Loci Identifies Multiple Genomic Regions with Potentially Functional Variants. *J.R. Gilbert[1], A.C. Naj[1], L. Wang[1], M.A. Kohli[1], K.L. Hamilton[1], R. Rajbhandary[1], P.L. Whitehead[1], R.M. Carney[2], B. Levin[2], E.A. Crocco[2], C.B. Wright[3], G.W. Beecham[1], E.R. Martin[1], S. Zuchner[1], J.L. Haines[4], M.A. Pericak-Vance[1].* 1) Hussman Inst for Human Genomics, University of Miami Miller School of Medicine, Miami, FL; 2) Department of Psychiatry & Behavioral Sciences, University of Miami, Miami, FL; 3) Evelyn F. McKnight Brain Institute, Department of Neurology, University of Miami, Miami, FL; 4) Center for Human Genetics Research, Vanderbilt University, Nashville, TN.

Besides *APOE*, recent genome-wide association studies (GWAS) have identified variants in nine genes (*CR1*, *CLU*, *PICALM*, *BIN1*, *EPHA1*, *MS4A4A/6E*, *CD33*, *CD2AP*, and *ABCA7*) that confer LOAD risk. However, it is unclear if there are additional novel or rare variants in or near these genes that contribute to LOAD. We resequenced these genes and surrounding regions to identify functional polymorphisms in 291 LOAD cases and 103 cognitively normal controls. We performed targeted capture (NimbleGen SeqCap EZ probe libraries) and next-generation sequencing on Illumina's HiSeq 2000. Target regions included variants in all LD blocks (r^2>0.80) overlapping the nine genes, with an average span of 368.3 kilobases. We filtered variants on three criteria: (a) functionality (missense, nonsense, or splice-site variant), (b) potentially damaging effect as defined by Polyphen-2, and (c) presence in one or more cases and absence in controls. We also examined common and rare variants together grouping by gene and function and using set-based multilocus association testing (RVASSOC). We identified 28,894 variants across the nine regions, out of which 21,203 were variants of minor allele frequency (MAF) <5% (19,100 with MAF <2%). 14,863 variants (51.4%) were novel. Polyphen-2 rated 247 variants as “probably” or “possibly” damaging, of which 180 were novel. Most of the 247 damaging variants were rare (222 with MAF <5%; 209 with <2%), and a subset (n=145; 129 novel) were present in one or more cases while absent in all controls (average MAF = 0.002). Among the rare, potentially damaging functional variants were three missense variants in *CR1*, four in *BIN1*, three in*CD2AP*, six in *EPHA1*, six in the *MS4A* region, 14 in *ABCA7*, and one each in *PICALM*, *CLU*, and *CD33*. Association testing identified several clusters of variants associated with LOAD including seven variants near *MS4A* (*P*=0.0049), seven variants upstream of *BIN1* (*P*=0.0067), and two sets of missense variants near *CD33* (one in *CD33*, *P*=0.031; one in nearby *CLDND2*,*P*=0.045). Deep resequencing identified multiple rare and novel variants near known LOAD susceptibility genes, suggesting that at least some of these may be contributing to risk of LOAD.

2584T

Whole Exome Sequencing in Childhood-onset Schizophrenia. *K. Ahn[1], H. Sung[2], T. Anderson[1], Y. Yao[1], J.L. Rapoport[1].* 1) Child Psychiatry Branch, National Inst Mental Health, NIH, Bethesda, MD; 2) Genometrics Section, Inherited Disease Research Branch, National Human Genome Research Institute, NIH, Baltimore, MD.

Genetic studies of neuropsychiatric illnesses such as childhood onset schizophrenia (COS) have shown that the common disease, common variant hypothesis contributes less to the etiology of neuropsychiatric disorders than previously thought. In contrast, studies of large copy number variants have provided evidence for a larger role of rare, highly penetrant rare variants in COS, which can only be fully explored using the next generation sequencing technology. Whole exome sequencing in 14 COS patients with their unaffected parents was performed with Agilent SureSelect 50 mb capture region and Illumina Hiseq sequencing platform. To test of association for inherited sequence variants in individuals with COS are more likely to harbor inherited rare variants, we employ a few statistical methods of collapsing rare variants by region or gene. These include SKAT, Rare Variant in FBAT, and KBAC. In addition, we explore the parents-origin effects if there is evidence from the rare variant association study. Finally, we perform the pathway analysis for all sequence variants to determine whether the sequence variants identified in COS individuals show enrichment for relevant pathways.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2585F

Complete Genome Sequence Based Genetic Analysis of Monozygotic Twins Discordant For Schizophrenia. *C. Castellani, R. O'Reilly, S. Singh.* Department of Biology and Department of Psychiatry, The University of Western Ontario, London, Ontario, Canada.

The reality of individual genome sequencing now offers a new hope in search of the cause(s) of complex diseases. When combined with genetic relationships, individual sequences add an unrivaled proficiency. Given the near identical genetic structure of monozygotic (MZ) twins, any difference between MZ twins discordant for a disease will have a high likelihood of being causal. With this in mind we have sequenced the DNA of six individuals, which includes two pairs of MZ twins discordant for Schizophrenia and one set of parents. Sequencing was carried out using the Complete Genomics Analysis system (Drmanac et al, 2010). The Complete Genomics platform has a 99.999% sequencing accuracy and calls >99% of the reference genome. The sequences were further assessed for accuracy in relation to Affymetrix SNP Array 6.0 results. Genome wide variations including SNPs, indels, and CNVs were assessed. It has allowed us to evaluate the similarities and differences across unrelated individuals, parents and children, as well as between MZ twins. The results show that an individual carries approximately 3.7 million SNPs, 150 CNVs, 400,000 indels. Also, two unrelated individuals differed for 1.5–1.8 million SNPs (~45%), a parent and child differed for 0.9–1.0 million SNPs (~30%) and a pair of MZ twins differed for 100,000 (~3%) SNPs. Differences in the identity of CNVs for the three comparisons were ~45%, ~30% and ~4%, respectively. It should be noted however that a number of these differences will be the result of sequencing errors. Interestingly, CNV and SNP differences between MZ twins affect a set of genes enriched in neurodevelopmental genes, as well as genes that have been already implicated in Schizophrenia. Further, potentially deleterious mutations (in PDE4DIP and LOC339742) were found in both affected patients but not in their unaffected co-twins or parental samples. The results support our strategy and identify patient specific genetic changes that may lead to Schizophrenia. The novel results re-enforce that individual genomes harbor extensive variability, some inherited and others acquired during parental meiosis and/or mitosis during ontogeny. Even MZ twins are not identical and each individual may be a mosaic, potentially carrying different sequence variations in different cells. This is supported by a high mutation rate and the persistence of genetic diseases with a severely reduced fecundity in all human populations.

2586W

Exome Sequencing of Extended Pedigrees with Late-Onset Alzheimer Disease (LOAD) Identifies *TTC3* as a Candidate Gene for AD. *K. Hamilton[1], M.A. Kohli[1], A.C. Naj[1], R. Rajbhandary[1], T. Plitnik[1], K. John-Williams[1], P.L. Whitehead[1], J.R. Gilbert[1], E.R. Martin[1], G.W. Beecham[1], J.L. Haines[2], S. Zuchner[1], M.A. Pericak-Vance[1].* 1) John P. Hussman Institute for Human GenomJics, University of Miami Miller School of Medicine, Miami, FL; 2) Center for Human Genetics Research, Vanderbilt University, Nashville, TN.

Identifying risk genes for Alzheimer's disease (AD) has focused on testing the `common disease-common variant' (CDCV) hypothesis. While common variants like the ε4 allele of *APOE*clearly play a role in AD, there is a growing realization that the CDCV hypothesis is unlikely to explain all the genetic effect underlying AD. One alternative hypothesis invokes multiple rare variants (RV) in one or more genes, each with stronger individual effects than common variants in genes identified under CDCV. We identified a subset of 6 pedigrees from our collection of 61 extended multi-generational, late-onset AD (LOAD) families to test the rare variant hypothesis. The pedigrees have an average six AD-affected individuals, and have been screened to exclude known familial AD mutations in *APP*, *PSEN1*, and *PSEN2* genes. All affected individuals underwent genome-wide high-density genotyping. Whole-exome sequencing was performed on 4–9 affected individuals per pedigree comprising 2–6 cousin or avuncular pairs. The BWA/GATK package was applied for single nucleotide and insertion-deletion variant calling. Genome-wide genotyping results were used to identify regions of high identity-by-descent (IBD) sharing across the genome. This identifies regions of complete IBD sharing among affected samples, indicating regions most likely to harbor segregating LOAD variants. Promising variants were validated by Sanger sequencing. Using this approach, we identified a novel missense variant (Ser1038Cys) in the gene chromosome 21 gene *TTC3*, which encodes tetratricopeptide repeat domain 3 in one extended pedigree examined. This variant is located at a phylogenetically highly-conserved site, and was not identified in an independent set of ~10,000 control alleles. Co-segregation of the risk genotype with affection status within the pedigree was strong (10/10 affected and 2/6 unaffected carriers) within. We have identified a potentially damaging missense change in *TTC3*, a novel biological candidate for AD that warrants further genetic and functional examination. *TTC3*, which encodes an akt- kinase-regulated factor, is located in the Down syndrome critical region, and has been implicated in neuronal differentiation and learning and memory impairment. This work supports the utility of whole-exome filtering approaches for identifying novel candidates for common diseases in heavily-affected pedigrees.

2587T

Exome sequencing reveals de novo gene disruptions in children on the autistic spectrum. *I. Iossifov[1], M. Ronemus[1], D. Levy[1], Z. Wang[1], I. Hakker[1], J. Rosenbaum[1], B. Yamrom[1], Y. Lee[1], G. Narzisi[1], A. Leotta[1], J. Kendall[1], E. Grabowska[1], B. Ma[1], L. Rodgers[1], J. Troge[1], P. Andrews[1], E. Ghiban[1], J. Parla[1], R. Demeter[2], L.L. Fulton[2], R.S. Fulton[2], V.J. Magrini[2], K. Ye[3], J.C. Darnell[4], R.B. Darnell[4,5], E.R. Mardis[2], R.K. Wilson[2], M.C. Schatz[1], W.R. McCombie[1], M. Wigler[1].* 1) Cold Spring Harbor Laboratory, Cold Spring Harbor, NY; 2) The Genome Institute, Washington University School of Medicine, St. Louis, MO; 3) Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY; 4) Laboratory of Molecular Neuro-oncology, Rockefeller University, New York, NY; 5) Howard Hughes Medical Institute, Rockefeller University, New York, NY.

The genetic component of autism can be transmitted or acquired through de novo mutation. Previous studies have focused on large-scale de novo copy number events, either deletions or duplications, and have identified a large number of autism candidate genes. Because copy number events often span many genes, discerning which of the genes in the target region contribute to the disorder requires complex network analysis and inference. In contrast, with high-throughput DNA sequencing we can readily search for de novo single nucleotide and small insertion or deletion mutations that affect a single gene. Such mutation is fairly common, approximately one hundred new mutations per child, but with only a few—on the order of one per child—falling in coding regions.

We are collaborating in an ongoing large project for sequencing the exomes of 2,800 families from the Simons Simplex Collection. Our preliminary analysis (Iossifov et.al., *Neuron*, 2012) of ~350 of these families, as well as the reports of three other groups, demonstrated the power of identifying de novo mutations through exome sequencing to implicate autism genes. Among the major results were: (1) strong evidence for the role of likely gene-disrupting (LGD) mutations (nonsense, splice site and frame shifts), with affecteds having twice as many LGDs compared to unaffected siblings; (2) the identification of 127 de novo LGD mutations across the four reports with five genes (*CHD8*, *DYRK1A*, *KATNAL2*, *POGZ*, *SCN2A*) having two de novo mutations in unrelated individuals (`double hits'); (3) estimation of a total number of about 400 autism target genes, with a prediction of ~100 double hits after the whole set of 2,800 families has been processed; and (4) the discovery of a strong association between the targets LGD mutations in autism and *in vivo* targets of the RNA-binding translational regulator FMRP (encoded by FMR1, which results in Fragile X Syndrome when silenced or mutated). In our preliminary assessment, we have not seen evidence for a role of de novo missense mutations or inherited LGD variants in autism pathogenesis, but with a greater sample size we hope to detect such effects (if present). We expect to present the results of the joint analysis of ~1000 families by November 2012.

2588F

Secretin receptor and the associated molecular processes relevant to autism spectrum disorder. *K. Kojima, T. Yamagata, A. Matsumoto, M. Saito, E. F. Jimbo, M. Y. Momoi.*Pediatrics, Jichi Medial University, Shimotsuke-shi, Tochigi, Japan.

Secretin is a peptide hormone released from duodenum to stimulate exocrine secretion from pancreas. It also functions as a neuropeptide hormone in the brain. Because secretin receptor (Sctr) deficient mice showed the impairment of social interaction, the involvement of secretin to the pathogenesis of autism spectrum disorder (ASD) is suggested. To investigate the contribution of secretin and its pathway to ASD, we screened SCTR gene for mutations in ASD patients and also analyzed genes regulated by secretin in the mouse brain. (Subjects and methods) (1) We analyzed 200 DNA samples from Japanese ASD patients with the parent's informed consent and Caucasian patients from the Autism Genetic Resource Exchange (AGRE). Each exon and adjacent intron of SCTR was amplified by PCR and subjected to direct sequencing. Base changes detected in ASD samples were screened in 200 Japanese control samples. (2) We injected 1g of secretin into lateral ventriculus of C57BL/6J male adult mice. Three hours after the injection, mRNA was extracted from hippocampus. The expression level of genes was measured using Agilent Expression Array, and data were analyzed by GeneSpring (Agilent Technology). (Results) (1) We detected a missense mutation, P90L in SCTR in one Japanese patient. (2) In mouse hippocampus, expression of oxytocin and vasopressin were markedly elevated after the secretin injection compared to the saline injection. (Discussion) The P90L mutation in SCTR is located in hormone binding domain. Because of the conservative nature of this residue, this mutation can have effect to alter the secretin binding to the receptor and its function. Secretin was reported to activate supraoptic nucleus oxytocin and vasopressin neurons. We showed that secretin stimulated oxytocin and vasopressin expression in hippocampus. These findings suggested the importance of seccretin pathway in the pathogenesis of ASD, that is probably through oxytocin and vasopressin.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2589W

Exome Sequencing of Multiplex, Schizophrenic Families Implicate Variants from Chromosome 5q in Neurocognitive Performance. *M.Z. Kos[1], M.A. Carless[1], J. Peralta[1], E.E. Quillen[1], R.C. Gur[2], M.F. Pogue-Geile[3], K. Prasad[4], J. Blangero[1], V.L. Nimgaonkar[4], R.E. Gur[2], L. Almasy[1].* 1) Dept Gen, Texas Biomedical Research Institute, San Antonio, TX; 2) Department of Psychology, Neurology and Radiology, University of Pennsylvania Medical Center, Philadelphia, PA; 3) Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA; 4) Department of Psychiatry, University of Pittsburgh Medical Center, Pittsburgh, PA.

Schizophrenia (SCZ) is a highly heritable and severe mental disorder characterized by hallucinations, delusions and cognitive deficits, with a lifetime prevalence of 0.4–0.8%. Although a major genetic underpinning has been incontrovertibly demonstrated by twin, family and adoption studies, with an estimated heritability of up to 80%, the genetic architecture of SCZ is unclear, as results for common variants from genome-wide association studies (GWAS) account for only a small fraction of the disorder's additive genetic variation. To better elucidate the biological pathways underlying familial liability of SCZ, we target rare variants via whole exome sequencing of affected families ascertained by the Multiplex-Multigenerational Genetic Investigation (MGI) of Schizophrenia. The MGI families were selected for study on the presence of two or more cases of SCZ or schizoaffective disorder, depressed type (SAD). In addition to clinical evaluations, subjects were tested with a computerized neurocognitive battery that assesses accuracy and efficiency of the following cognitive domains: abstraction and mental flexibility (ABF); attention; verbal, face and spatial memory; language and reasoning; spatial and emotion processing; and sensorimotor dexterity. Exome sequencing was performed on the Illumina HiSeq2000 at 50x coverage. Previous analyses of this data set have revealed significant evidence for linkage to SCZ at chromosome 19q, as well as significant and nominal signals overlapping at chromosome 5q for various neurocognitive measures. Preliminary exome sequence results (n=29 subjects) were examined for these two linkage regions, revealing a number of genome-wide significant associations at chromosome 5q for tests of abstraction and mental flexibility, and attention. For measures of attention, variants from the gene FAM153B are significantly associated with accuracy (rs148943943; P=2.63×10-6) and efficiency (rs148943943; P=7.76×10-6). For ABF, a variant from IL9, a putative candidate gene for SCZ, exhibits significant association with speed (rs1478080; P=3.36×10-6). Previous gene expression analyses have implicated MCAT in ABF performance, with a number of variants from this gene showing evidence of association in our data (best P=1.94 x10-3). Exome sequencing will ultimately be performed on 144 subjects from the MGI samples, with the objective of investigating the influence of exome variants on SCZ and related neurocognitive phenotypes on a genome-wide scale.

2590T

Mining for Rare Genetic Variation underlying Psychiatric Disorders using Family Based Sequencing. *S. McCarthy[1], J. Badner[2], D. Morris[3], M. Ayub[4], M. Kramer[1], D. Blackwood[5], A. Corvin[3], W. Byerley[6], W.R. McCombie[1].* 1) Stanley Institute for Cognitive Genomics, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA; 2) Department of Psychiatry and Behavioral Neuroscience, University of Chicago, Chicago, IL, USA; 3) Department of Psychiatry, Trinity College Dublin, Dublin, Ireland; 4) School of Medicine and Health, University of Durham, Durham, UK; 5) Division of Psychiatry, University of Edinburgh, Edinburgh, UK; 6) Department of Psychiatry, University of California San Francisco, San Francisco, CA, USA.

Interest in the contribution of rare genetic variation to the risk of psychiatric disorders has been revitalized by recent genome wide association and copy number variation surveys of the human genome. Next generation sequencing affords a unique opportunity to explore the impact of rare genetic variation even further at nucleotide resolution. Although targeted and whole genome sequencing provide comprehensive catalogues of rare sequence variation at nucleotide resolution, evaluating which of the thousands of variants may be risk factors for schizophrenia is an arduous and difficult task. Family based strategies such as parent-offspring trios and extended outbred or consanguineous pedigrees are powerful approaches to increase the signal to noise of risk variants for bipolar disorder, major depression and schizophrenia above the background of benign or insignificant genomic variation. Here we shall present the principles and hypotheses of family based sequence analyses, and present our findings from human exome sequencing of parent-offspring trios with schizophrenia as well as findings from whole genome sequencing in extended pedigrees with schizophrenia and consanguineous families with bipolar disorder and major depression. We shall highlight the importance of rare, private de novo or inherited variants in risk of schizophrenia in over 60 schizophrenia trios. In addition we will present our analysis of >15 whole genomes representing two extended pedigrees with schizophrenia and one large consanguineous family with bipolar disorder and major depression. Our findings illustrate the significance of family specific variants with potentially damaging effects in genes with neurobiological function but also highlight the genetic overlap between major psychiatric illnesses. In conclusion our results demonstrate how advancements in next generation sequencing provide the opportunity to disentangle the complex genetics of psychiatric disorders and how family based approaches are a powerful resource to enrich and harrow rare genetic variation contributing to the risk of schizophrenia, bipolar disorder and major depression.

2591F

Identification of the causative gene for SPG27 by Exome capture. *A. Noreau[1], I.A. Meijer[1], S.L. Girard[1], D. Spiegelman[1], A. Szuto[1], P. Cossette[1], P.A. Dion[1,2], G.A. Rouleau[1,3,4].* 1) The Centre of Excellence in Neuroscience, CRCHUM, Montreal, Quebec, Canada; 2) Université de Montréal, Department of pathology and cellular biology, Montreal, Quebec, Canada, H3T 1J4; 3) University of Montreal, Faculty of Medicine, Department of Medicine, Montreal, Quebec, Canada, H3C 3J7; 4) Research Center, CHU Sainte-Justine, Montreal, Quebec, Canada, H3T 1C5.

Hereditary spastic paraplegia (HSP) is a motor neuron disease for which the key symptoms are lower limb spasticity and weakness because of progressive neurodegeneration events. The group of HSP is particularly heterogeneous because different modes of inheritance were observed (dominant, recessive and X-linked) as well as various levels of symptoms complexity; pure and complicated forms were described. To date, over 45 spastic paraplegia locus were identified and 22 genes identified. A previous linkage analysis on an autosomal recessive pure HSP family identified a new locus (SPG27) on the 10q22.1-10q24.1 region; a locus partially overlapping the SPG9 locus. We proposed to perform use an exome capture approach with the Agilent Technology SureSelect Human All Exon Kit (50MB) and proceed to resequencing with ABI SOLID4 DNA sequencer. Four affected individuals, the unaffected-carrier father and another unaffected sibling carrying the same disease-allele as the father will be used. As we may be looking for compound-heterozygote mutations, we will develop an algorithm to determine the parental origin of variants and the more promising mutations will be validated through Sanger sequencing. By this method, it will be possible to cover the 40 genes, find the causative mutation and maybe solve the SPG9 clue.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2592W

DEC2 Gene Mutations Associated with distinct Sleep Phenotypes and Sleep length. *R. Pellegrino[1, 2, 3], I.H Kavakli[4, 5], N. Goel[6], C. Cardinale[1], D. Dinges[6], S. Tufik[3], J. Hogenesh[4], A. Pack[2], H. Hakonarson[1].* 1) Center for Applied Genomics, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA; 2) Center for Sleep and Circadian Neurobiology, University of Pennsylvania - School of Medicine, PA, USA; 3) Universidade Federal de São Paulo - UNIFESP; 4) Department of Pharmacology, Institute for Translational Medicine and Therapeutics, University of Pennsylvania School of Medicine, PA, USA; 5) Department of Chemical and Biological Engineering, Koc University, Istanbul, Turkey; 6) Division of Sleep and Chronobiology, Department of Psychiatry, School of Medicine, University of Pennsylvania, PA, USA.

Introduction: DEC2 gene is a negative component of the circadian clock by suppressing BMAL1-CLOCK transactivation. It belongs to a basic helix-loop-helix protein family and they dimerize with each other and can affect gene transcription by binding to specific DNA sequences. DEC2 and DEC1 repress Clock/Bmal1-induced transactivation of the Per2 promoters through direct protein-protein interaction and/or competition for E-boxes. DEC2 is probably a component of the mammalian clock system. It has been reported that a mutation in DEC2 is associated with a human short sleep phenotype. Methods: Given the complexity of phenotyping to study sleep homeostasis and genetics, we intend to evaluate differences in the phenotype between monozygotic and dizygotic twins. In addition, we also included unrelated samples and controls. We sequenced 589 (total) individuals for 5 exons covering the DEC2 gene. In order to examine the effect of the mutations on DEC2 transactivation, wild-type (WT) and mutants were used in a luciferase assay to evaluate the activity of Clock/Bmal-Per2-mediated transactivation. Results and discussion: We discovered a new Dec2 mutation (Dec2-Y362H) in a genetically related cohort which associated with a short sleep assessed by actigraphy and polysomnography. Results indicated that both mDec2-Y362H and mDec2-P384R have comparable attenuated repressive activity on BMAL1-CLOCK transactivation compared with wild type mDec2. On the other hand, mDec2-P384Q mutants displayed the opposite effect on BMAL1-CLOCK transactivation where it strongly inhibited BMAL1-CLOCK transactivation on mPer2 promoter. The Y362H and P384R is correlated to a decrease of sleep length and the individuals with P384Q mutations demonstrated normal or increased sleep length. Conclusion: We observed the DEC2 region in exon 5 as a “hot spot” for mutations. This region can be associated with important changes in phenotypes according with the position of the mutation and amino acid substitution. It is remarkable how recent data on the genetics of human sleep and increased knowledge of regulatory mechanisms of sleep quality and quantity will facilitate the development of prospective interventions in situations relating to sleep deprivation.

2593T

Common and rare variants in obsessive-compulsive disorder (OCD) identified by exome and targeted resequencing. *D. Trujillano[1], S. Ossowski[1], P. Alonso[2], C. Tornador[1], J.M. Menchón[2], M. Gratacòs[1].* 1) Genes and Disease, Centre for Genomic Regulation (CRG-UPF), Barcelona, Catalonia, Spain; 2) Psychiatry Department, Bellvitge University Hospital, Barcelona, Catalonia, Spain.

Obsessive-compulsive disorder (OCD) is a mental disease characterized by unwanted and intrusive thoughts, images, or impulses (obsessions) and/or repetitive behaviors or mental acts (compulsions). Although OCD affects 2%–4% of the population worldwide, very little progress has occurred in the elucidation of the underlying risk-conferring genetic variation by means of candidate-gene, linkage, and GWAS. We hypothesize that a wide spectrum of common and rare genetic variants contributes to OCD. In this study, taking advantage of NGS technologies, we have examined the extent to which common and rare variation might play a role in OCD. We used a two-step approach to study OCD. We begun with the screening of the exomes by NGS of 40 extremely well characterized OCD patients (clinical evaluation, course of the disease and MRI) to identify candidate genes affected by deleterious common and/or rare single nucleotide variants (SNV). For this, we used the SureSelect Human All Exon Kit (Agilent) in combination with the Illumina's GAIIx sequencer. Then, we followed this initial screening with association studies of the selected SNVs in a larger cohort of 500 OCD cases and 500 controls to identify statistically significant variants by means of direct genotyping for common SNVs, and NGS targeted resequencing for novel and rare SNVs. From the high depth of sequencing data generated, we selected 436 genes recurrently affected by deleterious novel and rare SNVs across the 40 patients of the discovery set. The resulting list of genes was targeted resequenced in 500 OCD cases and controls. In parallel, we also identified 60 deleterious common SNVs with significantly altered allelic frequencies in our initial sample set when compared to the normal population. Preliminary results of this study have identified a large number of SNVs that show significant different allelic frequencies between OCD cases and control samples. By combining targeted resequencing of genes enriched in novel and rare variants with genotyping of common SNVs, we will provide a comprehensive characterization of the genetic basis of. Coexistence of common and rare SNVs in the same genes and functional predicted consequences should help to prioritize functional studies, and hopefully facilitate to define a set of genes for comprehensive evaluation of OCD risk assessment and fine diagnosis. The knowledge acquired in this project should provide a grounding basis for new treatment strategies for OCD.

2594F

Deep whole genome sequencing in 18 Irish individuals with schizophrenia. *B.T. Webb[1,2], E. Loken[1], T.B. Bigdeli[1], B. Wormley[1,2], F.A. O'Neill[4], D. Walsh[5], K.S. Kendler[1,2], B.P. Riley[1,2,3].* 1) Virginia Institute for Psychiatric and Behavioral Genetics; 2) Departments of Psychiatry and; 3) Human and Molecular Genetics, Virginia Commonwealth University, Richmond, VA; 4) Department of Psychiatry, Queens University, Belfast; 5) The Health Research Board, Dublin, Ireland.

Background: Recent evidence supports the involvement of both common and rare variation to the genetic architecture of schizophrenia. We have previously conducted GWAS on the Irish Study of High Density Schizophrenia Families and the Irish Schizophrenia Case Control samples. We now seek to understand the contribution of rare and alternative classes of variation in these samples. Whole genome sequencing can be used to; 1) estimate the frequency of novel or very rare high impact variants, 2) measure mutational load in sporadic versus familial cases, 3) assess additional classes of variation including structural variation and mobile insertion element, 4) potentially discover population specific haplotypes and 5) provide additional Irish specific reference samples for more accurate imputation low frequency and rare variation. Methods/ Results: Sequencing, alignment, variant calling, and annotation of 18 case samples was performed by Complete Genomics. The median amount sequence where both pairs mapped was 162.8 Gb (55.2X) with 0.69, 0 .82, 0.92 of the genome covered at >40, 30, and 20X, respectively. The SNP transition/ transversion ratio for high quality calls was 2.14 and 3.04 for the genome and exome, respectively. Novel SNPs (unobserved in dbSNP132) were rare (0.034) while novel insertions (0.195), deletions (0.229), and substitutions (0.354) were more common. A median of 0.029 of the genome was associated with novel Copy Number Variation (CNV). Each sample had 1548 high quality structural variant junctions on average. A median of 4414 Mobile Element Insertions (MEI) per sample was observed with most MEIs being novel (0.767). Novel high impact variation including nonsense, frameshift, misStart, and splice site disruption was common with 352 candidate events per genome Conclusion: Metrics for Complete Genomics genomes are in agreement with recent reports from the 1000 genomes and NHLBI Exome sequencing projects for measures such as rare functional variation frequency. Further, the addition of population or study specific deep sequenced genomes yields low frequency variation not in current reference panels and therefore can improve imputation. Deep sequencing also offers the opportunity to investigate additional classes of variation that may contribute to the genetic architecture of schizophrenia. Finally, the identification of loci with multiple high impact rare variants can be used to improve pathway and gene-set enrichment analysis of GWAS data.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2595W

The Genetic Architecture of Severe Non-Syndromic Sporadic Intellectual Disability. *T. Wieland[1], D. Wieczorek[2], E. Graf[1], S. Endele[3], T. Schwarzmayr[1], J. Beygo[2], N. Di Donato[4], A. Dufke[5], M. Hempel[6], D. Horn[7], P. Joset[8], A. Röpke[9], U. Moog[10], E. Wohlleber[11], C. Zweier[3], A.B. Ekici[3], A.M. Zink[11], A. Rump[4], C. Meisinger[12], H. Grallert[13], H. Sticht[14], A. Schenck[15], H. Engels[11], G. Rappold[10], P. Wieacker[9], O. Riess[5], T. Meitinger[1,6], A. Reis[3], T.M. Strom[1,6], A. Rauch[3,8].* 1) Institute of Human Genetics, Helmholtz Center Munich, Munich, Germany; 2) Institute of Human Genetics, Universitätsklinikum Essen, Essen, Germany; 3) Institute of Human Genetics, Universität Erlangen-Nürnberg, Erlangen, Germany; 4) Institute of Clinical Genetics, Technische Universität Dresden, Dresden, Germany; 5) Institute of Human Genetics, Eberhard Karls Universität Tübingen, Tübingen, Germany; 6) Institute of Human Genetics, Technische Universität München, Munich, Germany; 7) Institute of Medical Genetics, Charité, Universitätsmedizin Berlin, Berlin, Germany; 8) Institute of Medical Genetics, University of Zurich, Schwerzenbach-Zurich, Switzerland; 9) Institute of Human Genetics, Westfälische Wilhelms-Universität Münster, Münster, Germany; 10) Institute of Human Genetics, Ruprecht-Karls-Universität Heidelberg, Heidelberg, Germany; 11) Institute of Human Genetics, Rheinische Friedrich-Wilhelms-Universität Bonn, Bonn, Germany; 12) Institute of Epidemiology II, Helmholtz Zentrum München, Neuherberg, Germany; 13) Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, Neuherberg, Germany; 14) Institute of Biochemistry, University of Erlangen-Nuremberg, Erlangen, Germany; 15) Department of Human Genetics, Nijmegen Centre for Molecular Life Sciences, Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands.

Background The etiology of intellectual disability (ID) remains elusive in the majority of patients due to the absence of morphological or positional clues and lack of suitable screening methods.

Methods We used exome sequencing in parent-child trios to identify *de novo* variants in 51 individuals with sporadic non-syndromic intellectual disability. Twenty trios were investigated as controls.

Results We defined 88 *de novo* variants in the case group resulting in an exomic mutation rate of 1.73 per individual per generation. The corresponding figures in the control group were 24*de novo* variants and 1.14 events per individual per generation. Most strikingly, a considerably higher number of individuals in the case group had loss-of-function variants (cases= 20/51 versus controls=2/20) indicating their contribution to disease etiology. Evaluation of the single variants revealed that 16 patients carried *de novo* variants in known ID genes with three recurrently mutated genes (STXBP1, SYNGAP1, and SCN2A). Focusing only on loss-of-function variants, we considered at least 7 mutations in 7 novel genes to be disease causing. We further identified several missense alterations with potential pathogenicity. Interestingly, our study revealed no evidence for a sizable contribution of autosomal recessive inheritance.

Conclusions Our results demonstrate that *de novo* point mutations and small indels are a major cause of severe, sporadic non-syndromic ID with high locus heterogeneity. The high number of *de novo* variants in known ID genes is only partially attributable to known unspecific phenotypes. Rather, several patients did not meet the expected syndromic manifestation indicating a strong bias in the current clinical syndrome descriptions.

2596T

A spectrum of mutations in 12 PLP1-related disorders patients from SARAH network of rehabilitation hospitals - Brazil. *S.C.L. Santos, D.R. Carvalho, R.I. Takata, C.E. Speck-Martins.* Human Genetics, SARAH Network of Hospitals, Brasília, DF, Brazil.

PLP1-related disorders represent a group of neurological features ranging from Pelizaeus-Merzbacher disease (PMD, MIM# 312080) with severe CNS involvement to the milder spastic paraplegia type 2 (SPG2, MIM# 312920). Mutations in PLP1 gene (MIM# 300401) lead to the arrest of myelination and can be grouped into two main categories: PLP1 gene duplications (60–70%) and intragenic sequence variants (15–20%). Twelve male patients from 11 unrelated families were screened for PLP1 gene duplications using SALSA MLPA kit P022-B1. PLP1 gene sequencing (ABI3130) was performed in negative patients for duplication. PLP1 duplications were identified in 4 (36%) unrelated patients whereas two of them presented duplications in flanking MLPA probes besides PLP1 gene probes. The duplication sizes were confirmed by using aCGH [(arr Xq22.1q22.3(RP3-347M6->RP1-75H8)x3 and arr Xq22.1q22.3(RP3-347M6->RP1-75H8)x3]. Seven different PLP1 intragenic mutations were found in the remaining 8 (64%) patients. So far, 4 out of 7 PLP1 intragenic mutations were not described yet, including two missense mutations (p.Tyr58Cys, p.Lys151Arg) and two microdeletion (c.216_217delC, c.354_355delA) resulting in frameshift (Ile72AlafsX44, Thr118GlyfsX28, respectively). Functional significance of the novel missense mutations was analyzed by the MutPred and SIFT tools. Both tools predicted a potencial deleterious effect for p.Tyr58Cys mutation and a non-deleterious effect for p.Lys151Arg mutation. Further investigation must be applied to confirm these data. All PLP1-duplication patients presented the classical form of PMD, but the patient with the largest PLP1 duplication showed a more severe phenotype presenting microcephaly, a high frequency of epileptic episodes and brain atrophy. Peripheral neuropathy was absent in PLP1-duplications patients, but was present in two patients with intragenic mutations. Our study detected novel intragenic mutations in PLP1-related disorders patients, as well as confirmed some genotype-phenotype correlations in a series of Brazilian patients.

2597F

Association between the NMDA Receptor Subunit 2B Gene (GRIN2B) gene and Alzheimer's Disease. *F. Trecroci, R. Cittadella, M. Liguori, F. Condino, P. Spadafora, E.V. De Marco, G. Di Palma, V. Andreoli.* Institute of Neurological Sciences-National Research Council, Pianolago di Mangone, Cosenza., Italy.

Various mechanisms may contribute to neurodegeneration in Alzheimer's disease (AD), the most common cause of dementia, including glutamate-mediated excitotoxicity. While normal glutamatergic neurotransmission is required for learning, memory and information processing of human brain, excessive glutamatergic activity may cause neuronal dysfunction and eventually degeneration. This excitotoxic effect appears to be mediated, in part, by the N-methyl-D-aspartate receptors (NMDAR). The NMDA receptor subunit 2B (NR2B) has attracted more attention due to its characteristic distribution and selective reduction in AD brain. The human gene encoding this subunit, named GRIN2B (chr 12p12), consists of 13 exons and it is expressed in the hyppocampus, basal ganglia, and cerebral cortex. We hypothesized that GRIN2B represents a candidate gene for AD and we carried out systematic mutation search of the GRIN2B gene in an independent series of Southern Italy samples. Clinical data and blood sample were collected from 270 selected AD patients, after inform consent. All coding exons and exon-intron junctions were amplified using specific primers and a mutational screening was done by DHPLC and DNA direct sequencing. At moment, five single nucleotide polymorphisms (SNPs) were detected. Although these variants are in the coding sequence, they are silent polymorphisms: c.15G>T-Ala5; c.366C>G -Pro122; c.1665C>T-Ser555; c.2514C>T-Cys838; c.3534C>T-His1178. First, we investigated three SNPs, Pro122 in exon 2; Ser555 in exon 8 and His1178 in exon 13 of the GRIN2B gene, in our AD patients and 250 healthy Caucasian age-matched controls. The distribution frequency of neither GRIN2B genotype nor allele (Pro122 P=0.337; Ser555 P=0.142 and His1178 P=0.868) was significantly different between AD patients and normal controls, even when the subjects were stratified by gender, APOE and age of disease onset in AD patients. The results of the Ala5 and Cys838 polymorphisms were omitted because of the low frequencies of the detected alleles. The present study, based on a case-control design, was performed on a cohort of unrelated Italian patients affected with AD. Furthermore, it appears that the GRIN2B polymorphisms analysed here did not show an association with the disease in our sample of AD patients. However, we will continue with a more comprehensive screening of GRIN2B polymorphisms that might be useful to determine the involvement of this gene in AD.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2598W

Rare missense variants in *CHRNA5* and *CHRNB3* are associated with increased risk of alcohol and cocaine dependence. *G. Haller[1,2], J. Budde[1,2], M. Kapoor[1,2], A. Agrawal[1], H. Edenberg[3], J. Kramer[4], L. Bierut[1], A. Goate[1,2].* 1) Deparment of Psychiatry, Washington University, Saint Louis, MO; 2) Deparment of Genetics, Washington University, Saint Louis, MO; 3) Deparment of Psychiatry, Indiana University, Indianapolis, IN; 4) Deparment of Psychiatry, University of Iowa, Iowa City, IA.

Previous findings have demonstrated that variants in nicotinic receptor genes are associated with alcohol and cocaine dependence. It is unclear, however, whether this is independent of their effect on nicotine dependence. Recent studies have shown that multiple nicotinic receptor genes harbor rare missense variants associated with reduced risk of nicotine dependence when taken in aggregate. In order to investigate the possible contribution of rare variants to the development of substance dependence, specifically alcohol and cocaine dependence, we undertook pooled sequencing of the coding regions and flanking sequence of the *CHRNA5*, *CHRNA3*, *CHRNB4*, *CHRNA6* and *CHRNB3* genes in 287 African American and 1028 European American individuals from the Collaborative Study of the Genetics of Alcoholism (COGA). All members of families for whom any individual was sequenced (2504 African Americans and 7318 European Americans) were then genotyped for all variants identified by sequencing. For each gene, we then tested for association between carrying at least one rare missense variant and each substance dependence phenotype, incorporating family structure using a kinship matrix. For European Americans, we find increased DSM-IV alcohol dependence symptoms among carriers of missense variants in *CHRNA5* ($p = 0.003$, $\beta = 0.35$) and increased DSM-IV cocaine dependence symptoms among carriers of missense variants in *CHRNA5* ($p = 0.006$, $\beta = 0.28$) and*CHRNB3* ($p = 0.004$, $\beta = 0.24$). Furthermore, these associations remained significant, when the average number of cigarettes smoked per day was included as a covariate, suggesting that these findings are not mediated through the effect of these variants on nicotine consumption alone. African American carriers of missense variants in *CHRNA5* also showed increased DSM-IV alcohol dependence symptoms ($p = 0.007$, $\beta = 1.29$). As a note, carriers of the GWAS significant SNP in *CHRNA5* (rs16969968/D398N) were not included in any analyses so as to insure novel findings. These are the first results to implicate rare variants in *CHRNB3* or *CHRNA5* in risk for alcohol or cocaine dependence and the first results to implicate *CHRNB3* generally in alcohol or cocaine dependence.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2599T

Excess of Rare Variants from Targeted Resequencing of the NKAPL Gene in Schizophrenia. *E.K. Loken[1], D. Brohawn[1], D. Walsh[2], F.A. O'Neill[3], K.S. Kendler[1], B.P. Riley[1].* 1) Virginia Commonwealth University School of Medicine, Richmond, VA; 2) Health Research Board, Dublin, Ireland; 3) Queens University, Belfast, United Kingdom.

Background: The most robust signal observed in genome-wide association studies (GWAS) of schizophrenia is in a region between 27 and 32 Mbp on chromosome 6p22. Although these signals are due to common variants in the population, sequencing GWAS genes or intervals to identify rare, functional and deleterious variation in cases is increasingly important in corroborating associations in specific genes. A GWAS in a Han Chinese sample implicated the NFKB activating protein-like (*NKAPL*) gene as a schizophrenia common variant locus. Through resequencing *NKAPL* in Irish schizophrenia cases and controls we demonstrate validation of variants identified in high-throughput sequencing and novel functional variant detection in a schizophrenia GWAS locus.

Methods: We sequenced 48 cases from the Irish Case-Control Study of Schizophrenia and 44 Irish controls from the Trinity Biobank. We amplified a 10kb region containing the *NKAPL* locus using long range PCR. We then performed multiplexed sequencing on the Ion Torrent PGM. We called variants using SAMtools. We compared the burden of rare variants (minor allele frequency < 5% in cases or controls) between cases and controls using permutation testing.

Results: The associated functional SNP rs1635 from the study in the Han Chinese population in *NKAPL* is rarer in the Irish population, with a 4.2% allele frequency in cases and no alleles in controls. A higher burden of rare variation was observed in the cases with 59 alleles observed in cases and 24 alleles observed in controls. Permutation testing of this difference yielded a p-value = 0.0246, suggesting rare variation across the entire loci may contribute to schizophrenia risk.

Discussion: We demonstrated detection of rare variation in a potential schizophrenia risk loci. There is an emerging pattern of rare variation in the form of rare variants occurring more frequently in schizophrenia cases in the *NKAPL* locus. We are currently sequencing *NKAPL* in a total of 500 Irish cases and 500 ethnically matched controls. Discovering excess functional rare variation in *NKAPL* provides independent evidence implicating this gene as a schizophrenia risk locus containing both common and rare variation contributing to risk.

2600F

Exome sequencing of singleton trios and quads to reveal *de novo* mutations in schizophrenia. *A.C. Watts, T. Walsh, M.K. Lee, S. Gulsuner, A.M. Thornton, J. McClellan, MC. King, Genes in Schizophrenia (GENESIS) Consortium.* University of Washington, Seattle, WA.

Schizophrenia is a debilitating neuropsychiatric disorder with a worldwide prevalence of about 1%. Although the heritability of schizophrenia is high, there are many sporadic cases; that is, affected individuals with no family history of mental illness. We are sequencing persons with schizophrenia in the absence of family history of mental illness (probands), their unaffected parents, and whenever possible, an unaffected older sibling, in order to identify *de novo* mutations that may be responsible for the disorder. Our hypothesis is that *de novo* events will be observed in many different genes, and that the functions of the impacted genes will differ between probands and their unaffected siblings. From exome sequence, we filter variants to include only those predicted to be both *de novo* and damaging to the protein product. Next, we validate all candidate variants passing these filters by Sanger sequencing, using diagnostic PCR primers to amplify genomic DNA of probands, siblings, and parents. Results to date suggest that approximately 50% of probands and siblings carry at least one damaging *de novo* mutation. Thus far, every mutation has been found in a different gene. The analysis of pathways in which the genes converge is in progress. Next, we will target these candidate genes and determine their complete variant profiles in 5000 cases and 5000 controls. These results can help us identify critical pathways that are affected in schizophrenia.

2601W

Novel Missense Mutations and Partial Deficiency of Collagen VI in Ullrich Congenital Muscular Dystrophy. *J. Yuan[1], I. Higuchi[1], Y. Sakiyama[1], Y. Inamori[1], E. Matsuura[1], Y. Higuchi[1], A. Yoshimura[1], A. Hashiguchi[1], K. Higashi[1], Y. Koreeda[2], K. Arimura[3], H. Takashima[1].* 1) Department of Neurology and Geriatrics, Kagoshima University 8-35-1 Sakuragaoka, Kagoshima City, Kagoshima 890-8520, Japan; 2) Division of Respiratory Medicine, Respiratory and Stress Care Center, Kagoshima University 8-35-1 Sakuragaoka, Kagoshima City, Kagoshima 890-8520, Japan; 3) Division of Neurology, Okatsu Hospital 3-95 Masagohonmachi, Kagoshima, 890-0067, Japan.

Objective To establish a diagnostic procedure and find new molecular mechanism in the patients with early-onset multi-joint contracture/overextension and muscle weakness. Methods A group of 9 patients were selected. Histochemical and immunohistochemical studies with collagen VI, collagen IV, laminin α2, adhalin, dysferlin, dystrophin-C, biglycan, perlecan, and HSP47 were performed on biopsied skeletal muscle. Using a next-generation sequencer (GS Junior platform), all coding regions and flanking intron sequences of 8 candidate genes including COL6A1, COL6A2, COL6A3, COL14A1, COL15A1, SERPINH1, BGN, and FMOD were screened, then suspected mutations were confirmed by Sanger sequencing. Results Histochemical stains indicated dystrophic changes, and no definite abnormality was observed in immunohistochemical stains. The high-throughput pyrosequencing revealed 67 single nucleotide polymorphisms or synonymous variants, and two novel heterozygous missense mutations, c.1499G>A (p.500G>E) in COL6A1 and c.801G>T (p.267K>N) in COL6A2 from patient 1 and 3, respectively. Patient 1, 28-year-old male, was suffered with proximal muscle weekness, atrophy, Achilles tendon contracture, and laxity of the distal joints. Patient 3, 25-year-old male, showed general loss of muscle strength, multiple joint contracture (foot, knee, elbow and wrist), and overextension in distal interphalangeal joints. Using a confocal microscopy, partial deficiency of collagen VI in basal lamina was revealed in double-immunofluorescent staining with collagen VI and IV. Conclusions We suggest high-throughput sequencing for genotyping the patients with early-onset multi-joint contracture/overextension and muscle weakness. We also demonstrated two novel missense mutations in collagen VI genes, and rare partial deficiency of collagen VI on basal lamina in two patients with relatively benign phenotype of Ullrich congenital muscular dystrophy.

2602T

Searching for the genes of Tourette's: The Tourette International Collaborative Genetics (TIC Genetics) Study. *G.A. Heiman[1,2], T.V. Fernandez[3], P.J. Hoekstra[4], A. Dietrich[4], R.A. King[3], M.W. State[5], J.A. Tischfield[1,2], the TIC Genetics Team.* 1) Dept. of Genetics, Rutgers University, Piscataway, NJ, USA; 2) Human Genetics Institute of New Jersey, Rutgers University, Piscataway, NJ, USA; 3) Yale Child Study Center, Yale University School of Medicine, New Haven, CT, USA; 4) Department of Psychiatry, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands; 5) Department of Genetics, Yale University School of Medicine, New Haven, CT, USA.

The Tourette's International Collaborate Genetics (TIC Genetics) Study unites a group of highly experienced clinicians specializing in Tourette's disorder (TD) with statistical and molecular geneticists with a strong record of collaboration across the United States, Europe, and South Korea. The goal is to conduct a comprehensive genomics study of TD concentrating on the potential contribution of rare and structural variants. We will ascertain over 1500 individuals over three years, focusing on the identification of TD pedigrees with 3 or more affected members and family-based association trios for pedigree-based gene discovery. Blood for DNA, LCLs and iPSCs are being collected as are PAXgene tubes for eventual RNA analyses. The DNA and cell lines will become part of the National Institute of Mental Health (NIMH) Repository at Rutgers University as an international resource. The samples will be initially analyzed by TIC Genetics Study researchers and later made public by the NIMH to qualified scientists around the world to study how genes cause TD. We will present the structure and clinical assessments of the TIC genetics collaboration. In addition, we will present results of the initial multiplex families using a gene discovery method of first genotyping all members of a family with one million SNP probes, using Illumina HumanOmni1-Quad microarrays, to detect copy number variants and to perform an affected-only parametric linkage analysis. Subsequently, we performed whole-exome sequencing of the two most distantly related affected individuals within a family using Illumina GAIIx to assess for shared disruptive variants within the identified linkage peaks. Variants are analyzed against the RefSeq hg19 gene definitions and determined to be nonsense, missense, splice-site, or indels.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2603F

Evidence for involvement of *EIF4G1*, not *VPS35* variations in Parkinson Disease. *J. Vance[1,2,3], K. Nuytemans[1,2], G. Bademci[1,2], V. Inchausti[1,2], S. Zuchner[1,2,3], A. Dressen[1,2], C. Jauregio[1,2], D.D. Kinnamon[1], A. Mehta[1,2], Y. Pasco[1], A. Aviram[1], A. Diaz[1], L. Wang[1,2,3], F. Nahab[2,4], C. Singer[2,4], W. Hulme[1], I. Konidari[1], Y. Edwards[1,2], G. Beecham[1,2,3], E.R. Martin[1,2,3], W.K. Scott[1,2,3].* 1) John P. Hussman Institute for Human Genomics, University of Miami, Miami Florida, USA; 2) Udall Center, University of Miami, Miami Florida, USA; 3) Dr. John T. Macdonald Foundation Department of Human Genetics, University of Miami, Miami Florida, USA; 4) Department of Neurology, Miller School of Medicine, University of Miami, Miami Florida, USA.

INTRODUCTION: Currently 29 genes are reported to be related to Parkinson Disease (PD) through causal mutations or susceptibility factors. It is important for both clinical diagnosis and research to clarify the mode of action (causal or susceptibility factor for PD) of the individual variants in these genes. Recently, *VPS35* and *EIF4G1* have been suggested to act as causal PD genes. Whole exome sequencing (WES) allows for quick, parallel analysis of multiple genes in a large number of individuals. MATERIAL AND METHODS: We performed WES in 213 PD patients and 272 control individuals. Those variants in *VPS35* and *EIF4G1* with <5% frequency in the exome variant server database and our own control data were considered for analysis. We performed single-variant analysis as well as gene-based joint tests of association in the PD cases and controls. RESULTS: We identified three novel VPS35 missense variations in three PD cases. Two were identified in multiplex families and did not segregate with PD. In *EIF4G1*, we identified 11 rare variants (9 nonsynonymous rare variants (nsRV) and 2 small indels), including the reported causal mutation p.R1205H. This variant p.R1205H segregated in all affected individuals in a large family, but also in one unaffected 86 year old family member, suggesting that its mode of action may not be causal with complete penetrance, but rather a strong risk factor. Several novel nsRV were identified in multiplex families but did not segregate with disease. Considering all rare variants and all rare synonymous variants (<5%) in *EIF4G1*, we found evidence of association between the gene and PD (p=0.06 / p=0.02 respectively). DISCUSSION: No strong evidence for a contribution of *VPS35* variability to PD development could be identified in the patient and control groups. In contrast, we confirmed one variant of*EIF4G1* as a very strong risk factor for PD, but WES in our samples demonstrates that other *EIF4G1* variants cannot be considered as having the same major effect. Further data is needed to strengthen the role of these other *EIF4G1* variants in PD.

2604W

Revealing the complex nature of a monogenic disease: exome sequencing of Rett syndrome. *F. Ariani[1], M.A. Mencarelli[1,2], E. Grillo[1], L. Bianciardi[1], I. Meloni[1], C. Di Marco[1], C. Lo Rizzo[1], F. Mari[1,2], A. Renieri[1,2].* 1) Medical Genetics, Dept Biotechnology, Univ Siena, Siena, Siena, Italy; 2) Medical Genetics, Azienda Ospedaliera Universitaria Senese, Siena, Italy.

Diseases previously thought to be monogenic, may be more complex due to interactions among genes, the contribution of environmental factors, incomplete penetrance and expression variability. Phenotypically discordant patients with the same mutation may be used to test the contribution of expression variability. We used “whole exome sequencing” (Illumina platform) to analyze the functional portion of the genome of two pairs of sisters with Rett syndrome. Although each pair of sisters had the same mutation in MECP2, the gene associated with this syndrome, one pair of sisters could not speak or walk and had a profound intellectual deficit (classical Rett syndrome), while the other pair of sisters could speak and walk and had a moderate intellectual disability (Zappella variant). We show that in addition to the MECP2 mutation each patient had hundreds of “private” variations. Three-hundred of these mutations were predicted to impair protein function. Seventy-six damaging variations were exclusive to the classical girls; many were in genes correlated with ATP synthesis, oxidative stress, myopathy, and intellectual disability and/or autism. Sixty damaging variations were exclusive to the Zappella girls; many were in genes associated with immune system modulation (e.g., interleukin and chemokine receptors). These results suggest that the final phenotype in the patients may be the result of mutations in the main gene (MECP2) in combination with other mutations. Specific arrays of mutations may alter clinical outcomes related to speaking, walking and intellectual capability.

2605T

***De novo* indels in Autism Spectrum Disorders.** *S. Dong[1,2], S. Sanders[1], A.J. Willsey[1], M. Murtha[1], N. DiLullo[1], L. Wei[2], M. State[1].* 1) Program on Neurogenetics, Child Study Center, Department of Psychiatry, Department of Genetics, Yale University School of Medicine, New Haven, CT. USA; 2) Center for Bioinformatics, State Key Laboratory of Protein and Plant Gene Research, College of Life Science, Peking University, Beijing, China.

Autism spectrum disorder (ASD) is a highly heritable syndrome characterized by defects in communication, social interaction, and the presence of a restricted or repetitive range of interests or behaviors. Recent research has demonstrated that *de novo* mutations, including large copy number variations (CNVs) and loss of function single nucleotide variations (SNVs), are strongly associated with ASD. Furthermore, recurrent *de novo* mutations have been used to identify specific risk-associated loci such as 16p11.2 microdeletions/duplications and the gene SCN2A. Small *de novo* insertions and deletions (indels) can be predicted with whole-exome sequencing, however this is more technically demanding process than detecting *de novo* SNVs. Using Dindel (Albers *et al.* 2011) we have developed a pipeline for the reliable detection of such *de novo* indels and confirmed the accuracy of this technique with PCR and Sanger sequencing. We present the use of this pipeline for the comprehensive analysis of indels from whole-exome sequencing of 800 individuals from 200 simplex ASD families with discordant sibling pairs. A higher rate of *de novo* indels was observed in affected individuals compared with their unaffected siblings (1.27 × 10-9 vs. 0.38 × 10-9; OR 3.00 (95% CI: 1.09–8.73); P=0.018). The majority of *de novo*indels observed in affected individuals resulted in a reading frame shift and predicted premature translation termination (17 out of 22). No such difference in the rate of *de novo* indels between probands and siblings was observed for non-coding regions (0.78 × 10-9 vs. 0.62 × 10-9; OR 1.15 (95% CI: 0.37–3.60); P=0.79). A *de novo* frameshift indel was observed in DYRK1A (Dual specificity tyrosine-phosphorylation-regulated kinase 1A), a gene in the Down's syndrome critical region on chr21; this observation, alongside two previously identified *de novo* loss of function mutations in ASD cases suggests that this locus is strongly associated with ASD (P<0.005). In summary, we demonstrate that *de novo* indels, identified through whole-exome sequencing, are associated with ASD and can be used to identify genes mediating ASD risk including DYRK1A.

2606F

Resequencing and promoter methylation analysis of the dopamine D2 receptor gene in schizophrenia. *Y. Chuang[1], M. Cheng[1], C. Lu[2], M. Tsai[3], Y. Shen[4], S. Hsu[1], C. Chen[3,5,6].* 1) Department of Psychiatry, Yuli Veterans Hospital, Hualien, Taiwan; 2) Department of Psychiatry, Hualien Armed Forces General Hospital, Hualien, Taiwan; 3) Division of Mental Health and Addiction Medicine, Institute of Population Health Sciences, National Health Research Institutes, Zhunan, Taiwan; 4) Department of Psychiatry, Tzu-Chi General Hospital-Hualien, Hualien, Taiwan; 5) Department of Psychiatry, Chang Gung Memorial Hospital at Linkou and Chang Gung University School of Medicine, Taoyuan, Taiwan; 6) Institute of Medical Sciences, Tzu-Chi University, Hualien, Taiwan.

Background: Aberrant dopamine D2 receptor (DRD2) function has been implicated in the pathogenesis of schizophrenia, as all the antipsychotics bind to dopamine D2 receptor. The study aimed to examine whether there are genetic variants in the DRD2 gene associated with schizophrenia. Methods: We resequenced the promoter and all the exons of the DRD2 gene in 560 Han Chinese patients with schizophrenia from Taiwan and conducted a case-control association analysis. We also conducted a methylation analysis of partial promoter region of the DRD2 gene of genomic DNA prepared from peripheral blood lymphocytes in 50 patients with schizophrenia and 50 control subjects using pyrosequencing method. Results: We identified 13 known SNPs in this sample, but SNP-based association analysis did not detect an association of these SNPs with schizophrenia. We also identified a rare variant (g.113346421T>C) that was located at the transcription factor AP-2 binding site of the putative core promoter region of the DRD2 gene in 3 patients and 1 control subject. Reporter gene assay showed that the mutant allele has significantly lower activity than the wild type. In the methylation analysis, we found significantly increased methylation of two CpG residues at the GCF and MAZ transcriptional factor binding sites in 50 schizophrenic patients compared to 50 control subjects. Conclusion: Our study suggests that genetic variants of the DRD2 gene appear not to play a major role in conferring susceptibility to schizophrenia in our population. In contrast, change of methylation status of the promoter of the DRD2 gene in patients with schizophrenia warrants further study.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2607W

A comprehensive genetic analysis of 421 Japanese patients with Charcot-Marie-Tooth disease using DNA microarray chips. *A. Hashiguchi, A. Yoshimura, Y. Higuchi, T. Nakamura, S. Tokunaga, Y. Okamoto, H. Takashima.* Neurology and Geriatrics, Kagoshima University Graduate School of Medical and Dental Science, Kagoshima, Japan.

[Objective] To perform genetic diagnosis in Japanese Charcot-Marie-Tooth disease (CMT) patients, and identify the molecular epidemiology in Japan, we created DNA microarray chips with 28 CMT-causative genes identified until 2004, and performed a genetic test in which CMT-causative genes could be identified fast and at low cost. [Methods] We conducted a genetic test in 421 Japanese patients with CMTfrom April 2005 to the end of 2011. Patients with median nerve motor conduction velocities less than 38 m/s and those with median nerve motor conduction velocities more than 38 m/s were classified into a demyelinating type and an axonal type, respectively. Demyelinating type CMT was conditional on absence of PMP22 gene duplication through the FISH method. Genetic abnormalities detected by a test with DNA microarray chips were confirmed with the Sanger method. When new genetic mutations were detected, they were identified as genes responsible for CMT by a segregation study. [Results and Discussion] Electrophysiologic and genetic classification of the 421 patients before genetic diagnosis showed demyelinating type CMT in 106 patients, axonal type CMT in 198, CMTX in 24, and unclassified type CMT in 93. Because patients with CMT1A due to PMP22 gene duplication among those with demyelinating type CMT were excluded, the number of patients with CMT1B due to MPZ gene mutations was 8, the largest number of patients. In the patients with axonal type CMT, the number of those with CMT2A2 due to MFN2 gene mutations was 17, the largest number of patients, and 9 of the patients with CMTX had GJB1 gene mutations. Some of the patients with demyelinating type CMT and axonal type CMT also showed GJB1 gene mutations. It has become possible to analyze 28 causative genes for CMT within a minimum of two days. The positive rate of genetic diagnosis of CMT is as low as 10 to 20%, and many causative genes other than the 28 genes can be considered to exist. In the future, a comprehensive genetic diagnosis system using next-generation sequencing technology is expected. [Conclusion] An analysis of DNA microarray chips enabled a well-known comprehensive CMT screening test, which led to the confirmation of distribution of genetic abnormalities in Japanese patients with CMT. Further improvement of genetic diagnosis systems is necessary. A comprehensive genetic diagnosis system using next-generation sequencing technology is expected.

2608T

A new frameshif mutation in the dopamine receptor D4 exon III VNTR 7-repeat allele. *D.S.S. Lobo[1, 2], M. Tampakeras[1], N. Freeman[1], J.L. Kennedy[1, 2].* 1) Neuroscience Dept., Neurogenetics Lab., Centre for Addiction and Mental Health, Toronto, Ontario, Canada; 2) Dept. of Psychiatry, University of Toronto, Toronto, Ontario, Canada.

A 48bp VNTR in exon III of the DRD4 gene has been implicated as an important genetic factor in psychiatric disorders. Several studies, including LaHoste et al.(1996), have shown a significant association between the 7-repeat allele (7R) and attention-deficit hyperactivity disorder (ADHD). Previously, studies have reported substantial sequence variation within the exon III VNTR repeat motif and across the various alleles characterized on the basis of length (Lichter et al 1994, Ding et al. 2002). In a recent study of DRD4 exon III VNTR sequence variation in subjects with ADHD, Tovo-Rodrigues et al.(2012) found that the 7R allele had significantly more sequence variation than the 4-repeat allele (4R), and suggested that the DRD4 VNTR region be routinely examined for both size polymorphisms and sequence variants in studies of ADHD and other impulse control disorders. Our objective was to sequence the DRD4 exon III VNTR region for new sequence variants and to investigate which of the previously reported variations within the 48-bp motif were present in a sample of impulsive subjects diagnosed with pathological gambling. Methods: We sequenced 156 individuals who were previously genotyped as 2R(N=4), 4R(N=141) or 7R(N=11) homozygotes to assess the amount of sequence variation in this region using standard Sanger sequencing. Results: Sequencing the DRD4 region in our sample yielded similar results to Ding et al.(2002) and Tovo-Rodrigues et al.(2012); individuals homozygous for the number of repeat motifs showed sequence variation within motifs and in the arrangement of the motifs within alleles characterized previously by the number of repeats. Of particular interest was a 1bp heterozygous deletion observed in 2 subjects homozygous for the 7R allele. In the first subject the 21st bp of the 1st repeat motif was deleted. In the second subject the 21st bp of the 3rd repeat motif was deleted. In both cases the deleted nucleotide was an adenine. Although the deletions are located in 2 different repeat units, both deletions occur at the same position within the repeat units. In silico analysis showed that this deletion causes a premature stop codon and truncated predicted amino acid sequence. No deletions were found in 2R or 4R homozygotes. The high frequency of this frameshift mutation in 7R homozygotes (2/22 alleles= 9%) is noteworthy and may explain, at least in part, conflicting results in the association of this particular allele with impulse control disorders.

2609F

SPG4 mutations in Brazilian patients with hereditary spastic paraplegia. *I. Lopes-Cendes[1], D.B. Dogini[1], C. Cartaxo[4], S. Raskin[4], H.A.G Teive[4], J. Saute[3,6], L.B. Jardim[3,5,6], M.C. França Jr[2].* 1) Department Medical Genetics, Faculty of Medical Sciences, University of Campinas - UNICAMP, Campinas, SP, BRAZIL; 2) Department of Neurology, Faculty of Medical Sciences, University of Campinas - UNICAMP, Campinas, SP, BRAZIL; 3) Post-graduation program in Medical Sciences, Universidade Federal do Rio Grande do Sul, RS, BRAZIL; 4) Department of Internal Medicine,Universidade Federal do Paraná, Curitiba, PR, BRAZIL; 5) Department of Internal Medicine, Universiade Federal do Rio Grande do Sul. Porto Alegre, RG, BRAZIL; 6) Medical Genetics Service, Hospital de Clinicas de Porto Alegre, RS, BRAZIL.

Hereditary spastic paraplegia (HSP) belongs to a group of neurodegenerative diseases that represent a significant number of patient visits to a neurogenetics clinic. With a prevalence of 3.8 cases:100,000. HSP has a predominant manifestation of progressive weakness associated with spasticity in the lower limbs and a strong genetic substrate. The pattern of inheritance may be autosomal dominant (AD), autosomal recessive (AR) or X-linked. The objective of this study is to determine the frequency of SPG4-HSP among Brazilian patients with AD-HSP and to identify mutations associated with the disease in this population. We recruited 34 families with AD-HSP at the neurogenetics clinics of three university hospitals; patients underwent a complete clinical evaluation as well as detailed family history. Genomic DNA was extracted from lymphocytes and used in PCR reactions with primers designed to cover the 17 exons of the SPG4 gene. Mutation screening was performed by automatic sequencing and MLPA. We found five potential deleterious mutations in SPG4: c.1413 +5 G> A (splicing-site mutation), c.1255G> T (nonsense), c.1267G> T (missense), c.1849T> G (nonsense) and c.839delA (frameshift). Each variant was found in only one single family. Among the variants found, the mutation c.1255G> T, is reported here for the first time. We found SPG4 mutations in approximately 15% of the families with AD-HSP studied; this represents a smaller proportion compared to the frequency reported in Europe and the USA.

2610W

Changes in the Human Transcriptome Caused by LMNB1 Duplication in a Case of Autosomal Dominant Leukodystrophy: An RNASeq study. *P. Cherukuri[1], D. Simeonov[1], K.V. Fuentes-Fajardo[1], P. Zumbo[2], C. Mason[2], S. Lin[3], Y.H. Fu[3], C. Boerkoel[1], T. Markello[1], W. Gahl[1], C. Toro[1].* 1) Undiagnosed Diseases Program, NHGRI / NIH, Bethesda, MD; 2) Department of Physiology and Biophysics, Weill Medical College, Cornell University, New York, NY; 3) Department of Neurology, University of California, San Francisco, San Francisco, California.

Autosomal dominant leukodystrophy (ADLD) is a rare, progressive, adult-onset demyelinating disorder caused by duplications of the nuclear lamina gene, lamin B1 (LMNB1). LMNB1 is part of the stable fibrous meshwork of intermediate filaments (IFs) that underlie the nuclear membrane, forming the lamina. In addition, LMNB1 is tethered to chromatin at lamina-associated domains and has a role in the regulation of gene expression and gene silencing. In this study, we investigated the global impact of LMNB1 locus duplication on the expression profile of the poly-adenylated RNA fraction of the human transcriptome using massively parallel sequencing (RNASeq). 100ng of high-quality total RNA (RIN>0.8) for each experimental condition was isolated from primary fibroblast cell lines, and cDNA was prepared with the TruSeq mRNA prep kit (v2). Multiplexed cDNA libraries for each triplicate sample (2 samples (1 case & 1 normal); 6 datasets) were sequenced to 50×50 bp on the Illumina HiSeq2000, and post-processed with CASAVA (v1.8.2). Raw data were filtered for high-median quality (Q-value > 20) and then a total of 426 million paired-end (PE) reads (~71 million PE reads per-replicate) were processed on a High Performance Computing (HPC) cluster (SGE) through different RNASeq analysis pipelines (PPBS and ERANGE). Gene expression was quantified as reads-per-kilobase-per-million (RPKM). To test for dysregulation or disruption of gene-expression profiles of clusters of genes in genomic neighborhoods, we developed a computational approach to detect movement of genes' transcription outside of a stable boundary condition. Using this approach, we found that ~80 highly reproducible gene-clusters were significantly dysregulated. These clusters were significantly enriched (P-value=4.7e-29; hypergeometric distribution) for chromatin regions associated with interaction with the lamina. We are pursuing exploring the mechanism underlying these alterations in expression and evaluating genes within these boundaries for robustness of differential expression. This study presents a hypothesis for the mechanism underlying the neurodegeneration.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2611T

Exome sequencing of young onset sporadic amyotrophic lateral sclerosis trios. *A.E. Renton[1], J.R. Gibbs[1], J.O. Johnson[1], M.A. Nalls[1], C. Sassi[1], Y. Abramzon[1], C. Edsall[1], G. Restagno[2], M. Sabatelli[3], A. Chiò[4], B.J. Traynor[1].* 1) Neuromuscular Diseases Research Unit, Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD; 2) Molecular Genetics Unit, Department of Clinical Pathology, A.S.O. O.I.R.M.-S. Anna, Turin, Italy; 3) Institute of Neurology, Catholic University of the Sacred Heart, Rome, Italy; 4) Department of Neuroscience, University of Turin, Turin, Italy.

Amyotrophic lateral sclerosis (ALS) is a fatal neurodegenerative disease. Known pathogenic mutations explain ~65% of familial ALS (FALS) and ~10% of sporadic ALS (SALS). In order to identify *de novo* and recessive mutations putatively underlying SALS, we undertook exome and dideoxy sequencing in a cohort of 19 disease trios from Italy. Each trio consisted of two unaffected parents and one offspring diagnosed with young onset SALS. Gene validation was performed using exome and dideoxy sequence data from 146 unrelated FALS probands. All novel variants were filtered against 1000 Genomes, 150 in-house population exome controls, and 5600 population controls obtained from the Exome Sequencing Project cohort. We confirmed eleven *de novo* and eight recessive non-synonymous mutations in nine trios. Almost all of these were single base alterations, and seven trios carried multiple changes. Two genes exhibiting single *de novo* mutations each harboured two heterozygous non-synonymous mutations in two FALS cases. These findings suggest that we have identified one or more novel ALS genes, and demonstrate the power of the trio-exome paradigm for elucidating the genomic architecture of complex disorders.

2612F

PRRT2 gene missense mutations account for both paroxysmal kinesigenic dyskinesia and infantile convulsions in a Chinese pedigree. *C. Cai[1], W.D. Li[2].* 1) Department of Surgery, Tianjin Children's Hospital, Tianjin, Tianjin, China; 2) Center of Basic Medical Sciences, Tianjin Medical University, Tianjin, China.

Paroxysmal kinesigenic dyskinesia is an autosomal dominant dystonia induced by sudden voluntary movements. Recently, proline-rich transmembrane protein 2 gene mutations, especially frameshift mutations, were described for PKD. Methods: A three-generation paroxysmal kinesigenic dyskinesia -infantile convulsions pedigree was collected in Tianjin, North China. The symptoms of six patients varied; age of onset decreased in each generation. Mutations in the proline-rich transmembrane protein 2 gene in nine PKD family members were screened by PCR sequencing of genomic DNA samples. Results: Missense mutations of the proline-rich transmembrane protein 2 gene were found in all four PKD patients and two children with infantile convulsions. All six individuals carried heterozygous codon 138 (Pro/Ala) and codon 306 (Ala/Asp) mutations. Conclusions: Mild mutations of the proline-rich transmembrane protein 2 gene other than truncate and frameshift mutations were account for PKD and/or infantile convulsions. Age of onset and symptoms were not necessarily associated with proline-rich transmembrane protein 2 mutations.

2613W

Whole-exome sequencing and homozygosity analysis to dissect autism, a complex genetic disorder. *M.H. Chahrour[1,2,3,4,5], T.W. Yu[1,2,3,4,5,6], E.T. Lim[5,7,8], B. Ataman[9], M.E. Coulter[1], R.S. Hill[1,2,3], C.R. Stevens[5], C.R. Schubert[1,2,3,4,5], ARRA. Autism Sequencing Collaboratio[10], M.E. Greenberg[9], S.B. Gabriel[5], C.A. Walsh[1,2,3,4,5,6].* 1) Division of Genetics, Department of Medicine, Children's Hospital Boston, Harvard Medical School, Boston, MA; 2) Manton Center for Orphan Disease Research, Children's Hospital Boston, Boston, MA; 3) Howard Hughes Medical Institute; 4) Department of Pediatrics, Harvard Medical School, Boston, MA; 5) Program in Medical and Population Genetics, Broad Institute of Massachusetts Institute of Technology and Harvard University, Cambridge, MA; 6) Department of Neurology, Harvard Medical School, Boston, MA; 7) Biological and Biomedical Sciences Program, Harvard University, Boston, MA; 8) Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA; 9) Department of Neurobiology, Harvard Medical School, Boston, MA; 10) The ARRA Autism Sequencing Consortium.

Autism is a neurodevelopmental disorder characterized by impaired communication skills, social behavior abnormalities, and stereotypies. Despite the clear genetic component of autism, extreme genetic heterogeneity has made it difficult to identify causative genes. We quantified runs of homozygosity in probands from 1000 families from the Autism Genetic Research Exchange (AGRE) collection, in order to identify those with potential recessive mutations due to distant shared ancestry. Whole-exome sequencing of the top sixteen probands with the highest percent of homozygosity (~1–9%) in their genomes revealed homozygous, potentially pathogenic recessive mutations in four families. The mutations were extremely rare (absent from the 1000 Genomes project, dbSNP130, NHLBI ESP, and a database of 831 exomes), validated, and segregated perfectly with disease. We screened our candidate genes in a larger independent cohort of whole exome data from 418 autism cases and 371 controls, sequenced as part of the ARRA Autism Sequencing Consortium, and identified a modest excess of rare recessive variants (homozygous or compound heterozygous) in cases (24/418, 5.7%) versus controls (11/371, 3.0%) (P = 0.042, Fisher's exact test, one-tailed). The novel candidate genes (*UBE3B*, *CLTCL1*, *NCKAP5L*, *ZNF18*) encode proteins involved in proteolysis, GTPase-mediated signaling, cytoskeletal organization, and other pathways. We show that the transcription of these genes is regulated by neuronal depolarization, suggesting potential activity-dependent roles in neurons. One of the identified mutations is a homozygous missense change in the HECT domain of UBE3B, ubiquitin protein ligase E3B, a member of the E3 ubiquitin-conjugating enzyme family that is highly expressed in the brain. Its paralog *UBE3A* is disrupted in Angelman syndrome, a neurodevelopmental disorder characterized by intellectual disability, movement or balance problems, abnormal behaviors, and speech and language impairment. Our strategy of combining homozygosity analysis with whole-exome sequence data identifies candidate recessive mutations in autism, and may have broader applicability to other complex, heterogeneous disorders.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2614T

Gene-Based Rare Variant Analyses Reveal Loci Associated with Autism Spectrum Disorder. *N. Dueker[1], E.R. Martin[1,2], A.J. Griswold[1], H.N. Cukier[1], S. Slifer[1], J. Jaworski[1], I. Konidari[1], P.L. Whitehead[1], M. Schmidt[1], D.J. Hedges[1,2], D. Martinez[1], S. Clarke[1], M.L. Cuccaro[1,2], J.R. Gilbert[1,2], J.L Haines[3], M.A. Pericak-Vance[1,2].* 1) John P. Hussman Institute for Human Genomics, University of Miami, Miami, FL; 2) Dr. John T. Macdonald Department of Human Genetics, University of Miami, Miami, FL, USA; 3) Vanderbilt University School of Medicine, Vanderbilt University, Nashville, TN, USA.

Background: Autism spectrum disorder (ASD) is a developmental disorder characterized by social, communicative and behavioral impairments leading to average yearly medical costs that are 4.1–6.2 times higher for those with an ASD compared to those without. Despite demonstrating high heritability estimates, only a small proportion of the genetic risk for ASD is explained. Rare variants are hypothesized to have large effect sizes and may contribute to ASD risk. Therefore, we analyzed rare variant data to identify loci associated with ASD. Methods: Participants were drawn from a large, family-based study of ASD and included 944 unrelated cases and 467 controls that were genotyped using the Illumina HumanExome-12v1 Array. This array includes coding region variants across the genome, with a majority of variants (88%) being rare (MAF <5%). We performed gene-based analyses where we tested autosomal genes genotyped on the Array for association with ASD. The sequence kernel association test (SKAT) was used for these analyses, with covariate adjustment for the first three principal components. We also tested 21 pathways previously associated with ASD in copy number variant (CNV) analyses, for association in our study using SKAT. Results: Study participants were primarily male (73.1%) and of self-reported White race (87.4% of cases and 82.4% of controls). Mean age between cases and controls was similar (16.0y for cases and 15.4y for controls). Of the 16,906 genes analyzed, 15 showed p-values <0.001 in the total sample analyses. Two of these genes fall within pathways associated with ASD in CNV analyses, *PODN*(p=0.0002) and *GPSM3* (p=0.0003). In race-stratified analyses, 19 genes showed p-values <0.001 among Whites and 31 genes showed p-values <0.001 among Blacks. None of the 21 tested were associated with ASD. Discussion: Autism spectrum disorders are complex traits, known to be influenced by several genes. Rare variants are thought to exert large influences on disease risk therefore making their study of great importance. In our analysis of rare variants, we identified several genes of interest. Two moderately associated genes fall within two pathways associated with ASD in CNV analyses;*PODN* which falls within a cell motion pathway (GO:0006928) and *GPSM3* which falls within a GTPase activator activity pathway (GO:0005096). These results suggest that rare variants play a role in ASD.

2615F

Data sharing in the genomic age: Genome Variant Database for Neuromuscular Diseases. *M. Gonzalez, R.F. Acosta Lebrigio, S. Zuchner.* John P. Hussman Institute for Human Genetics and Genomics, Unviersity of Miami Miller School of Medicine, Miami, FL.

Many rare neuromuscular diseases (NMDs) show an impressive Mendelian locus and allelic heterogeneity. Yet, known genes still explain only a fraction of many phenotypes. New genomic approaches, such as exome sequencing have significantly increased our ability to identify the remaining, often very rare genes. To further facilitate the study of rare genes we have identified the need for a data-sharing platform for exome/ genome sequencing results. Such a platform requires: 1) ease of access to allow the participation of a wide range of investigators (geneticists, biologists, physician scientists), 2) data protection to prevent non-intended mass downloads and genetic profiling, 3) central data processing under a standard pipeline to guarantee high data quality, and 4) openness to the participation of multiple investigators. We have developed the Genome Variant Database for Neuromuscular Diseases (GVD-NMD), which represents an online public database containing thoroughly annotated de-identified whole exome data from a number of neuromuscular disorders. Our current focus is on inherited disorders of the longest neuronal axons, which are clinically categorized under hereditary spastic paraplegias (HSP) and inherited peripheral neuropathies, also known as Charcot-Marie-Tooth disease (CMT). Currently, GVD-NMD contains 245 samples and hopes to encompass 500–750 exomes by the end of 2012. A conservative set of filters has been applied to limit the number of rare variants to those of interest to the majority of rare variant Mendelian studies. Current filters are as follows: minor allele frequency in NHLBI EVS <1%, minor allele frequency in GVD-NMD <5%, GERP score > 0 OR PhastCons score > 0.3, GATK quality scores (DP ≥ 10, QUAL ≥150, GQ ≥75), and segregation in available families. To date GVD-NMD contains 77,238 variants (SNVs and indels) in 14,926 genes, which represent 51,812 unique variants (48,783 SNVs, 3,029 indels). We see three principle uses of GVD-NMD: 1) human geneticists discover rare variants that further support their top candidate genes, 2) basic scientists identify disease associated variants in genes of functional interest, and 3) physician scientists and medical geneticists identify rare alleles in known genes not yet reported. We believe that sharing of genome-scale data in innovative ways will foster collaborations and lead to more efficient identification of novel disease alleles and genes.

2616W

X-linked Charcot-Marie-Tooth disease type 1 in Japan: genetic, clinical, and electrophysiological study of 33 cases. *Y. Higuchi, A. Yoshimura, T. Nakamura, S. Tokunaga, A. Hashiguchi, Y. Okamoto, H. Takashima.* Department of Neurology and Geriatrics, Kagoshima University Graduate School of Medical and Dental Science, Kagoshima City, Kagoshima, Japan.

Objective: To investigate the clinical heterogeneity of X-linked Charcot-Marie-Tooth disease type 1 (CMTX1), which is caused by various mutations in the gap junction beta 1 (*GJB1*) gene.**Methods**: Thirty-three Japanese patients with CMTX1 were genetically, clinically, and electrophysiologically examined. All *GJB1* mutations were screened by an originally designed microarray resequencing DNA chip and confirmed by Sanger sequencing. **Results**: Thirty-three patients showed mutations in GJB1. The mutations included 14 missense mutations, one nonsense mutation, four deletions, and one insertion. Six mutations were not reported previously: N2T, S42C, L76R, K103N, S138R, and V192D. Heterozygous female were asymptomatic or less severely affected than male. The severity of the CMT clinical phenotype did not correlate with the location and type of mutation in GJB1. A distinct phenotypic variability was observed even among patients with the same genotype (e.g., R22Q). In addition to peripheral neuropathy, nine patients had symptoms and signs of central nervous system involvement, presenting with hearing loss, transient episodes of hemiparesis, ataxia, and dysarthria. One patient had chronic inflammatory demyelinating polyneuropathy. Electrophysiological study of 27 patients revealed that 14 showed demyelinating neuropathies, six showed intermediate, and seven showed axonal neuropathies. **Conclusion**: We identified six novel mutations in GJB1. The clinical phenotype of CMTX1 was variable, and significant genotype-phenotype correlation was not possible. These findings suggest that not only loss of function of GJB1 but also some gain of function may be related to the pathogenesis of CMTX1.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2617T

De novo mutations in autism spectrum disorders revealed by whole genome sequencing. *Y. Jiang[1*], X. Jin[2*], J. Yu[3,11*], R.K. Yuen[4*], D. Cao[2], M. Wang[2], S. Walker[4], L. Lau[4], D. Merico[4], Y. Shi[2], Q. Xu[6], B. Zhou[6], Z. Hu[5], B. Cheng[3], N. Chen[2], J. Mei[2], K. Xia[5], X. Xu[6], Y. Wang[6], Y. Wu[10], J. Wang[2], J. Wang[2], C. Lajonchere[7], G. Dawson[7,8], H. Yang[2], A. Shih[7], B. Devlin[9], Z. Sun[3,11#], S.W. Scherer[4§], Y. Li[2&].* 1) Department of Pediatrics Medical Genetics and Neurobiology, Duke University School of Medicine, Durham, NC; 2) BGI-Shen Zhen, China; 3) Beijing Institutes of Life Science, Chinese Academy of Sciences, Beijing, China; 4) McLaughlin Centre and The Centre for Applied Genomics, The Hospital for Sick Children and the University of Toronto, Toronto, Ontario, Canada; 5) The State Key Laboratory of Medical Genetics, Xiangya School of Medicine, Central South University, China; 6) Children's Hospital of Fudan University, Shanghai, China; 7) Autism Speaks, New York, NY; 8) University of North Carolina at Chapel Hill, NC; 9) University of Pittsburgh School of Medicine, Pittsburgh, PA; 10) Yuning Psychiatry Clinic, Growing With You Mental Health Center, Taipei 10664, Taiwan; 11) Institute of Genomic Medicine, Wenzhou Medical College, Wenzhou, China. *equal contribution; #corresponding for Chinese samples, §AGRE/TASC samples, & sequencing production.

Autism spectrum disorders (ASD) are a group of neurodevelopmental conditions characterized by impaired social interaction and communication and by restricted and repetitive behaviors. The latest statistics from the Center for Disease Control indicate one in 88 children in the United States is diagnosed with ASD with worldwide data showing similar trends. Genetic variation contributes an important role in the etiology of ASD, and both de novo and inherited variations have now been identified affecting ASD risk. Recent estimates place the number of susceptibility genes for ASD in the hundreds, and discovering these genes is an urgent goal in the field. We have initiated an international endeavor called the Autism Genome 10K-Project, which aims to sequence the genomes from 10,000 individuals belonging to 2,000 ASD families from Autism Speaks Autism Genetic Resource Exchange (AGRE) and 1000 Chinese ASD families. In an initial pilot study, whole genome sequencing (WGS) was performed in 33 father-mother-ASD child trios of European ancestry (99 individuals), and in a second phase in 33 father-mother-ASD child trios of Chinese ancestry (99 individuals). All ASD subjects were diagnosed using gold standard tools, Autism Diagnostic Interview Revised and Autism Diagnostic Observation Schedule. WGS was performed at BGI using Illumina HiSeq technology yielding ~30X coverage used for the initial analyses. For the first 33 families of European ancestry our first pass analysis used the GATK program to detect SNP and INDEL variants. For each individual, ~4.2M putative SNPs and ~0.89M putative INDELs were identified, of which 1% were annotated in exonic intervals of the genome. The analysis of copy number variants is in progress. We detected roughly 1 de novo SNP in the exome of the individuals with ASD, findings consistent with other (mainly exome sequencing) studies. By scanning across the genome we can infer 400 candidate de novo SNPs, on average, in each trio (we expect this number to be diminished by additional filtering, which is in progress). We are currently experimentally validating all de novo exonic variants and interesting rare inherited variants affecting relevant coding and non-coding segments of the genome. A comparative genome analysis of Caucasian and Chinese ASD families is planned. We will present our complete data analysis from this pilot phase of project, as well as additional progress in the Autism Genome 10K-Project.

2618F

Analysis of low-frequency, protein altering variation in 13,000 individuals from a Swedish schizophrenia cohort on the exome array. *B.M. Neale[1,2,3], J. Goldstein[1,2], C. O'Dushlaine[3], M. Fromer[1,3,4], G. Genovese[3], J.L. Moran[3], K. Chambert[3], C.M. Hultman[5], P. Sklar[4], S. Purcell[1,2,3,6], M.J. Daly[1,2], P.F. Sullivan[7], S. McCarroll[3,8].* 1) Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA; 2) Program in Medical and Population Genetics, the Broad Institute of Harvard and MIT, Cambridge, MA; 3) Stanley Center for Psychiatric Research, the Broad Institute of Harvard and MIT, Cambridge, MA; 4) Division of Psychiatric Genomics, Mount Sinai School of Medicine, New York, NY; 5) Department of Medical Epidemiology, Karolinska Institutet, Stockholm, Sweden; 6) Deparment of Psychiatry, Harvard Medical School, Boston, MA; 7) Department of Genetics, Psychiatry, and Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill,; 8) Deparment of Genetics, Harvard, Boston, MA.

The recently developed exome chip array enables the assessment of low-frequency coding variation. Containing approximately 250,000 variants drawn from whole genome and exome sequencing of 12,000 individuals, this array enables assessment of ~200,000 missense, ~5,500 nonsense, and ~10,000 splice site mutations. Consequently, we assess the role of coding variation in a large population derived sample of schizophrenia. We genotyped over 13,000 individuals drawn from the Swedish Schizophrenia Study on the Illumina Infinium HumanExome BeadChip. Using a subset of samples for whom exome sequencing and exome chip genotype data are available, we designed zCall to improve low-frequency variant genotyping. Based on comparisons with sequencing, ~99% of variants with a singleton on the chip (i.e. only one heterozygote observed) in 947 individuals were called with a single heterozygote in the correct individual. Similarly, approximately 75% of the coding variation within an individual was captured. The overwhelming majority of the variation on the assay is low-frequency, with ~46% of the sites invariant and another ~25% variable sites having fewer than 10 copies. We identified a significant subset of the sample of apparent Finnish ancestry enabling the identification of low-frequency functional variation significantly increased in Finland due to the limited founder population. We present single SNP association analysis using both logistic regression with principal components and a mixed model incorporating genome-wide similarity. Beyond single locus association, we also describe regional association analyses, combining evidence for association across a gene, testing for recessive disease models, and critically evaluating loss of function alleles. While our analysis is provisional as of the date of submission of this abstract, the results to date suggests that at most a few (and possibly no) protein-altering alleles influence schizophrenia in Sweden with an effect size that makes them detectable at genome-wide significance in a cohort of 13,000. The exome array allows large scale analysis of functional variation. The data generated here demonstrate that this assay performs well and captures most of an individual's variation. Furthermore, these results bound how much standing coding variation influences schizophrenia in Sweden and provide context for previously reported genome-wide association results, partly addressing the question of synthetic association.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2619W

Search for rare-variant risks of Parkinson's disease by sequencing of candidate genes and exome sequencing. *W. Satake[1], Y. Suzuki[2], Y. Ando[1], T. Nishioka[1], K. Syoji[1], H. Tomiyama[3], M. Yamamoto[4], M. Murata[5], N. Hattori[3], S. Sugano[2], T. Toda[1].* 1) Div of Neurology, Kobe Univ Grad Sch Med, Kobe, Japan; 2) Dept of Med Genome Sci, the Univ of Tokyo, Japan; 3) Dept of Neurology, Juntendo Univ Sch Med, Tokyo, Japan; 4) Dept of Neurology, Kagawa Pref Cent Hosp, Takamatsu, Japan; 5) Dept of Neurology, Natl Cent Hosp of Neurology and Psychiatry, Kodaira, Japan.

Parkinson's disease is one of the most common neurodegenerative diseases worldwide, mainly manifesting motor impairment due to degeneration of dopamigenic neurons. Heterozygous mutations in the GBA gene, whose homozygous mutations cause one of lysosomal disorders Gaucher's disease, have been reported as strong risk factors for Parkinson's disease. In addition, some patients with Gaucher's disease have been reported to manifest parkinsonism as a part of Gaucher's symptoms. Here, we examined whether other causative genes for lysosomal diseases or diseases showing parkinsonism can be genetic risks for Parkinson's disease. We selected 13 genes as candidates of genetic risks for Parkinson's disease; 8 lysosomal disease (GM1-gangliosidosis, GLB1; Tay Sacks, HEXA; Sandhoff, HEXB; Fabry, GLA; Metachromatic leukodystrophy, ARSA; Krabbe, GALC; Niemann-Pick, NPC1; Adrenoleukodystrophy, ABCD1) and 5 diseases with parkinsonism in part of the symptoms (Wilson, ATP7B; Creutzfeldt-Jacob, PRNP; FTDP-17, MAPT; Hallervorden-Spatz, PANK2; Choreacanthocytosis, CHAC). We further selected 34 exons from these 13 genes, which have major and high-frequent causative mutations for each disease, including 714 known causative mutations and variants, and re-sequenced them by Sanger method using DNA samples of 541 Parkinson's disease cases (500 sporadic and 41 familial cases). In our patients' sample, the mutations and variants were detected in 7 genes; HEXB, GLA, ARSA, GALC, ATP7B, PRNP, and MAPT. We tested allele frequencies between cases and controls by x2 statistics; however no significant association was obtained in each SNV; P= 0.934, 0.808, 0.359, 0.097, 0.059, 0.502, and 0.154 respectively. Moreover, we tested exome-sequensing for a pair of an affected sibling in order to extract candidate SNVs as Parkinson's disease-risks. We extracted 50Mb exome by Agilent Sureselect and performed massive parallel sequencing using Hiseq2000. After BWA mapping, GATK SNV calling, and dbSNPs filtering, we detected novel 690 and 694 SNVs in each case sample. Of these, 321 SNVs were matched in the sibling, which may contain candidates for risk variants for PD. Furthermore, to identify rare-variant risks, we are now performing exome-sequencing of several hundreds of Parkinson's disease cases.

2620T

Rare combination of inherited mutations in a large multigenerational pedigree with autism spectrum disorder (ASD) and co-morbid neurodevelopmental disorders. *P. Szatmari[1], O. Migita[2], I. O'Connor[1], A. Prasad[2], D. Sato[2], S. Walker[2], A. Thompson[1], S.W. Scherer[2], D. Pinto[2,3].* 1) Dept Psychiatry, Chedoke Site, McMaster Univ, Hamilton, ON, Canada; 2) The Centre for Applied Genomics and Program in Genetics and Genome Biology, The Hospital for Sick Children, Toronto, ON, Canada; 3) Departments of Psychiatry, and Genetics and Genome Sciences, Mount Sinai School of Medicine, NY, USA.

Autism spectrum disorders (ASD) are a heterogeneous group of neurodevelopmental disorders with a complex inheritance pattern. Rare copy number variations (CNVs) and sequence-level mutations have been identified as etiologic factors in ASD. Here we report the genetic analyses of a large multigenerational family composed of an index case presenting autism, low functioning and epilepsy. Initial whole-genome SNP array and whole-exome resequencing analyses found that the index case carries two inherited rare variants in keeping with the `multiple-hit model' for ASD: a maternally inherited *NRXN1* premature stop mutation not seen in dbSNP, 1000GP and 6,000 unrelated controls, and a paternally inherited 15q11.2 (BP1-BP2) deletion. Neither of these rare variants was seen in the two unaffected siblings. The *NRXN1* gene encodes an adhesion protein important in synapse development/function, and has been associated with neurodevelopmental disorders. The 15q11.2 deleted region has been associated with a number of neurological disorders, including epilepsy, and contains the *CYFIP1* gene that codes for a protein that binds FMRP, the protein missing in fragile X syndrome. Clinical history and phenotype assessment are being collected for both maternal and paternal sides of the extended family. At least 8 additional nuclear pedigrees with autism across four-generations have already been identified. On the father's side, there are also two kinships with epilepsy and a kinship with Tourette syndrome. To evaluate the segregation of these rare variants in the extended family, we further genotyped nuclear families with affected children with Illumina 1M or 2.5M arrays, and whole-genome sequencing is planned for two kinships. We expect that integration of the various genetic analyses coupled with deep phenotyping will aid in establishing genotype-phenotype correlations.

2621F

Resequencing of 17 genes in the N-methyl-D-aspartate (NMDA) glutamate system to identify rare variants affecting risk of substance dependence. *P. Xie[1,4], H.R. Kranzler[5], J.H. Krystal[2,4], L.A. Farrer[6], H. Zhao[1,3], J. Gelernter[1,2,4].* 1) Departments of Genetics, Yale University School of Medicine, New Haven, CT; 2) Departments of Psychiatry, Yale University School of Medicine, New Haven, CT; 3) Department of Epidemiology and Public health, Yale University School of Medicine, New Haven, CT; 4) VA CT Healthcare Center, West Haven, CT; 5) Department of Psychiatry, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA; VISN 4 MIRECC, Philadelphia VAMC, Philadelphia, PA; 6) Departments of Medicine (Biomedical Genetics), Neurology, Ophthalmology, Genetics and Genomics, Epidemiology, and Biostatistics, Boston University School of Medicine and Public Health, Boston, MA.

As for most common traits, common variants found by association studies explain only a small fraction of the heritability of substance dependence (SD) risk. In this study, we aimed to identify rare variants in genes important in the NMDA glutamate receptor system, and explore their possible relationship to alcohol, cocaine and opioid dependence (AD, CD, and OD). Glutamate is the major excitatory neurotransmitter in the brain. NMDA receptors are known to play important roles in SD. We selected 17 genes involved in the system: three NMDAR subunit genes (*GRIN1*, *GRIN2A*, and *GRIN2B*), ten genes encoding protein kinases, phosphatase and palmitoyltransferase that perform post-translational modifications of NMDAR (*FYN*, *SRC*,*PRKCG*, *PTPN5*, *PPP1CA*, *PPP1CB*, *PPP1CC*, *PPP1R1B*, *PPP1R9B*, and *ZDHHC3*), and four genes encoding proteins that regulate NMDAR function (*DISC1*, *NRG1*, *DAOA*, and *DTNBP1*). We used a pooled sequencing design (20 subjects of the same sex and population and with phenotypes of AD, OD, and CD in each pool; Illumina HiSeq), and sequenced the coding regions and the flanking intronic regions with high coverage in 760 cases with all three substance dependence diagnoses and 760 healthy controls. A total of 454 variants, including 380 rare variants (minor allele frequency ≤ 1% in both EAs and AAs) were identified. The transition/transversion ratio was 3.50, consistent with results from the 1000 Genomes Exon Pilot Project, indicating likely low false-positive variant identification. Based on case-control allele count differences, we selected 17 exonic rare variants and genotyped them using the TaqMan method in 6751 subjects for follow-up and the 1520 subjects from the sequencing stage for validation. The 17 rare variants were found in at least three more case pools than control pools, or vice versa. Eleven of them were successfully genotyped, and all variants called in the sequencing stage were validated. Using the Sequence Kernel Association Test (SKAT; Wu M.C., *et al.* 2011) in the 8271 subjects, we found a statistically significant cumulative association between the 11 rare variants and OD in AAs (P=0.00080). Results from gene-based association tests showed that the association signal derived mostly from *DISC1* (P= 0.0010) and *GRIN2B* (P=0.00085). The three *DISC1* rare variants investigated were largely limited to AAs. We conclude that rare variation in the set of genes we selected contributes significantly to OD risk in AAs and potentially in other populations.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2622W

Widespread novel RNA editing in human brain tissue, identified by RNA-Seq. *L. Hou[1], N. Akula[1], J. Wendland[1], D.T. Chen[1], X. Jiang[1], K. Choi[2], B.K. Lipska[3], J.E. Kleinman[3], F.J. McMahon[1].* 1) Human Genetics Branch, National Institute of Mental Health Intramural Research Program, National Institutes of Health, US Dept. of Health and Human Services, Bethesda, MD USA; 2) Department of Psychiatry, Uniformed Services University of the Health Sciences, Bethesda, MD USA; 3) Section on Neuropathology, Clinical Brain Disorders Branch, NIMH-IRP, National Institutes of Health, US Dept of Health and Human Services, Bethesda, MD USA.

RNA editing is a post-transcriptional process that alters the sequence of primary RNA transcripts. Previous studies have shown that RNA editing occurs in the brain and may be a key regulator of neural development. Dysregulation of RNA editing might also be associated with some psychiatric disorders. Recent advances in high-throughput sequencing technology have made it possible to study RNA editing on a genome-wide scale. In this study, poly(A)$^+$ RNA extracted from postmortem dorsolateral prefrontal cortex of 11 unaffected controls and 10 individuals with bipolar disorder (BP) was sequenced on an Illumina GA-IIx or HiSeq 2000 platform. Paired-end reads were aligned to the reference human genome (hg19) by GSNAP, with all known splice sites provided (based on RefSeq Genes). Single nucleotide variants (SNVs) were called with GATK Unified Genotyper. To reduce false positive calls from paralogous sequences, only uniquely mapped reads were used. Additionally, a series of filters based on read depth, strand bias, mapping quality, and read position were used to remove false positive calls stemming from mapping or sequencing errors. The remaining SNVs were filtered against known variants in dbSNP135, the 1000 Genomes Project, and the NHLBI Exome Sequencing Project. The mRNA levels of the adenosine deaminase encoding genes, ADAR1, ADAR2, and ADAR3, were quantified by Cufflinks. On average, 9516 ± 248 RNA editing sites were identified in each individual, about 10% of which could be found in the DAtabase of RNa EDiting (DARNED). Most of the putative RNA editing sites were canonical A-to-I editing sites, and were located in 3'-UTR, consistent with known RNA editing sites. Individuals with BP did not differ in the total number of RNA editing sites or in the proportion of identified sites in DARNED. Overall, the proportion of edited reads was lower in individuals with BP. Consistent with this, we observed decreased expression of ADAR2 in brain tissue from individuals with BP (p=0.008 and p=0.038 in the GA-IIx and HiSeq data, respectively). This study shows that RNA-seq data can be used to identify RNA editing sites in postmortem brain, these sites appear to be widespread across the genome, and many are novel. Very stringent data analysis methods are required to avoid false calls, especially when as in this study DNA sequencing data are not available. Decreased ADAR2 expression, if replicated, suggests that RNA editing might be involved in BP.

2623T

Association of the polymorphisms of ITPK1 gene in inositol metabolic pathways with NTDs. *Z. Guan, JH. Wang, J. Guo, F. Wang, XW. Wang, GN. Li, Q. Xie, X. Han, B. Niu, T. Zhang.* Department of biotechnology, Capital Institute of Pediatrics, beijing, CH., China.

Neural tube defects (NTDs) are common and severe malformations that is multifactorial, involving the combined action of both genetic and environmental factors. Maternal nutritional status(myo-inositol) is associated with the occurrence of NTDs. Myo-inositol and related genetic factors are suggested to be implicated in the etiology of NTDs. Inositol, 1,3,4-trisphosphate 5/6-kinase(ITPK1) is a very important enzyme in inositol metabolism. To investigate the role of genetic variants in ITPK1 and NTDs risk, a case-control study of women with NTD-affected pregnancies (n=200) and controls (n=328) from a high-risk area for NTDs in China was carried out. The study protocol was reviewed and approved by the ethic evaluation committee of Capital Institute of Pediatrics, Beijing, China, and we obtained written informed consent from all participants. DNA of maternal peripheral blood was extracted. The 13 Snps of ITPK1 were genotyped by the Sequenom MassArray system, and EMSA(Electrophoretic mobility shift assay) was applied to verify the binding activity between wild and mutated oligonucleotides probes. We found that 4 SNPs (rs4586354, rs2236131, rs3783903, rs170689) were statistically associated with NTDs occurrence (p<0.05). After stratifying participants by NTD phenotypes, the significant association only existed in cases with spina bifida. The result of the EMSA showed a different allelic binding capacity of activator protein-1 in the intron region of the ITPK1 gene, which is affected by an A→G exchange at rs3783903 (OR=2.72, 95% CI=1.33–5.54, p<0.005). The results suggested that the polymorphism in ITPK1 may be a genetic risk factor for NTDs in a high-risk area of China. Zhen Guan, Jian-Hua Wang, contributed equally to this manuscript.

2624F

High apolipoprotein E4 allele frequency in FXTAS patients. *M. Mila[1,2,3], F. Silva[1,4], I. Madrigal[1,2], MI. Alvarez-Mora[1,2], R. Oliva[1,3,5], L. Rodriguez-Revenga[1,2].* 1) Hospital Clinic, Barcelona, Spain; 2) CIBER de Enfermidades Raras (CIBERER), Barcelona, Spain; 3) IDIBAPS (Institut d'Investigacions Biomediques August Pi I Sunyer), Barcelona, Spain; 4) Fundacio Clinic per a la Recerca Biomedica, Barcelona, Spain; 5) Faculty of Medicine. University of Barcelona, Barcelona, Spain.

Fragile X-associated tremor/ataxia syndrome (FXTAS) is a late-onset neurodegenerative disorder that occurs in *FMR1* premutation carriers. I it is well known that the ApoE ε4 allele is a risk factor for neurodegenerative disease. The main goal of this work was to evaluate the ApoE genotypes and allelic distribution among FXTAS patients. A total of 44 unrelated *FMR1*premutation carriers (22 presenting with FXTAS and 22 non-FXTAS) have been genotyped. All the ApoE ε4/4 genotype carriers detected (100%) and 6 out of the 7 ApoE ε4/3 genotype carriers (85.7%) are patients presenting with FXTAS symptoms, while only 40% of the ApoE ε3/3 genotype carriers belong to the FXTAS group. The results showed that the frequency of the ApoE ε4 allele in the FXTAS group is statistically significant when comparing with non-FXTAS group (P = 0.023). On the basis of these results, we conclude that the presence of at least one ApoE ε4 allele might act as a genetic factor predisposing to develop FXTAS.Acknowledgments We would like to thank the Fragile X families, the “Associació Catalana Síndrome X fràgil”, and the Federación Española del Síndrome X Frágil for their cooperation. This work was supported by FIS (PS09/00413) from Instituto de Salud Carlos III, AGAUR SGR1337 from Generalitat de Catalunya and Fondo Europeo de Desarrollo Regional (Fondos FEDER) and FSE (SFRH/BD/81271/2011). The CIBER de Enfermedades Raras is an initiative of the ISCIII.

2625W

Generation and characterization of induced Pluripotent Stem (iPS) cells from patients with the autosomal dominant disorder Spinocerebellar Ataxia Type 1 (SCA1). *N. Mollema[1], A. Shanmugam[2], E. Leathley[1], D. Kaufman[2], H.T. Orr[1].* 1) Institute for Translational Neuroscience, University of Minnesota, Minneapolis, MN; 2) Stem Cell Institute, University of Minnesota, Minneapolis, MN.

Generation of induced Pluripotent Stem (iPS) cells is a viable way to study human disease processes in an *in vitro* culture system, and iPS cells have been created from a variety of neurodegenerative disease patients. Spinocerebellar Ataxia Type 1 (SCA1) is a neurodegenerative disease that results from a polyglutamine expansion in the gene *ATXN1*. The expanded protein is inherited in an autosomal dominant manner, and results in late-onset degeneration of cerebellar Purkinje cells. Mutant ATXN1 protein is stabilized by phosphorylation at location S776, and must enter the nucleus to cause disease. We set out to generate iPS cells from SCA1 patients for the long-term goal of utilizing these cells in the drug screening for inhibitors of ATXN1 expression, and gene expression studies. In these studies, human dermal fibroblasts from SCA1 patients and unaffected controls were infected with a lentivrial vector containing the four classic transcriptional reprogramming factors (*OCT4*, *SOX2*, *C-MYC*, *KLF-4*). The infected cells were cultured in fibroblast media for six days and then transferred to mouse fibroblast feeder layers in iPS cell media. Colonies emerged after approximately 26 days in culture and were manually selected. Cells were assessed by immunohistochemistry, and karyotyping. Human embryonic stem (hES) cells (NIH approved line H9) were used as controls for these experiments. hES and iPS cells were differentiated to a mitotic population of neural epithelial cells, and these cells were maintained in culture for further differentiation. SCA1 patient and control iPS lines showed hallmark morphology of stem cells, and stained for markers of pluripotency. Mutant protein did not appear to inhibit the reprogramming process, although some variations in *ATXN1* repeat length were noted after cellular reprograming. We determined that ATXN1 is not expressed in hES, iPS cells, or subsequently differentiated neuroepithelial cells. We are in the process of differentiating human ES- and iPS-generated neuroepithelial cells along a ventral/midbrain pathway, to generate cell types where ATXN1 is normally expressed *in vivo*. These studies will allow us to determine when ATXN1 expression turns on during the process of differentiation from stem cell to neuron, and identifying when the protein is phosphorylated.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2626T

Quality of Life in Machado-Joseph Disease/Spinocerebellar Ataxia Type 3 (MJD/SCA3) is related to depressive and extracerebellar disease features. *J.A. Saute[1,2], R. D Ávila[2], R.M. Castilhos[2], T.L. Monte[3], A.F.S. Schuh[4], K.C. Donis[2], G.N. Souza[2], A.D. Russo[2], T.C. Gheno[4], G.V. Furtado[4], M.L. Saraiva-Pereira[2,4,5], C.R.M. Rieder[1,3], L.B. Jardim[1,2,4,6].* 1) Post-Graduate Program in Medical Sciences - UFRGS, Porto Alegre, Rio Grande do Sul, Brazil; 2) Medical Genetics Service, Hospital de Clínicas de Porto Alegre, Rio Grande do Sul Brazil; 3) Neurology Service, Hospital de Clínicas de Porto Alegre, Rio Grande do Sul Brazil; 4) Post-Graduate Program in Molecular Biology and Genetics - UFRGS, Porto Alegre, Rio Grande do Sul, Brazil; 5) Biochemistry Department - UFRGS, Porto Alegre, Rio Grande do Sul, Brazil; 6) Internal Medicine Department - UFRGS, Porto Alegre, Rio Grande do Sul, Brazil.

MJD/SCA3, caused by a CAG repeat expansion (CAGn) at ATXN3 gene, is an untreatable condition characterized by a relentless deterioration of cerebellar and other neurological functions. Few studies have explored the quality of life (QoL) of MJD/SCA3 patients. Aim: to describe the main factors influencing patient's QoL through the WHOQoL-Bref questionnaire. Methods: molecularly diagnosed MJD/SCA3 patients were recruited from our Neurogenetics Outpatient Clinic in South Brazil. WHOQoL-Bref and Barthel score (both with higher scores indicating better performance); Neurological Examination Score for Spinocerebellar Ataxia (NESSCA), Scale for the Assessment and Rating of Ataxia (SARA) and Beck Depression Inventory (BDI) - all crescent in severity - were applied. Results: 62 patients (30 males) were included, with the following characteristics (m±sd): 40.5±9 years of age; age at disease onset (AO) of 34.4±8.8 years; disease duration (DD) of 6±2.5 years; CAGn of 75.4±3; NESSCA of 14.3±4.7; SARA of 10.8±4; Barthel index of 94.2±8; BDI of 12.3±9; and WHOQol-Bref - total score of 68.3±10.5. WHOQol-Bref total score and its four sub-scores were not associated with DD, AO, CAGn or SARA scores. WHOQol-Bref total score was correlated to NESSCA (R=-0.323, p=0.01). NESSCA was then divided in its cerebellar (up to 14 points) and extra-cerebellar items (up to 26 points): the correlation with WHOQol-Bref was maintained only by the extra-cerebellar items (R=-0.35, p=0.005). When WHOQol-Bref subdomains were analyzed, the association with extra-cerebellar items of NESSCA was evidenced with physical health (R=-0.36, p=0.005) and environment domains (R=-0.30, p=0.016). WHOQol-Bref total score was strongly associated to BDI (R=-0.71, p=0.0001), specially its psychological domain. Finally, the physical health domain of WHOQol-Bref was associated with severity of gait ataxia (R=-0.282, p=0.02) and with Barthel index (R=-0.284, p=0.02) Discussion: As expected, quality of life is reduced in MJD/SCA3, even in early stages of disease. The losses seemed to be mainly related to depressive manifestations and to extra-cerebellar dysfunctions as measured by NESSCA, with only a minor role for ataxic features. These findings point to the importance of depressive and extra-cerebellar manifestations of this disease. Acknowledgments: FIPE-HCPA, FAPERGS, CNPq, INAGEMP.

2627F

A health non penetrant pedigree derives from a late onset Parkinsonism founder index case with SCA2 and SCA8 expansions. *V. Volpini, H. San Nicolas, J. Corral, L. De Jorge, B. Campos.* Center For Molecular Genetics Diagnosis (CDGM), Idibell, L' Hospitalet De Llobregat, Barcelona, Spain.

Late onset cerebellar ataxias are a group of severe neurodegenerative disorders with an estimated prevalence in our population of 1 in 100,000. Ataxia with progressive course is the most characteristic sign of these disorders. Age at onset varies from 5 to 70 years, but it usually begins around 30 to 40y. Spinocerebellar ataxia type 2 (SCA2) is characterized by progressive cerebellar ataxia, including nystagmus, slow saccadic eye movements and, in some individuals, ophthalmoparesis or Parkinsonism. Pyramidal findings are present; deep tendon reflexes are brisk early on and are absent later in the course. Age of onset is typically in the fourth decade with 10 - to 15 - year disease duration. SCA2 expanded alleles with moderate size presenting CAA interruptions are associated with late onset Parkinsonism. Is generally known that SCA2 expanded alleles without CAA interruptions are associated with pathological development of the disease and that the interruptions would prevent elongation of the expanded sequence. Spinocerebellar ataxia type 8 (SCA8) is caused by CAG/CTG repeat expansion, but SCA8 diagnosis is not clear due to reduced penetrance of some expanded alleles present in asymptomatic cases. In this study we show one pedigree in which segregate SCA2 expanded alleles with CAA interruptions coexisting with SCA8 expanded alleles in several members. The index and founder case shows a SCA2 interrupted expanded allele of 38 CAG repeats coexisting with an SCA8 allele of 89 repeats. He presents late onset essential tremor and cerebellar atrophy. No mutations were found in LRRK2 and Parkin genes associated with Parkinson's disease to explain his etiology. One of his sons and two of his grandchild's share the same expanded interrupted SCA2 allele, coexisting with SCA8 expansions from 78 to 84 repeats, and all of them remain asymptomatic. In conclusion, SCA2 expanded alleles in pathological range with CAA interruptions could be associated with late onset Parkinsonism with uncertain penetrance. The role of SCA8 expansion coexisting with SCA2 expanded alleles remains unknown.

2628W

Impaired TFEB signaling accounts for autophagy pathway dysfunction in the neurodegenerative proteinopathy X-linked spinal & bulbar muscular atrophy. *C. Cortes[1], Y. Batlevi[1], J.E. Young[2], H. Frankowski[1], B.L. Sopher[3], G.A. Garden[3], A.R. La Spada[1,2,4].* 1) Pediatrics, University of California, San Diego, San Diego, CA; 2) Cellular & Molecular Medicine, University of California, San Diego, San Diego, CA; 3) Neurology, University of Washington, Seattle, WA; 4) Neurosciences, University of California, San Diego, San Diego, CA.

Autophagy is a highly conserved pathway of cellular macromolecule and organelle turnover, with a recently defined role in degrading aggregate-prone proteins. Autophagy dysfunction occurs in many neurodegenerative disorders, but its role in polyglutamine (polyQ) disease remains unclear. To address this, we studied autophagy in the neurodegenerative proteinopathy X-linked spinal & bulbar muscular atrophy (SBMA), a rare motor neuron disorder caused by a polyQ repeat expansion in the Androgen Receptor (AR). We found that primary neurons expressing AR-polyQ displayed altered lysosomal staining, but maintained robust induction of autophagy. Biochemically, we observed a significant decrease in the levels of cathepsin D, a key lysosomal enzyme, and a concomitant decrease in long-term protein turnover and lysosomal function in MN-1 AR66 cells (a classic in vitro SBMA model). Expression of full-length AR led to impaired autophagosome maturation, and defects in autophagosomal turnover were detected. Ultrastructural analysis of lumbar spinal cord sections from SBMA transgenic mice revealed that upregulation of functional autophagy in motor neurons correlated with maintenance of normal neural function (during the presymptomatic stage). However, emergence of disease in these mice was associated with a block in autophagy pathway progression. To determine the basis of autophagy pathway dysregulation in SBMA, we considered a role for Transcription Factor E-B (TFEB), a glutamine-tract containing master regulator of lysosomal function and autophagy. We found an interaction between AR and TFEB by co-immunoprecipitation, and observed AR polyQ length-dependent repression of a TFEB response element (4X CLEAR) promoter-reporter in cell co-transfection assays. This effect was not due to impaired trafficking of TFEB into the nucleus, but rather occurred after both transcription factors were activated and translocated into the nucleus. Overexpression of TFEB rescued impaired autophagic flux in AR66 cells, and restored 4X CLEAR transactivation in the face of lysosomal stress. Our results thus implicate AR-mediated TFEB transcriptionopathy in autophagy pathway dysregulation in SBMA, and pinpoint altered CLEAR signaling pathway function as a pathological turning point in SBMA and potentially in other related neurodegenerative proteinopathies.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2629T

The histone demethylase *KDM5C* gene is a direct target of the ARX homeobox transcription factor. *L. Poeta[1,2], F. Fusco[1], D. Drongitis[1], C. Shoubridge[3,4], G. Manganelli[1,5], A. Padula[1], S. Filosa[1,5], M. Courtney[6,7], P. Collombat[6,7], M.B. Lioi[2], J. Gecz[3,4], M.V. Ursini[1], M.G. Miano[1].* 1) Institute of Genetics and Biophysics "Adriano Buzzati Traverso", CNR, Naples 80131, Italy; 2) University of Basilicata, Potenza 85100, Italy; 3) Department of Paediatrics, University of Adelaide, Australia, South Australia 5006, Australia; 4) Department of Genetics and Molecular Pathology, SA, Pathology, Adelaide, Australia; 5) IRCCS, Neuromed, Pozzilli 86077, Italy; 6) Universitè de Nice-Sophia Antipolis, FR-06108, Nice, France; 7) Inserm U1091, IBV, Diabetes Genetics Team, FR-06108 Nice, France.

Intellectual Disability (ID) and Epilepsy often occur together with a dramatic impact on the development and quality of life of the affected children. Mutations of the *Aristaless-related homeobox* gene (ARX) cause a spectrum of X-chromosome phenotypes with ID (XLID) and various forms of malignant pediatric epilepsy, including infantile spasms. Located at Xp22.13,*ARX* is a homeobox gene encoding a paired-type-HomeoDomain transcription factor of crucial significance for brain ontogenesis. We show that *lysine (K)-specific demethylase 5C*(*KDM5C/JARID1C/SMCX*), a known XLID gene involved in chromatin remodeling, is regulated by ARX. By in vitro assays, we demonstrate that ARX directly regulates and binds the 5' region of the *KDM5C* gene. We studied ARX mutant proteins carrying various PolyAlanine expansion mutations, identified in XLID/Epilepsy patients. They are in frame duplications affecting two out of the four ARX PolyAlanine amino acid tracts and represent disease hot spot sites with a strong tendency to add coding trinucleotide repeats. We prove that the changes in the repeat units in the PolyAlanine tracts cause hypomorphic *ARX* mutations, which exhibit a decreased trans-activating activity and a reduced binding to the *KDM5C* regulatory region. Furthermore, the altered functioning of the ARX PolyAlanine mutants tested seems to correlate with the degree of XLID/Epilepsy severity. We found a down-regulation of the *Kdm5C* mRNA and protein levels in the murine Arx KO neuron-oriented ES cells. Additionally, we observed that *Arx* KO produced an increase of the Kdm5C-mediated levels of H3K4me3, potentially due to a compromised activity of Kdm5C. In summary, we discover a new ARX-dependent disease path and suggest that the molecular pathogenesis of the *ARX* PolyAlanine mutations may be in part caused by an aberrant histone demethylation as a result of the KDM5C defect. Even though many other ARX targets could have important roles in the expansion phenotype, a further development of our findings could open up the possibility of additional studies aimed to elucidate the symptoms associated with the *ARX* PolyAlanine-related neurophenotypes.

2630F

15q11.2 microdeletion and *FMR1* premutation in a family with intellectual disabilities and autism. *L. Rodriguez-Revenga[1,2], I. Madrigal[1,2], M. Mila[1,2,3].* 1) Hospital Clinic, Barcelona, Spain; 2) CIBER de Enfermedades Raras (CIBERER), Barcelona, Spain; 3) IDIBAPS (Institut d'Investigacions Biomèdiques August Pi i Sunyer). Barcelona, Spain.

Genomic rearrangements of chromosome 15q11-q13 are responsible for diverse phenotypes including intellectual disabilities and autism. 15q11.2 deletion, implicating common PWS/AS breakpoints BP1–BP2, has been described in patients with delayed motor and speech development and behavioural problems. Here we report the clinical and molecular characterization of a maternally inherited BP1–BP2 deletion in two siblings with intellectual, motor and speech delay, autistic syndrome disorder and several dysmorphic features. One of the patients was also a carrier of an *FMR1* allele in the low premutation range. The four genes within the deletion were underexpressed in all deletion carriers but *FMR1* mRNA levels remained normal. Our results suggest that BP1–BP2 deletion could be considered as a risk factor for neuropsychological phenotypes and that it presents with variable clinical expressivity. Acknowledgments: This work was supported by FUNDACIÓN ALICIA KOPLOWITZ (IP: Laia Rodriguez-Revenga) and AGAUR SGR1337 from Generalitat de Catalunya. The CIBER de Enfermedades Raras is an initiative of the ISCIII.

2631W

FRAXE molecular diagnosis in individuals referred for FRAXA screening. *A. Javed[1], G. Ali[1], L. Caicedo[1], I. Marques[2], R. Santos[2], P. Jorge[2].* 1) R& D, Gene Link, Inc, Hawthorne, NY; 2) UI&D, Centro de Genetica Medica Dr. Jacinto de Magalhaes - Porto, Instituto Nacional de Saude Dr. Ricardo Jorge, INSA I.P., Porto, Portugal.

FRAXE mental retardation is a form of mild to moderate intellectual disability generally associated with learning difficulties, communication deficits, attention problems, hyperactivity and autistic behavior. FRAXE (AFF2/FMR2 gene) a folate-sensitive fragile site in Xq28 ~600 kb distal to the FRAXA (FMR1 gene) site is the most common form of inherited mental retardation. Molecular characterization revealed that individuals expressing FRAXE had amplifications of a CCG repeat adjacent to a CpG island. Normal individuals showed 4–39 copies of the polymorphic FRAXE CCG repeat, while individuals expressing the fragile site had >200 copies and their CpG island was fully methylated. These findings are similar to those found for folate-sensitive fragile X site FRAXA. Reports of FRAXE full expansions and pre-mutations are rarely documented. In this respect, it has been very difficult to determine to what extent the alleles, with CCG repeats in the range of 36 and 199, have a pathogenic effect. Intellectually disabled individuals are primarily referred for FRAXA screening and individuals who are negative for FRAXA are possible candidates for FRAXE screening. Traditionally in some laboratories AFF2 molecular analysis is performed by PCR, it is known that CCG repeats in the range of ~80 and above are not reliably amplified. We embarked on an effort to supplement our PCR analysis by Southern blot and cloned a segment of the AFF2 gene that can be used by appropriate labeling as a probe to determine expansion of the CCG repeats in the AFF2 gene. We have developed a probe to be used for Southern blot analysis that reliably detects the AFF2 CCG triple repeat amplification. We present data of AFF2 molecular analysis in a subpopulation of 5,000 individuals referred for FRAXA screening. The presence of pre-mutated and fully expanded alleles in either gender, were confirmed by Southern blot analysis, which also enabled exclusion of methylation or repeat number mosaics as well as PCR failure. We recommend the use of this probe as suitable for genotyping of pre-mutations, full mutations, and mosaics specifically for individuals presented for FRAXA screening with negative results to determine FRAXE status.

2632T

An isogenic, epi-isoautosomal fibroblast model to investigate *FMR1* premutation-associated cellular dysfunction. *K.A. Koscielska, S. Roenspie, P.J. Hagerman.* Department of Biochemistry and Molecular Medicine, University of California, Davis, School of Medicine, Davis, CA.

The *FMR1* gene contains a CGG trinucleotide repeat in its 5' UTR. The premutation alleles of the gene (55–200 repeats) give rise to several different human disorders, including primary ovarian insufficiency (POI) and the late-onset neurodegenerative disorder, fragile X-associated tremor ataxia syndrome (FXTAS). The molecular mechanism of premutation-induced cellular dysregulation is not well understood, but is a consequence of the elevated levels of the CGG-containing *FMR1* mRNA, while FMRP protein levels remain close to the normal range. Cellular dysregulation has been observed in human and mouse neurons, as well as in human skin fibroblasts, the latter readily obtained from premutation carriers.

To better understand the cellular phenotype in humans, cloned female fibroblasts were generated to establish a case-control model of the premutation in an isogenic, epi-isoautosomal (differing active X chromosomes) genetic background. Each clonal line expresses exclusively either a normal or an expanded CGG repeat, depending on whether the maternal or the paternal X chromosome is active. In addition, fibroblasts from a mosaic male patient were cloned, in order to established an isogenic cell culture model with differing CGG repeats. The CGG repeat size was shown to be stable throughout long-term cell culture, and was also unchanged by the iPSC reprogramming and neuronal differentiation. The cloned fibroblast cell lines were characterized with respect to their *FMR1* mRNA and FMRP levels, and are a promising human cell culture model for the study of premutation disorders.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2633F

Formation of an R-loop at the *FMR1* 5'UTR: Unraveling the extra thread of *FMR1* expression. *E.W. Loomis[1], F.L. Chédin[2], P.J. Hagerman[1].* 1) Department of Biochemistry and Molecular Medicine, UC Davis School of Medicine, Davis, CA; 2) Department of Molecular and Cellular Biology, UC Davis, Davis, CA.

The $(CGG)_n$ trinucleotide repeat in the 5' untranslated region (5'UTR) of the *FMR1* gene is responsible for a number of heritable disorders, albeit by remarkably distinct pathogenic mechanisms: gene silencing for fragile X syndrome; RNA toxicity for FXTAS. In our studies of expression in the "premutation" repeat range (55–200 CGG repeats), we have discovered that long, three stranded, R-loop structures form during transcription through the guanine-rich sequences within the 5'UTR region. The G-rich transcribed mRNA forms a stable RNA/DNA hybrid with the template DNA strand, displacing the coding DNA strand. Recent research demonstrates that R-loops form throughout the human genome at a majority of CpG island promoters, including *FMR1*, and play an important role in the genome. However, the consequences of CGG repeat expansion on R-loop formation have not been considered previously.

To better understand the relationship between R-loop formation and CGG-repeat size, we have begun to characterize R-loop formation in human fibroblast cell lines containing alleles across the range of transcribed expanded *FMR1* CGG repeats. Immunoprecipitation with an RNA/DNA hybrid-specific antibody (DRIP) of genomic DNA from cultured human skin fibroblasts with both normal (CGG ≈ 30) and premutation (55 < CGG > 200) repeat alleles demonstrates a significant enrichment of *FMR1* over background, indicative of R-loop formation. Additionally, non-denaturing bisulfite mapping of the RNA-displaced single-stranded DNA spanning the 5'UTR and the CGG repeats in genomic DNA reveals clear differences in secondary structure within the displaced (G-rich) DNA strand. R-loop formation through expanded CGG repeat alleles is a previously uncharacterized phenomenon at *FMR1* that is likely to be involved in the complex molecular pathology observed in association with repeat expansion.

2634W

The role of AGG interruptions, alternative splicing and bidirectional transcription in the stability and function of the *FMR1* gene. *C.M. Yrigollen[1], H. Tang[1], E. Loomis[1], G. Filippova[2], P.J. Hagerman[1, 3], F. Tassone[1,3].* 1) Biochemistry and Molecular Medicine, University of California Davis, School of Medicine, Davis, CA, United States; 2) Division of Human Biology, Fred Hutchinson Cancer Research Center, Seattle, WA, United States; 3) MIND Institute, University of California Davis, Sacramento, CA, United States.

The CGG repeat element in the 5'UTR of the *FMR1* gene normally contains 2–3 AGG interruptions, although this number decreases in expanded alleles. We have determined that the presence and number of AGG `interruptions' strongly influences the likelihood of generating a full mutation allele in a premutation mother-to-offspring transmission. This observation will have an important impact on genetic counseling of carrier women who are contemplating the risk of having a child with fragile X syndrome. In our study of 267 maternal premutation alleles, considering the number and positions of AGG interruptions, and the length of pure and total CGG repeats, we found that the presence of AGG interruptions significantly reduces the risk of transmission of a full mutation for a maternal premutation allele of fewer than ~100 CGG repeats. Additionally, alternative splicing in the coding region of the *FMR1* gene together with alternative start and polyadenylation site usage gives arise to multiple transcripts, which generate multiple protein isoforms whose altered distribution in the premutation range may have profound biological consequences. Our preliminary quantitative RT-PCR (real time PCR) and sequencing (single molecule real time (SMRT) sequencing) data indicate that some isoforms are differentially overexpressed with increasing CGG repeat number and show specific tissue expression. Finally, the antisense FMR1 mRNA (ASFMR1 mRNA) has been reported to be similarly overexpressed in premutation carriers and silenced in full mutations. We further observed that the levels of ASFMR1 mRNA are also elevated in gray zone alleles and that altered ASFMR1 splicing/expression might play a role in Fragile X associated tremor ataxia syndrome (FXTAS). Indeed, we show that the expression of a splice isoform in the ASFMR1 transcript, located near the ASFMR1 promoter is elevated in those with a premutation and particularly in those with FXTAS. This absent/minor splice isoform in normal *FMR1* alleles may represent an important outcome measure associated with FXTAS. These molecular mechanisms contribute to cell dysfunction and to the variable phenotypes associated with the *FMR1* gene particularly for alleles in the premutation range.

2635T

MutLα is required to repair short slipped-DNAs of trinucleotide repeats. *G. Panigrahi[1], M. Slean[1, 2], J. Simard[1], C. Pearson[1, 2].* 1) Dept Genetics & Genomic Biol, Hosp Sick Children, TMDT Building 101 College St., 15th Floor, Room 15–312 East Tower, Toronto, Ontario, CanadaM5G 1L7; 2) Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada.

Mismatch repair is required for proper maintenance of the genome by protecting against mutations, as well as for driving certain mutations, including disease-associated trinucleotide repeat instability. We recently revealed a role of hMutSβ in the repair of short slip-outs containing a single CTG repeat units (Panigrahi et al., 2010, PNAS, 107(28)12593–12598). The involvement of other MMR proteins in short slip-out repair is unknown. Here we show that hMutLα is required for the highly efficient in vitro repair of single CTG repeat slip-outs. HEK293T cell extracts, deficient in hMLH1, are unable to process single-repeat slip-outs, but are functional when complemented with hMutLα. The MMR-deficient hMLH1-mutant, T117M, with a point mutation proximal to the ATP-binding domain, was defective in slip-out repair; further supporting a requirement for hMLH1 in processing of short slip-outs and possibly the involvement of hMHL1 ATPase activity. The hPMS2-deficient HEC-1-A cell extracts, cells express hMLH1, hMLH3 and hPMS1, are only functional when complemented with hMutLα, indicating that neither hMutLβnor hMutLγ are sufficient to repair short slip-outs. Furthermore, clustered slip-outs are poorly repaired, and this low level of repair (<15%) appears to depend upon the presence of a functional hMutLα in addition to hMutSβ.

2636F

Regional alterations of transcriptome-wide AGO:miRNA binding profiles in Huntington's disease patient brains. *R.L. Boudreau[1], P. Jiang[2], B.L. Gilmore[1], Y. Xing[2], B.L. Davidson[1].* 1) Internal Medicine, University of Iowa, Iowa City, IA; 2) Biostatistics, University of Iowa, Iowa City, IA.

MicroRNAs (miRNAs) are small RNAs which regulate gene expression and play important roles in development and disease. Mammalian genomes encode hundreds of miRNAs which, in conglomerate, regulate up to one-third of the transcriptome. miRNAs are excised from larger stem-loop containing transcripts and incorporated into Argonaute (AGO) proteins, the core component of the RNA Induced Silencing Complex (RISC), producing functional entities capable of silencing target transcripts via sequence complementarity, often by as few as 6–7 base-pairs. To better understand the biological function of miRNAs and their contributions to disease pathogenesis, target identification is of paramount importance. Although bioinformatic prediction and wet-lab validations have identified hundreds of individual miRNA:target interactions, this represents a limited scope considering the complex nature of the miRNA regulatory network. To address this, recent efforts have focused on high-throughput mapping of transcriptome-wide miRNA binding sites. Researchers have employed a crosslinking immunoprecipitation protocol coupled with high-throughput sequence analysis (CLIP-seq) to map AGO:miRNA binding sites to specific regions within target mRNAs. Here, we utilized AGO-CLIP-seq to investigate the pathogenic role of miRNAs in Huntington's disease (HD), a dominantly-inherited late-onset neurodegenerative disease. Prior data support that mRNAs and miRNAs are dysregulated in HD patient brains, and we hypothesized that complementary AGO-CLIP-seq data would facilitate the interpretation of these changes and assist with identifying novel pathogenic mechanisms. We performed AGO-CLIP-seq on control and HD patient brain samples isolated from either primary motor cortex or cingulate gyrus, regions with transcriptional changes implicated in HD phenotypes. We have sequenced and mapped over one million unique reads in each sample, and our preliminary evaluation reveals an abundance of brain-enriched miRNAs and thousands of miRNA binding sites throughout the brain transcriptome. Our immediate near-term goal is to perform a thorough bioinformatic comparison to identify disease- and regional-associated changes in miRNA expression and mRNA targeting. These data represent the first AGO-CLIP-seq study performed in post-mortem human brain tissue, and this innovative approach may serve to elucidate new therapeutic targets for HD and other human ailments with disease-associated transcriptional changes.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2637W

Large repeat expansions in the *C9ORF72* gene contribute to a spectrum of neurodegenerative disorders including Alzheimer disease. *M.A. Kohli[1], K. John-Williams[1], R. Rajbhandary[1], A. Naj[1], P. Whitehead[1], K. Hamilton[1], R.M. Carney[1], C. Wright[2], E. Crocco[3], H.E. Gwirtzman[4], R. Lang[5], G. Beecham[1,6], E.R. Martin[1,6], J. Gilbert[1,6], M. Benatar[2], G.W. Small[7], D. Mash[2], G. Byrd[5], J.L. Haines[4], M.A. Pericak-Vance[1,6], S. Züchner[1,6].* 1) University of Miami, Miller School of Medicine, John P. Hussman Institute for Human Genomics, Miami, FL, USA; 2) University of Miami, Miller School of Medicine, Department of Neurology, FL, USA; 3) University of Miami, Miller School of Medicine, Department of Psychiatry and Behavioral Science, FL, USA; 4) Center for Human Genetics Research, Vanderbilt University Medical Center, Nashville, TN, USA; 5) North Carolina Agricultural and Technical State University, Department of Biology, Greensboro, NC, USA; 6) University of Miami, Miller School of Medicine, Dr. John T. Macdonald Foundation Department of Human Genetics, Miami, FL, USA; 7) David Geffen School of Medicine at UCLA, Semel Institute for Neuroscience and Human Behavior, Los Angeles, CA, USA.

Progressive neurodegenerative diseases comprise clinically and genetically highly heterogeneous phenotypes. For instance, *PSEN1 and 2*, *MAPT* and *GRN* mutations are associated with clinical characteristics typical of Alzheimer disease (AD), frontotemporal dementia (FTD) and/or Parkinson disease (PD). These data point towards molecular, but also clinical, commonalities, which are important for our understanding of some of the most common neurodegenerative diseases of older age. Recently, an intronic hexanucleotide $(GGGGCC)_n$ repeat expansion in the*C9ORF72* gene has been identified to account for up to 50% of FTD/ALS families, 7% - 12% of FTD families, and 2% - 5% of sporadic FTD cases in Caucasians. Given the clinical, pathological and genetic overlap between FTD and AD, we hypothesized that *C9ORF72* expansions may also confer risk to AD. We analyzed 1,268 and 292 cases with a clinical and/or pathological diagnosis of AD and 1,039 and 620 neurologically healthy controls aged over 60 from European and African American ancestry, respectively. In Caucasians, we found *C9ORF72*expansions in the pathogenic range of FTD/ALS (>30 repeats) at a rate of 0.9% in AD cases versus zero in controls (p=0.0014, 1,268 cases, 1,039 controls). This case-control association became stronger in a subsequent meta-analysis comprising 2,177 AD cases and 6,441 neurologically healthy controls from literature (p=7.5E-05). In contrast, no large expansions were detected in individuals of African-American ethnicity (292 cases, 620 controls). Moreover, in the range of normal variation of *C9ORF72* expansions (0–23 repeat copies), we detected significant differences in distribution and mean repeat counts between Caucasians and African-Americans. These differences were unexpected, but likely reflect ancient haplotype distributions, including haplotypes at risk for a repeat expansion. *C9ORF72* expansions in the normal range did not significantly confer risk to AD in either ethnicity when compared to elderly controls. Most relevant, evaluation of AD cases with large *C9ORF72* repeat expansions revealed a phenotypic spectrum in affected families that includes ALS, FTD and classic AD, with autopsy confirmed diagnoses. Thus, these data support the hypothesis that *C9ORF72* expansions contribute to a broad phenotypic spectrum of neurodegenerative diseases and will be a key factor to further decipher the underlying molecular pathology of AD.

2638T

C9ORF72 repeat expansions are a genetic cause of parkinsonism. *S. Lesage[1], I. Le Ber[1,2], C. Condroyer[1], E. Broussolle[4], A. Gabelle[5], S. Thobois[4], A. Durr[1,3], A. Brice[1,3], Parkinson's Disease Genetics Study Group.* 1) Université Pierre et Marie Curie-Paris6, Centre de Recherche de l'Institut du Cerveau et de la Moelle épinière, UMR-S975; Inserm, U975, Cnrs, UMR 7225, Paris, France; 2) Centre national de référence des démences rares, Paris, France; 3) Département de Génétique et Cytogénétique, Hôpital de la Pitié-Salpêtrière, Paris, France; 4) Université Lyon I; Centre de Neurosciences Cognitives, CNRS UMR5229; Service de Neurologie C, Hôpital Neurologique Pierre Wertheimer, Lyon, France; 5) Département de neurologie, CHU Gui de hauliac, Montpellier, France.

BACKGROUND: The recently identified C9ORF72 gene accounts for a large proportion of amyotrophic lateral sclerosis (ALS) and frontotemporal lobar degenerations (FTLD). Since several forms of these disorders are associated with parkinsonism, we hypothesized that some patients with Parkinson's disease (PD) or other forms of parkinsonism might carry pathogenic C9OFR72 expansions. AIM: To evaluate C9ORF72 repeat expansions in 1,446 parkinsonian patients. PATIENTS AND METHODS: The study included a total of 1,446 unrelated patients: 1,225 clinically diagnosed with PD, 123 with progressive supranuclear palsy (PSP), 21 with corticobasal degeneration syndrome (CBDS), 43 with Lewy body dementia (LBD) and 25 with multiple system atrophy-parkinsonism (MSA-P). Of the 1,225 PD patients, 40% reported family histories of typical PD and 4% had family histories of other neurodegenerative disorders. The GGGGCC repeat expansions in the first intron of C9ORF72 were searched for, using the repeat primed PCR. RESULTS: Five of the 1,446 patients with parkinsonism had C9ORF72 repeat expansions. In the subgroup of the 1,225 PD patients, 3 carriers of C9ORF72 repeat expansions (0.2%) had isolated parkinsonian symptoms at onset (ages at onset: 29–64 years). Two cases had long disease durations of typical PD (5 to 8 years) before cognitive deficit or dementia developed; one case had no cognitive deficit, 12 years after PD onset. Although these 3 patients were clinically diagnosed with typical PD, they all had family histories of atypical PD, degenerative dementias or ALS. We also identified C9ORF72 repeat expansions in a patient with CBDS and another with PSP. No expansion was found in patients with MSA-P or LBD. CONCLUSION: This study shows that although rare, C9ORF72 repeat expansions may be associated with clinically typical PD, but also with PSP and CBDS. In several patients, the parkinsonism was levodopa-responsive and, like PD, remained pure, without associated cognitive deficit or dementia, for more than 10 years. Interestingly, all C9ORF72 mutation carriers had positive family histories of parkinsonism, degenerative dementias or ALS. The present study, therefore, suggests that patients with PD or other parkinsonian syndromes, particularly when a family history of other neurodegenerative disorders is present, should be screened for C9ORF72 expansions. This provides the tools for identifying potential carriers with important consequences for genetic counseling.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2639F

Autosomal dominant family history is a highly sensitive but not exclusive indicator of pathogenic C9ORF72 expansions in frontotemporal dementia and amyotrophic lateral sclerosis. *J.C. Fong[1], S.J. Sha[1], L. Takada[1], K.J. Rankin[1], J.S. Yokoyama[1], N.J. Rutherford[2], A.M. Karydas[1], B.K. Khan[1], M.C. Baker[2], M. DeJesus-Hernandez[2], M. Pribadi[3], G. Coppola[3], D.H. Geschwind[3], R. Rademakers[3], A.L. Boxer[1], B.L. Miller[1].* 1) Dept of Neurology, UCSF Memory and Aging Center, San Francisco, CA; 2) Dept of Neuroscience, Mayo Clinic Florida, Jacksonville, FL; 3) Depts of Psychiatry and Neurology, UCLA, Los Angeles, CA.

Background: Diagnosis of frontotemporal dementia (FTD) is complicated by overlap with other conditions. A related but distinct disorder, amyotrophic lateral sclerosis (ALS) co-occurs in 15% of behavioral-variant FTD (bvFTD) patients. 50% of ALS patients have frontal lobe deficits. Expansion of a hexanucleotide repeat in a noncoding region of C9ORF72 was identified as the cause of chromosome 9-associated FTD and ALS. C9ORF72 analysis offers definitive molecular diagnosis in patients with FTD, ALS or both, but the challenge is how best to identify who benefits from testing. Family history is a useful tool to identify disease risk. The practical utility of family history of neurodegenerative and psychiatric disease to predict C9ORF72-associated disease remains to be clarified. **Methods:** 384 subjects with a FTD-spectrum disorder were screened by a 2-step process for pathogenic C9ORF72 expansions. PCR products with a unique electrophoresis stutter pattern provided evidence of a mutation. Pedigrees from 92 unrelated subjects underwent blinded review. Family history assessment of neurodegenerative disease followed published criteria (Goldman score or GS 1–4). Positive psychiatric family history meant ≥1 1st-/2nd-degree relative with schizophrenia, bipolar disorder, substance abuse, personality change or death by suicide. **Results:** 23 unrelated subjects carried a C9ORF72 expansion, C9+ (C9- matched controls): 12 bvFTD (47), 3 ALS (6), 8 FTD/ALS (16). An autosomal dominant (ADo) family history of neurodegenerative disease (GS 1) had 100% sensitivity (95%CI, 60–100%) to identify expansions in 7 subjects. Familial aggregation not meeting ADo criteria (GS 2) had 29% sensitivity (8–64%), 82% specificity (72–89%), and 13% positive predictive value (PPV) (4–36%) to identify expansions among remaining C9+ subjects. Diagnosis of FTD/ALS with positive psychiatric family history had 56% sensitivity (27–81%), 80% specificity (55–93%), and 63% PPV (31–86%). 43% C9+ subjects had a noncontributory/unknown family history (GS 4).**Conclusions:** An ADo family history of FTD, ALS or both is indicative of a C9ORF72 mutation. High specificity reflects high penetrance of the expansion. Poor predictive value of GS 2 reflects subjective family history reporting. Family history of psychiatric disease should be carefully assessed in FTD/ALS patients, as this may predict an expansion. Presence of expansions among patients with GS 4 supports gene testing among sporadic FTD or ALS cases.

2640W

Walking down the molecular pathogenic road of ARX polyalanine tract expansions: where do they part? *C. Shoubridge[1, 2], K.P.Y Lee[1], T. Mattiske[1], J. Gécz[1, 2].* 1) Genetics & Molec Pathology, SA Pathology, Adelaide, South Australia, Australia; 2) Dept of Paediatrics, University of Adelaide, Adelaide, South Australia, Australia.

Intellectual disability (ID) is a highly prevalent disorder that affects 1–3% of the population, with a number of causative genes mapped to the X-chromosome. The Aristaless-related homeobox gene (*ARX*) encodes a transcription factor characterised by a homeodomain and four polyalanine (PolyA) tracts. ARX is indispensable for proper forebrain and testis development and is a frequently mutated X-Linked ID gene. Expanded PolyA tracts account for over half of all mutations in *ARX* and clinically give rise to ID with or without epilepsy. Mice modelling the human PolyA expansions ($Arx^{(GCG)7}$) recapitulate many of the clinical features seen in humans, including learning impairment and a predisposition to epileptic seizures[1]. To dissect the molecular basis of different PolyA expansions *in vivo*, we have analysed mice modelling the two most frequent PolyA expansion mutations ($Arx^{(GCG)7}$ and $Arx^{432–455dup24}$) (Kindly provided by K Kitamura)[1]. Given that *Arx* is expressed in the mouse brain from 9.5 dpc, we initiated our analysis from 12.5 dpc to capture early consequence(s) of PolyA mutations. Interestingly, we did not observe protein aggregation in cells expressing Arx in contrast to our, and others, in vitro studies. We confirmed that the marked reduction in Arx protein abundance noted at 14.5 dpc[1] is present in both strains at 12.5 dpc. This reduction could not be accounted for by a loss of cell mass or expression of *Arx* transcripts in either mouse models. We are currently investigating if this reduction in protein level is due to translation efficiency or differences in protein stability. To better dissect the mechanism underpinning the phenotypic differences due to PolyA expansions, we have assessed the transcriptional consequences of mutations in these mouse models. Our data indicate a selective effect of these mutations impacting on some but not all Arx downstream targets. Since Arx is essential for interneuron development, establishing the i*n vivo* distribution of interneuron subpopulations in the two mouse models is underway. This study provides further evidence that protein dynamics can be different between *in vitro* and in vivo systems. We demonstrate that in mice, PolyA expansion mutations lead to a marked reduction of Arx protein during early brain development, and identify affected gene targets that potentially contribute to the manifestation of clinical features of the associated disease. [1](Kitamura et al. Hum Mol Genet 18:3708–3724, 2009).

2641T

Copy number variation of the neuronal glucose transporter 3 and age of onset in Huntington's disease. *A. Vittori[1,2], C. Breda[2], M. Repici[2], T.F. Outeiro[1,3,4], F. Giorgini[2], E.J. Hollox[2], the REGISTRY investigators of the EHDN.* 1) Instituto de Medicina Molecular, Cellular and Molecular Neuroscience Unit, Instituto de Fisiologia, Faculdade de Medicina da Universidade de Lisboa, Lisboa, Portugal; 2) Department of Genetics, University of Leicester, Leicester, UK; 3) Instituto de Fisiologia, Faculdade de Medicina da Universidade de Lisboa, Lisboa, Portugal; 4) Department of NeuroDegeneration and Restaurative Research, University Medizin Gottingen, Gottingen, Germany.

Huntington's disease (HD) is a neurodegenerative disease caused by an expanded triplet repeat within the *HTT* gene. The length of this repeat is inversely correlated with the age of onset (AO) of the disease, which ranges from ~2 to 80 years old. The length of this repeat explains ~50–70% of the variance of AO with the remaining variation due to environmental and other genetic factors. We investigated copy number variation (CNV) which may affect AO in HD. Decreased expression of *SLC2A3*, which encodes the neuronal glucose transporter Glut3, in the caudate of the brain is prodromal to significant neurodegeneration in HD patients and a study in a fly (*Drosophila melanogaster*) HD model has shown that overexpression of the *SLC2A-3*functional homolog is protective against neurodegeneration. We thus analyzed copy number variation of the human SLC2A3 gene as a candidate genetic modifier of the AO in HD, as the*SLC2A3* diploid copy number can vary in the population from one to three copies, possibly affecting expression levels of the gene. We genotyped *SLC2A3* CNV in 989 HD patients from the European Huntington's Disease Network (EHDN) cohort using the paralog ratio test (PRT), a comparative PCR-based method. We found that deletions and duplications of *SLC2A3* are rare (1% and 3% respectively), consistent with the frequencies found in healthy Europeans, and a weak but significant association of copy number with AO in HD (ANOVAIII test, p value = 0.030). Individuals with three copies show a delay in the AO of ~3 years compared to individuals with one or two copies. To investigate the functional basis for this effect, we have constructed lines showing over- or under-expression of the functional homolog of *SLC2A3* in a fly model of HD. We are currently analysing several disease-relevant metrics, including neurodegeneration of the photoreceptors, longevity and locomotor activity, in order to elucidate how different expression levels of the neuronal glucose transporter modulate severity of the HD phenotype in vivo.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2642F

Spinocerebellar ataxias in Brazil: frequencies and potential modifiers of phenotype. *R.M. Castilhos[1], P. Schaeffer[1], T.C. Gheno[1,2], G.V. Furtado[1,2], K.C. Donis[1,2], O. Barsottini[3], J.L. Pedroso[3], D.Z. Salarini[4], F.R. Vargas[5], M.A.F.D. Lima[5], C. Godeiro[6], L.C.S. Silva[7], M.B.P. Toralles[8], H.Y. Wanderley[9], P.F.V. Medeiros[10], S. Santos[11], H. van den Linden[12], E.M. Ribeiro[13], E.T. Pereira[14], M.L. Saraiva-Pereira[1,2], L.B. Jardim[1,2], Rede Neurogenética.* 1) Medical Genetics Service, Hospital de Clínicas de Porto Alegre, Porto Alegre, Brazil; 2) Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil; 3) Universidade Federal de São Paulo, São Paulo, Brazil; 4) Santa Casa de Misericórdia de São Paulo, São Paulo, Brazil; 5) Universidade Federal do Estado do Rio de Janeiro, Rio de Janeiro, Brazil; 6) Universidade Federal do Rio Grande do Norte, Natal, Brazil; 7) Universidade Federal do Pará, Belém, Brazil; 8) Universidade Federal da Bahia, Salvador, Brazil; 9) APAE de Vitória, Vitória, Brazil; 10) Universidade Federal de Campina Grande, Campina Grande, Brazil; 11) Universidade Estadual da Paraíba, João Pessoa, Brazil; 12) Centro de Reabilitação Dr. Henrique Santillo, Goiânia, Brazil; 13) Hospital Infantil Albert Sabin, Fortaleza, Brazil; 14) Universidade Federal de Santa Catarina, Florianópolis, Brazil.

The frequency of SCAs in diverse populations is not well known, as well as the reason for the clinical heterogeneity in each disorder. Aims: to describe the proportion of SCA1, 2, 3(MJD), 6, 7, 10, 12, 17 in families from different regions of Brazil; and to study the effect of the expanded CAG repeat (eCAGn) of SCA1,2,3,6 and 7 as well as the normal CAG repeats (nCAGn) in age at onset (AO) and neurological characteristics of SCAs. Methods: patients with an adult onset, autosomal dominant ataxia were recruited in university hospitals of 12 Brazilian cities, distant up to 3,800 km. DNA samples were sent to the central laboratory, to target analysis of the related genes. Results: 502 patients (336 families) were recruited. SCA3/MJD was the most frequent (210 or 63.3% of all families), specially in South (Rio Grande do Sul, 79%, p<0,05). The following were SCA2 (25 or 7,3% of all families; but more frequent in São Paulo, 16,4%, than in other regions, p<0,05), SCA7 (15 or 4,5% overall; but more frequent in São Paulo and Rio de Janeiro, 22,6% and 10,9%, than in other regions, p<0,05), and SCA10 (11 families or 3%). 6 SCA1 and 5 SCA6 families were also detected. 62 families or 18,6% had these diagnoses excluded; this rate was higher in the Northeast Brazil (64,3%). SCA12, 17 and DRPLA were not identified. There were no differences in the nCAGn distribution of CAG loci with respect to geographical origin, this distribution being unrelated to the different frequencies of SCAs. AO was associated to the eCAGn in SCA2 (r=-0,77), SCA3/MJD (r=-0,54), and SCA7 (r=-0,93) (p<0.0001). AO were older in SCA6 (m±sd of 55±10 years) and younger in SCA2 and SCA7 patients (29±11 and 24±12 years) than the SCAs in general (33±12, p=0,005). The nCAGn of ATXN2, 3, 7 and CACNA1A did not influence AO in SCA2, SCA3/MJD and SCA7 (ns, linear regression). In SCA3/MJD, the finding of eyelid retraction, absent reflexes, rigidity and bradykinesia were associated to large nCAGn of the CACNAIA; and ophthalmoparesis was associated to large nCAGn of ATXN2 gene (p=0.05, logistic regression). Discussion: SCA3/MJD is the most frequent SCA in Brazil. The high frequency of SCA without diagnosis in Northeast Brazil points to the presence of other, non-studied SCA in those populations. Whereas the effect of eCAGn in the AO of each specific disease was confirmed, our results suggest that the normal alleles at ATXN2 and CACNA1A genes might modify the phenotype of SCA3/MJD.

2643W

Comprehensive haplotype analysis of the Huntington's disease gene. *J. Lee[1, 2], H. Li[3], J.C. Roach[3], N. Goodman[3], L. Hood[3, 4], J.A. Kaye[5], H. Zayed[5], I.H. Kratter[5], A.C. Daub[5], S. Finkbeiner[5], R.H. Myers[6], M.E. MacDonald[1,2], J.F. Gusella[1,2].* 1) Center for Human Genetic Research, Massachusetts General Hospital, Boston, MA 02114, USA; 2) Program in Medical and Population Genetics, Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, MA 02142, USA; 3) Institute for Systems Biology, Seattle, WA 98109, USA; 4) Luxembourg Center for Systems Biomedicine, University of Luxembourg, Luxembourg, Luxembourg; 5) Gladstone Institute of Neurological Disease, San Francisco, CA 94158, USA; 6) Department of Neurology, Boston University School of Medicine, Boston, MA 02118, USA.

Huntington's disease (HD) is a dominantly inherited neurodegenerative disease in which the same type of genetic alteration is responsible for all cases of the disease, an expansion of huntingtin (HTT) CAG repeat. In HD, HTT CAG repeat lengths greater than 35 cause characteristic symptoms including involuntary movements, cognitive decline and psychiatric disturbances. Importantly, the length of the CAG repeat also determines the age at onset of neurological symptoms as ~65% of the variance in age at onset of motor symptoms is explained by the CAG repeat length. Previously, we discovered based upon common SNP-based haplotype analysis that common HTT haplotypes on HD chromosomes do not differ in their CAG repeat length distribution or for the associated age at onset of motor symptoms. Here, we expand our haplotype analysis to understand individual haplotype composition. The top 10 most frequent haplotypes on HD chromosomes (i.e., mutation-bearing chromosomes) and on normal chromosomes accounted for 89.4% and 77.7%, respectively. The union of these (i.e., 14 haplotypes) made up 90% and 82.3% of HD expanded chromosomes and normal chromosomes, respectively. Individually, HD subjects with the most common HTT expanded haplotype and the most common normal chromosome haplotype account for 10.4% of all HD subjects. In addition, approximately 90% of all HD subjects carried two different HTT haplotypes, suggesting that the vast majority of HD subjects could be eligible for HTT expanded allele-specific gene silencing approaches. Furthermore, we obtained high resolution haplotypes by whole-genome sequencing of pedigrees. Among eight HD pedigrees, four carry the most common HTT haplotype, two carry a less common haplotype, and other pedigrees carry different haplotypes. Pedigrees with same haplotype can be further stratified by rare variants. Haplotype analysis based on the sequence data confirmed the common SNP-based haplotypes and further analysis of the haplotypes at the high resolution revealed novel genetic variants that differed between HTT expanded and normal haplotypes. Together, our haplotype analysis results demonstrate the power of integrated approaches of combining population-targeted and individual-targeted analysis to understand HTT haplotype structure. The novel variants discovered may be important for delineating the history of HTT haplotypes, understanding their effects on the disease process and developing allele-specific gene silencing strategies.

2644T

Brain-specific junctophilins: Function and role in neurodegeneration. *X. Sun, M. Bhat, C. Zhang, W. Zheng, R.L. Margolis, D.D. Rudnicki.* Johns Hopkins, Baltimore, MD.

The junctophilin family of proteins, characterized by C-terminal endoplasmic/sarcoplasmic reticulum (ER/SR) transmembrane domains and N-terminal MORN motifs that specifically bind with plasma membrane (PM), participate in the formation of junctional membrane complexes linking ER/SR and PM. Muscle-specific JPH1 and JPH2 are involved in calcium-induced calcium release (CICR) in cardiac and skeletal myocytes through physiological coupling between cell surface and intracellular channels; mutations in JPH2 cause hypertrophic caridomyopathy. Our laboratory has identified a CTG repeat expansion mutation in the brain-specific junctophilin JPH3 that causes the neurodegeneration disorder Huntington's Disease-like 2 (HDL2). The pathogenic mechanism appears to involved both a loss of JPH3 protein expression and toxic properties of the expanded CUG tract in the mutant JPH3 transcript. While the physiological function of JPH3 and the other brain-specific junctophilin, JPH4, is less clear, mouse knock-out models indicate that they play a role in neuronal plasticity and motor function. Using expression analysis, knockout mouse models and proteomic techniques, our data suggests that loss of JPH3 expression increases neuronal vulnerability to multiple insults, including RNA toxicity. In addition, the subcellular distribution of the JPH3 and JPH4 protein products, and their protein interacting partners, suggest that each may have functions beyond modulation of calcium flux.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2645F

Repeat associated non-AUG translation in repeat expansion disorders. *T. Zu[1], B. Gibbens[2], L.P.W. Ranum[1].* 1) Center for NeuroGenetics, Dept. Mol. Genetics & Microbiology, Univ Florida, Gainesville, FL; 2) Univ Minnesota, Minneapolis, MN.

A variety of neurodegenerative diseases, including myotonic dystrophy type 1 (DM1) and spinocerebellar ataxia type 8 (SCA8), are caused by expansion of unstable repeats. We discovered that the canonical rules of translation do not apply for CAG•CTG expansions and that in the absence of AUG or other known alternative start codons, expanded CAG and CTG trinucleotide repeats express homopolymeric expansion proteins in all three frames. This Repeat-Associated Non-AUG (RAN) translation occurs in mammalian tissue culture and in lentiviral-vector transduced cells and tissues. We also show that RAN translation across human spinocerebellar ataxia type 8 (SCA8) and myotonic dystrophy type 1 (DM1) CAG-expansion transcripts results in the accumulation of SCA8 polyAla ($SCA8_{GCA-Ala}$) and DM1 polyGln ($DM1_{CAG-Gln}$) expansion proteins in multiple organ systems in affected patients and mice. Experiments to understand the mechanisms of RAN translation show this process is repeat length-dependent and occurs across a variety of hairpin-forming repeat tracts. We have now performed additional cell-culture studies to characterize the RNAs capable of undergoing RAN translation and the cellular factors and conditions that favor this process. We show RAN translation occurs in two or more frames across a variety of triplet and tetranucleotide repeat motifs including the DM2 CCTG•CAGG repeat expansion motif. Analysis of RAN translation in different cell types shows that robust expression of polyGln, polyAla and polySer in HEK293T, HeLa and T98 cells but not in several other cell lines, suggesting the influence of specific cellular factors on RAN translation. Additionally we show that overexpression of the MBNL1 protein, which has been shown to sequester CUG expansion RNAs in the nucleus decreases RAN translation. RAN translation is also modulated in mouse embryonic fibroblasts generated from a series of knockout mice lacking Mbnl1, Mbnl2, and Mbnl3. In summary, the efficiency of RAN translation is affected by repeat length, repeat motif and cellular factors and the accumulation of the homopolymeric proteins appears to contribute to a large category of repeat expansion disorders.

2646W

MicroRNA Expression in Alzheimer's Disease: An Exploratory Investigation. *L.M. Bekris[1, 2], F. Lutz[1, 2], T.J. Montine[3, 7], C.E. Yu[1, 2], E.R. Peskind[3, 4], J.B. Leverenz[4, 5, 6, 7].* 1) Geriatric Research, Education, and Clinical Center (GRECC), VA Puget Sound Health Care System, Seattle, WA; 2) Department of Medicine, University of Washington, Seattle, WA; 3) Pathology, University of Washington, Seattle, WA, USA; 4) Northwest Network VISN-20 Mental Illness Research, Education and Clinical Center (MIRECC), VA Puget Sound Health Care System, Seattle, WA; 5) Psychiatry and Behavioral Sciences, University of Washington School of Medicine; 6) Department of Neurology, University of Washington School of Medicine, Seattle, WA; 7) Northwest Network VISN-20 Parkinson's Disease Research, Education and Clinical Center (PADRECC).

MicroRNA (miRNA) play an essential role in gene regulation in the brain. However, little is known about their role in neurodegenerative diseases, such as late-onset Alzheimer's disease (AD). Characterizing expression profiles of miRNA in AD brain may help elucidate the role of miRNA in the pathophysiology of AD and may lead to identification of AD specific biomarkers. The aim of this exploratory investigation was to determine if miRNAs are differentially expressed according to tissue type, disease status, neuritic plaque or Braak stage. In the first phase, a post-mortem brain sample set was pooled to measure multiple miRNA (n=377) using miRNA arrays. In the second phase, individual post-mortem brain AD and control samples from this sample set, were used in qRT-PCR assays to further explore differences in miRNA levels according to neuritic plaque score or Braak stage. In the third phase, hippocampus, cerebellum, cerebrospinal fluid and plasma miRNA found positive in the first and second phase were measured in a small independent sample set. Finally, in silico analyses were performed to predict target genes for differentially expressed miRNAs. Many miRNA were not detected in our brain sample and some were not detected in both the cerebellum and hippocampus leaving 247 miRNA for further analysis. MiRNA differentially expressed in AD hippocampus compared to control hippocampus were chosen for qRT-PCR validation (n=21). Five qRT-PCR validated miRNA showed significant sensitivity and specificity for disease status, plaque score or Braak stage. Most of these 5 miRNA were detected in CSF, plasma and human neuronal and astrocyte cell lines. Plasma miRNA-15a was found to be associated with plaque score in this small independent sample. In conclusion, this investigation demonstrates that miRNA represent potential biomarkers of AD, neuritic plaque score and Braak stage. MiRNA identified in this exploratory study are predicted to target AD relevant genes and thus may warrant further investigation into their functional impact on expression and their utility as biomarkers of neurodegenerative disease.

2647T

Behavioral abnormalities of knockout mice of *Grm3*, a candidate gene for schizophrenia susceptibility. *Y. Fukumaki[1], R. Fujioka[1], T. Nii[1], A. Iwaki[1], A. Shibata[1], I. Ito[2], K. Kitaichi[3,4], M. Nomura[5], S. Hattori[6], K. Takao[6], T. Miyakawa[6].* 1) Div. of Human Molecular Genetics, Medical Institute of Bioregulation, Kyushu Univ., Fukuoka, Japan; 2) Dept. of Biology, Faculty of Science, Kyushu Univ., Fukuoka, Japan; 3) Dept. of Pharmacology, Nagasaki International Univ., Nagasaki, Japan; 4) Dept. of Pharmacy, Gifu University Hospital, Gifu, Japan; 5) Graduate School of Medical Science, Kyushu Univ., Fukuoka, Japan; 6) Div. of Systems Medical Science, Institute for Comprehensive Medical Science, Fujita Health Univ., Toyoake, Japan.

On the basis of glutamatergic dysfunction hypothesis of schizophrenia pathogenesis, we previously performed systematic association studies of glutamate receptor gene family members with schizophrenia and found an association of *GRM3* (a gene of metabotropic glutamate receptor 3: mGluR3) with the disorder. To explore the physiological role of *GRM3* in brain functions and functional roles in schizophrenia pathogenesis, we have conducted comprehensive behavioral analyses of mGluR3 knockout (KO) mice. The KO mice showed hyperactivity in the open field and light/dark transition tests, impaired reference memory for stressful events in the Porsolt forced swim test and impaired working memory in the T-Maze forced alternation task. The hyperactivity and impaired working memory are known as endophenotypes of schizophrenia. Regarding memory impairments, we examined short- and long-term synaptic plasticities by assessing the paired-pulse facilitation (PPF) and long-term potentiation (LTP), respectively, in the CA1 region of the hippocampus of the 12–14 weeks old mice. We observed PPF in KO and WT mice, but detected no differences in the amplitude of PPF between the two genotypes. In the LTP assay, the EPSP slope at 40 min after tetanus was normalized using the mean slope before stimulation. There were no statistically significant differences in the LTP between two genotypes, suggesting that mGluR3 is not essential for LTP in the CA1 region of the hippocampus of mice. As hyperactivity is typically associated with increased dopaminergic transmission, we performed *in vivo* microdialysis measurements of extracellular dopamine in the central nervous system of the KO and WT mice. We observed enhancement of methamphetamine-induced release of dopamine in the nucleus accumbens of the KO mice, indicating the altered dopaminergic transmission in the KO mice. These results suggest that interaction of dopaminergic and glutamatergic systems is involved in pathophysiology of the schizophrenia-like behavior.

2648F

Repeat Expansion in C9ORF72 in Alzheimer's Disease. *E. Majounie[1], Y. Abramzon[1], AE. Renton[1], R. Perry[2], SS. Bassett[3], O. Pletnikova[3], JC. Troncoso[3], J. Hardy[4], AB. Singleton[1], BJ. Traynor[1].* 1) National Institute on Aging, Bethesda, MD; 2) University of Alabama at Birmingham, Birmingham, AL; 3) Johns Hopkins University, Baltimore, MD; 4) University College London Institute of Neurology, London, United Kingdom.

Alzheimer's disease is the most common progressive neurodegenerative disorder, with a prevalence of 35 million worldwide. Our understanding of the genetic etiology of the disease is still largely incomplete, and only the common genetic variant (E4) in the apolipoprotein E gene is widely accepted as risk factor for sporadic Alzheimer's disease. We recently demonstrated that a hexanucleotide repeat expansion (GGGGCC) within gene C9ORF72 accounts for approximately 40% of cases of familial amyotrophic lateral sclerosis and 30% of cases of frontotemporal dementia. Given the clinical and pathological overlap between frontotemporal dementia and Alzheimer's disease, we tested the hypothesis that the C9ORF72 hexanucleotide expansion may also be associated with susceptibility to Alzheimer's disease. We screened 342 late-onset Alzheimer's disease families using a repeat-primed PCR method. The cohort consisted of 771 cases who had received a probable diagnosis of Alzheimer's disease (on the basis of criteria of the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association) and 223 siblings who were assessed as being unaffected at the time of collection. We showed that the pathogenic C9ORF72 repeat expansions was present in 3 of 342 families (<1%) apparently affected with Alzheimer's disease. Postmortem analyses of two affected sisters showed neuropathological findings consistent with a primary diagnosis of frontotemporal dementia with ubiquitin-positive, tau-negative neuronal inclusions. These findings support the possibility that the subjects carrying the C9ORF72 repeat expansion had amnestic frontotemporal dementia that was misdiagnosed as probable Alzheimer's disease. As postmortem studies estimate the diagnostic accuracy of clinical Alzheimer's disease to be 83%, the availability of a test for the C9ORF72 repeat expansion may provide an opportunity to correct the misclassification of frontotemporal dementia as Alzheimer's disease in current and future patients.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2649W

Oxr1 Is Essential for Protection against Oxidative Stress-Induced Neurodegeneration. *P.L. Oliver, M.J. Finelli, K. Liu, B. Edwards, E. Bitoun, E.B. Becker, K.E. Davies.* MRC Functional Genomics Unit, Department of Phsiology Anatomy and Genetics, University of Oxford, Oxford, UK.

Oxidative stress is a common etiological feature of neurological disorders, although the pathways that govern defence against reactive oxygen species (ROS) in neurodegeneration remain unclear. We have identified the role of oxidation resistance 1 (Oxr1) as a vital protein that controls the sensitivity of neuronal cells to oxidative stress; mice lacking Oxr1 display cerebellar neurodegeneration, and neurons are less susceptible to exogenous stress when the gene is over-expressed. A conserved short isoform of Oxr1 is also sufficient to confer this neuroprotective property both *in vitro* and *in vivo*. In addition, biochemical assays indicate that Oxr1 itself is susceptible to cysteine-mediated oxidation. Finally we show up-regulation of Oxr1 in both human amyotrophic lateral sclerosis (ALS) as well as over-expression at the pre-symtomatic stage in several mouse models of neurodegeneration. These data suggest that Oxr1 is a novel early marker for stress in neurons and potentially a new neuroprotective factor in neurodegenerative disease.

2650T

Joint SNP analysis identifies multiple signals underlying association between *CHRNB3-CHRNA6* and nicotine dependence. *B. Sadler[1,2], E. Johnson[3], N. Saccone[2], L. Bierut[2], A. Goate[2].* 1) Arizona State University, Tempe, AZ; 2) Washington University, St. Louis, MO; 3) Research Triangle Institute, Research Triangle Park, NC.

Today, tobacco use and related traits are the number one cause of preventable deaths in the United States. The loci identified by GWAS to be associated with nicotine addiction mainly include genes encoding neuronal cholinergic nicotinic receptors *(CHRNs)* in two gene clusters, one on chromosome 15 containing *CHRNA5*, *CHRNA3* and *CHRNB4* and one on chromosome 8 containing *CHRNB3* and *CHRNA6*. The strongest association with nicotine addiction is a non-synonymous change (rs16969968) in *CHRNA5*. This strong association has been replicated and validated across several studies and populations, yet does not explain a large percentage of the overall heritability of nicotine dependence, accounting for less than 5% of the variance. Variants within the gene cluster on chromosome 8 were found to affect nicotine consumption in several GWAS, but the region has not been fine-mapped nor has a causal variant been found. Using imputed data from a GWAS of nicotine dependence including 2055 European Americans (1062 ND cases and 993 controls), haplotype and conditional analyses were run for the region on chromosome 8 encompassing the *B3-A6* nicotinic receptor subunits. Of the 558 SNPs in the region, only SNPs with a MAF >5% and imputation allelic R^2 values of >0.8 were considered in the conditional analyses (179 SNPs). To determine whether there was evidence for multiple functional variants at the locus contributing to risk for nicotine dependence, the most significant SNP in the region was added to the model as a covariate in a stepwise manner for three rounds of analysis (rs7017612, $p=3.15^{-5}$; rs13263434, $p=8.28^{-4}$; rs9298628, $p=8.43^{-3}$). Examination of LD in the region confirms that the r-squared values between these SNPs are sufficiently low so as to be representative of different association signals (r^2 range 0.43–0.64). These SNPs were then used in three-locus haplotype analyses. Six haplotypes were observed although 3 had a frequency below 2%. Three broad categories of risk were associated with these haplotypes: the most common haplotype (74.5%) increased risk (OR= 1.18, p=0.029); a second haplotype with a frequency of 2.57% was neutral and 2 additional haplotypes of low frequency were strongly protective (OR=0.324, $p=9.20^{-5}$; OR=0.442, p=0.002). These data strongly suggest that the *CHRNA6/B3* locus contains multiple alleles affecting risk for nicotine dependence. We are genotyping key imputed SNPs to confirm these associations.

2651F

Variation in Gene Expression Related to Schizophrenia-Associated Neurocognitive Endophenotypes. *E.E. Quillen[1], M.A. Carless[1], J. Neary[1], M.Z. Kos[1], R.C. Gur[2], M.F. Pogue-Geile[3], K. Prasad[4], J. Blangero[1], H.H.H. Göring[1], V.L. Nimgaonkar[4], R.E. Gur[2], L. Almasy[1].* 1) Department of Genetics, Texas Biomedical Research Institute, San Antonio, TX; 2) Departments of Psychiatry, Neurology and Radiology, University of Pennsylvania Medical Center, Philadelphia, PA; 3) Department of Psychology, University of Pittsburgh, Pittsburgh, PA; 4) Department of Psychiatry and Human Genetics, University of Pittsburgh Medical Center, Pittsburgh, PA.

The Multiplex-Multigenerational Genetic Investigation (MGI) of Schizophrenia is a family-based study of schizophrenia (SCZ) and related endophenotypes including measures of disordered cognition, a hallmark of SCZ. Scores on these neurocognitive tests are heritable and genetically correlated with SCZ (ρ_G = -0.27 to -0.56), suggesting overlapping genetic influences on these endophenotypes and SCZ. MGI families were ascertained on the basis of two affected probands and all family members over age 16 were invited to participate. A Computerized Neurocognitive Battery was utilized to assess abstraction and mental flexibility, verbal memory, face memory, spatial memory, spatial processing, sensorimotor dexterity, and emotion processing. For each domain, efficiency was calculated as the ratio of accuracy to speed and shown to differ among affected individuals, unaffected relatives, and healthy controls. For genetic analysis, the endophenotypes were prioritized using the Endophenotype Ranking Value. RNA was extracted from lymphoblastoid cell lines and genome-wide transcriptional profiling was completed for 587 individuals using Illumina HT-12 Expression BeadChips. 26,358 variably expressed transcripts were normalized and measured transcript association analysis was performed in SOLAR to assess the age- and sex-adjusted relationship between expression and response efficiency for the top endophenotypes while controlling for pedigree structure (N = 320). Diminished abstraction and mental flexibility (h^2 = 0.33) is associated with decreased expression of mitochondrial gene *MCAT* ($p <$ 5e-5 for all associations). Increased expression of*RAP1GDS1*, a gene known to be expressed in the mouse brain, with high activity in the hippocampus (functionally similar to the human frontal cortex), is associated with worse performance on attention (h^2 = 0.45). Poor sensorimotor skills (h^2 = 0.60) are associated with increased *MRPS12* and decreased *FAM108C1*, *TEK*, and *TOX2* expression. Bivariate polygenic analysis indicates a shared genetic basis between sensorimotor skills and *FAM108C1* (ρ_G = 0.21) and *MRPS12* (ρ_G = -0.31) and between attention and *RAP1GDS1* (ρ_G = -0.57). In humans, *TOX2*, a transcription factor, is expressed across many regions of the brain and *MRPS12* exhibits decreased expression while *FAM108C1* and *TEK* exhibit increased expression in the frontal cortex. These results suggest a mechanism for the observed increase in frontal lobe abnormalities among individuals with SCZ.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2652W

A genome-wide epigenetic analysis in monozygotic twins discordant for schizophrenia. *F. Nishimura[1], A. Yoshikawa[1], S. Nanko[2], A. Imamura[3], Y. Okazaki[4], K. Iwamoto[5,6], T. Kato[6], T. Sasaki[7], K. Kasai[1], C. Kakiuchi[1,6].* 1) Department of Neuropsychiatry, Graduate School of Medicine, University of Tokyo, Tokyo, Japan; 2) Department of Psychiatry, Teikyo University School of Medicine, Tokyo, Japan; 3) Department of Neuropsychiatry, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan; 4) Tokyo Metropolitan Matsuzawa Hospital, Tokyo, Japan; 5) Department of Molecular Psychiatry, Graduate School of Medicine, University of Tokyo, Tokyo, Japan; 6) Laboratory for Molecular Dynamics of Mental Disorders, RIKEN Brain Science Institute, Saitama, Japan; 7) Office for Mental Health Support and Graduate School of Education, University of Tokyo, Tokyo, Japan.

Introduction: Schizophrenia is one of the major mental disorders with approximately 1% prevalence rate. The family, twin and adoption studies have established that the genetics play a significant role in the etiology of schizophrenia. Disease concordance rate of schizophrenia within monozygotic twin is about 50% and on the other hand, disease concordance rate of schizophrenia within dizygotic twin is about 15%. This suggests the development of schizophrenia has a strong genetic determinant and concurrently non-genetic factors, epigenetic changes, are also important. We previously used DNA microarray analysis to examine the mRNA expression patterns in the lymphoblastoid cells derived from two pairs of monozygotic twins discordant for schizophrenia. We found the expression changes of the ADM, SEPX1, and CD200. The up-regulation of ADM and SEPX1 were previously reported and this indicated that there are some kinds of biological differences which make the mRNA expression differences within monozygotic twin discordant for schizophrenia. In this study, we reported the DNA methylation differences that may be directly involved in etiology of schizophrenia. Method: We performed a genome-wide DNA methylation analysis on peripheral blood obtained from two twin pairs discordant for schizophrenia using Illumina Infinium HumanMethylation450 BeadChip. The objective of this study was clearly explained and written informed consent was obtained from all subjects. The study was approved by the Ethical Committee of the Faculty of Medicine, the University of Tokyo. Results: The methylation status of 485,512 individual CpGs was detected from two twin pairs. We found 59 differentially methylated CpGs from one pair and 740 CpGs from another twin pair. A number of common differentially methylated CpGs between two twin pairs is 10 sites. Conclusion: We have undertaken the genome-wide methylation analysis in monozygotic twins discordant for schizophrenia and identified new potential susceptibility genes for schizophrenia.

2653T

MicroRNA-9 and -206 expression profile in human amyotrophic lateral sclerosis. *D. Dogini[1], A. Nucci[2], I. Lopes-Cendes[1], M. França Jr[2].* 1) Dept of Medical Genetics, FCM - UNICAMP, Campinas, Sao Paulo, Brazil; 2) Dept of Neurology, FCM - UNICAMP, Campinas, Sao Paulo, Brazil.

MiRNAs are a new class of small RNA molecules (21–24 nucleotide-long) that negatively regulate gene expression either by translational repression or target mRNA degradation. MiRNAs are involved in many important biological processes including cell differentiation, central nervous system development and neurodegenerative diseases. Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disease that specifically affects upper and lower motor neurons, leading to progressive paralysis and death. Recent data indicate that miR-9 and miR-206 are abnormally expressed in animal models of ALS. The main goal of this study is to investigate the expression pattern of the neuronal microRNA (miRNA) miR-9 and the muscle-specific miRNA, miR-206 in patients with ALS. Total RNA was isolated from skeletal muscle tissue of 7 patients with ALS and 5 healthy controls for comparison. Each patient fulfilled El Escorial diagnostic criteria for probable or definite ALS. RNA samples were used in real-time PCR reactions with TaqMan™ microRNA assays (Applied Biosystems) to quantify miR-9 and miR-206. MiR-9 and miR-206 were both over expressed in patients when compared to controls (5-fold and 3-fold, respectively). We have identified abnormal expression of miR-9 and miR-206 in human ALS, which may indicate a tight down regulation of genes regulated by these miRNAs in ALS. Our results should be explored in additional studies to find and validate possible gene targets that improve understanding of miRNA regulation in ALS and investigate miRNAs that can be used as biomarkers for early diagnosis of ALS disease.

2654F

Altered cell growth and mTOR signaling in idiopathic autism spectrum disorders. *A.M. Suzuki[1], C.O.F. Machado[2], K.G. Oliveira[1], E. Vadasz[3], M.R. Passos-Bueno[1], A.L. Sertie[2].* 1) Departamento de Genética e Biologia Evolutiva, Universidade de São Paulo, São Paulo, Brazil; 2) Hospital Israelita Albert Einstein, Instituto de Ensino e Pesquisa, São Paulo, Brazil; 3) Instituto de Psiquiatria, Faculdade de Medicina, Universidade de São Paulo, São Paulo, Brazil.

Autism Spectrum Disorders (ASD) are common neurodevelopmental conditions with high phenotypic and genetic heterogeneity. Although the genetic cause is mostly unknown, evidences support the hypothesis that common pathogenic pathways may be shared by ASD. The molecular mechanisms underlying the neurological defects in genetically defined syndromes with increased prevalence of ASD, such as Fragile-X and Tuberous sclerosis syndromes, may provide important clues for such common pathways. Currently, it is well known that these monogenic disorders present dysregulated mTOR signaling, which controls several biological processes, such as cell growth and protein synthesis. Therefore, the present study was undertaken to investigate whether altered mTOR signaling is also a feature of subjects with idiopathic ASD. Firstly, we compared the proliferative response of dental pulp stem cells (DPSCs) derived from 6 autistic male patients and 5 matched controls to different fetal bovine serum (FBS) concentrations. Interestingly, all patient-derived DPSCs showed significant increased or decreased mitogenic response to FBS compared to controls. Next, to investigate whether these responses were due to altered mTOR signaling, we evaluated by western blot the pattern of phosphorylation of pAktS473, pmTORS2448, and pS6S235/236 proteins in an expanded sample of 13 patients and 8 controls DPSCs after 48hr of starvation period followed or not by stimulation with 20% FBS for 30min. Under these conditions, 2 patient-derived cells, in which proliferative response to FBS was higher than in the controls, showed increased phosphorylation of components of the mTOR cascade compared to controls: in relation to FBS-starved condition, 1 patient showed 23.5-fold increase in pAkt, 9.8-fold increase in pmTOR, and 10.4-fold increase in pS6, and the other showed 6.9-fold increase in pS6; whereas controls had lower activation values: mean 3.8-fold (SD 2.7) for pAkt, mean 1.8-fold (SD 1.3) for pmTOR, and mean 3.1-fold (SD 1.1) for pS6. Our data suggest that dysregulation of mTOR pathway may play an important role in the pathogenesis of at least a subgroup of idiopathic ASD cases. In addition, our findings suggest that other pathways may also be involved in the altered cell proliferation phenotype observed in patient-derived cells. Collectively, these data support an important contributory role for pathways involved in cell growth and protein synthesis in the pathophysiology of ASD. FAPESP, CNPq.

2655W

Investigating the genetic susceptibility to epilepsy in animal models using gene expression studies. *A.H.B. Matos[1], V.D.B Pascoal[1], D.R. Nascimento[2], M. Martins[2], C.S. Rocha[1], J.F. Vasconcellos[1], M.T. Chamma[3], C.V. Maurer-Morelli[1], A.S. Martins[2], A.C. Valle[3], A.L.B. Godard[2], I. Lopes-Cendes[1].* 1) UNICAMP, Campinas, Brazil; 2) UFMG, Belo Horizonte, Brazil; 3) USP, Sao Paulo, Brazil.

BACKGROUND: Wistar audiogenic rat (WAR) is a genetic epilepsy model susceptible to audiogenic seizures, after high-intensity sound stimulation. Another genetic model we have recently identified is the generalized epilepsy with absence seizures (GEAS) rat. The aim of the present study was to characterize and compare the genetic profile of these two strains using gene expression analysis. METHODS: We obtained total RNA from five susceptible WAR (hippocampus and corpora quadrigemina), and two resistant WAR, as well as from hippocampus of three GEAS rats and three control Wistar. Gene expression analysis was performed using the GeneChip® Rat Genome 230 2.0 Array (Affymetrix™), and analyzed in R environment using the Affy and RankProd packages from Bioconductor. Overrepresented gene ontology categories were identified with DAVID software, and gene interactions and correlation networks were identified with Ingenuity Pathways Analysis software. RESULTS: Enriched gene ontology identified in WAR was involved in neuronal development and differentiation, regulation of synaptic transmission and neuron projection, as well as cell-cell signaling. The most active signaling pathways were development of the cochlear duct (Neurod1 up-regulated), and calcium and GABA receptor signaling (Gabra6 up-regulated). The main genes with differential expression and a possible biological role in epileptogenesis were, Apbb1 and Scn1A. In the GEAS rats we found differential expression of genes related to central nervous system development, activation of MAPK transcription factors, neuronal migration and apoptosis, such as Nrsn1, Hspb1, Fos, Twist1 and Krt18. The top enriched gene ontology categories included signal peptide, extracellular region and antigen processing and presentation. Among the most activated signaling pathways was neurosystem development. CONCLUSION: Our results clearly show a different molecular signature in the two genetic rat models analyzed. This study may help to clarify the underlying molecular mechanism that leads to the predisposition to seizures in these animals. SUPPORTED: FAPESP-FAPEMIG.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2656T

Effects of oral L-carnitine administration in narcolepsy patients. *T. Miyagawa[1], H. Kawamura[2], M. Obuchi[2], A. Ikesaki[2], A. Ozaki[3], K. Tokunaga[1], Y. Inoue[4, 5], M. Honda[4, 6].* 1) Department of Human Genetics, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan; 2) Clinical Trial Operations Division, Site Support Institute Co., Ltd. (SSI), Tokyo, Japan; 3) School of Nursing, Faculty of Nursing, Toho University, Tokyo, Japan; 4) Japan Somnology Center, Neuropsychiatric Research Institute, Tokyo, Japan; 5) Department of Somnology, Tokyo Medical University, Tokyo, Japan; 6) Sleep Disorders Project, Department of Psychiatry and Behavioral Sciences, Tokyo Metropolitan Institute of Medical Science, Tokyo, Japan.

Narcolepsy is a sleep disorder characterized by excessive daytime sleepiness, cataplexy, and rapid eye movement (REM) sleep abnormalities. A genome-wide association study identified a novel narcolepsy-related SNP rs5770917, which is located adjacent to the carnitine palmitoyltransferase 1B (*CPT1B*) gene encoding an enzyme involved in β-oxidation of long-chain fatty acids. The mRNA expression levels of *CPT1B* were associated with SNP rs5770917. In addition, we recently reported that acylcarnitine levels were abnormally low in narcolepsy patients. To assess the efficacy of oral L-carnitine for the treatment of narcolepsy, we performed a clinical trial administering L-carnitine (510 mg/day) to patients with the disease. The study design was a randomized, double-blind, cross-over and placebo-controlled trial. Thirty narcolepsy patients were enrolled in our study. Two patients were withdrawn and 28 patients were included in the statistical analysis (15 males and 13 females, all with *HLA-DQB1*06:02*). L-carnitine treatment significantly improved the total time for dozing off during the daytime, calculated from the sleep logs, compared with that of the placebo period (P=0.048). L-carnitine efficiently increased serum acylcarnitine levels, and reduced serum triglycerides concentration. Differences in the Japanese version of the Epworth Sleepiness Scale (ESS) and the Medical Outcomes Study 36-Item Short-Form Health Survey (SF-36) vitality and mental health subscales did not reach statistical significance between L-carnitine and placebo. However, the JESS scores and SF-36 subscale scores showed a tendency to improve in the L-carnitine period from the baseline. Although we examined the possible association of SNP rs5770917 with the clinical efficacy data, no significant associations were observed for any of the parameters. This study suggests that oral L-carnitine can be effective in reducing excessive daytime sleepiness in narcolepsy patients.

2657F

Discovery of genetic syndromes in patients manifesting with primary neuropsychiatric disorders. *J. So, J.L. Kennedy.* Neurogenetics Section, Neuroscience Research, Centre for Addiction and Mental Health, Toronto, ON, Canada.

Background: Many genetic syndromes are known to present with psychiatric features. Significantly, many syndromes require surveillance for other, sometimes preventable or treatable, clinical manifestations. Additionally, many metabolic conditions are treatable by dietary adjustments, medication and/or enzyme replacement therapy. However, many patients being treated for psychiatric symptoms are not routinely referred for genetic assessment, despite the presence of comorbidities that may suggest a genetic diagnosis. The goal of this study is to determine the role of hidden genetic conditions in patients with psychiatric features and comorbidities. By defining the prevalence of genetic syndromes within this population, we aim to increase awareness of potential genetic diagnoses in psychiatric patients, thereby potentially revolutionizing the standard of care, management and treatment of selected individuals with psychiatric impairment.**Methods**: Patients recruited are age 16 years and older with psychiatric features and at least one of 1) neurologic abnormality, 2) developmental delay (DD), autism spectrum disorder (ASD) or pervasive developmental disorder (PDD), 3) dysmorphic features, 4) congenital anomalies or 5) family history of DD, ASD or PDD. A database of phenotypic correlates will aid in determining the highest-yield data that lead to the most effective and efficient diagnosis of genetic syndromes in patients with neuropsychiatric disorders. Additionally, novel adult phenotypes in genes associated with known genetic syndromes can be delineated. If a diagnosis cannot be achieved by clinical means, samples from recruited patients are investigated for the presence of alterations and CNVs in high-priority candidate genes and for novel variants by high-resolution SNP array. **Results**: We present examples of genetic syndromes and variants identified in the patient cohort to date, including an unusual presentation of Fragile X syndrome in a female. Patient recruitment and sample analysis are ongoing. **Discussion**: This study will contribute to identifying genetic syndromes that can mimic or contribute to psychiatric illness, as well as determining the prevalence of these syndromes in psychiatric populations. It will facilitate education of primary care physicians, patients and families about, as well as establishment of standardized protocols for, the surveillance, management and treatment of genetic diagnoses in psychiatric patients with comorbidities.

2658W

Genomic and bioinformatic analyses of sex differences in autism spectrum disorder. *D.M. Werling[1], J.K. Lowe[2,3,4], V. Chandran[2], R.M. Cantor[4,5], D.H. Geschwind[2,3,4], ACE Genetics Consortium.* 1) Interdepartmental PhD Program in Neuroscience, Brain Research Institute, University of California, Los Angeles, Los Angeles, CA; 2) Program in Neurogenetics, Department of Neurology, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA; 3) Center for Autism Research and Treatment, Semel Institute, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA; 4) Center for Neurobehavioral Genetics, Semel Institute, University of California, Los Angeles, Los Angeles, CA; 5) Department of Human Genetics, University of California, Los Angeles, Los Angeles, CA.

Males are four times more likely to be affected by autism spectrum disorder (ASD) than females, an unexplained feature of ASDs likely to provide important pathophysiological clues. To investigate this sex difference, two methods, an unbiased genetic approach and a supervised bio-informatic analysis, were employed to identify new sex-specific genetic loci for ASD and examine known ASD-associated genes for evidence of sex-specific regulation. We first stratified multiplex families from the Autism Genetic Resource Exchange (AGRE) into two groups: those with affected male children only (male only, MO), and those with at least one affected female child (female containing, FC). Within each group, genome-wide SNP data were tested for association with ASD and 57,929 SNPs pruned for $r^2 \leq 0.1$ were used for nonparametric linkage analyses. Peak LOD scores were used to define genomic regions of interest, and haplotypes and imputed SNPs were tested for association with ASD within these regions. No genome-wide significant association was observed. However, significant linkage was observed in the MO group at chr1p31.3 (LOD 3.04; rs7521242). The strongest single-marker association signals within a 1-LOD interval around this peak (4.9Mb, 16 genes) were in *INADL*, a gene related to tight junctions, which is implicated in retinal degeneration, and interacts with the tuberous sclerosis complex. The strongest linkage signal in the FC group was at chr8p21.2 (LOD 2.51; rs10111167); association signals within a 1-LOD interval around this peak (3.3 Mb, 23 genes) implicate *EBF2*, a transcription factor involved in cerebellum development. For the second strategy, we used the JASPAR database of experimentally-validated transcription factor binding sites to assess transcription factor binding site enrichment in two sets of ASD-associated genes drawn from the literature. Compared with selected background sequences, the upstream regions (1 kb) of ASD candidate genes were enriched for binding site sequences for SRY, a chrY transcription factor best known for its critical role in sex determination of the gonads. This finding suggests that SRY-driven sex-specific transcriptional regulation of ASD-associated genes may play a role in males' greater vulnerability to ASDs. These complementary strategies, linkage and analysis of sex-specific gene regulation, show significant promise in identifying novel genetic loci and potential mechanisms that may contribute to the unexplained sex bias in ASD.

2659T

Analysis of RBFOX1 gene expression in lymphoblastoid cell lines of discordant ASD sibpairs. *C. Zusi, P. Prandini, G. Malerba, A. Marostica, E. Trabetti, P.F. Pigatti, The Italian Autism Network (ITAN).* Department of Life and Reproduction Sciences, University of Verona, Verona, Verona, Italy.

Several line of evidence suggest that 16p region is associated with ASDs (Autism Spectrum Disorders) and gene expression network studies of autistic brain vs controls show underexpression of genes related to neuronal marker RBFOX1 (Voineagu I et al, Nature 2011), a key specific alternative splicing regulator in neural cells. Here, we report on RBFOX1 gene expression in discordant ASD sibpairs using specimens which are more accessible than brain samples, i.e. lymphocytes from peripheral blood and lymphoblastoid cell lines (LCLs). Ten male sibpairs aged between eight and fifteen years (mean=10,8 years, SD 1.99) were selected from ITAN cohort. After RNA isolation from each sample, expression of RBFOX1 and housekeeping genes (DECR, FPGS, and TRAP1) was measured by semiquantitative RT-PCR using TaqMan assays. RBFOX1 gene expression was detected in LCLs but not in lymphocytes. No significant differences between ASD and non affected brothers were found in LCLs. Although LCLs are generally considered a valid approximation of in-brain expression, we were not able to confirm in ASD LCLs the previously reported gene expression dowregulation of RBFOX1. This maybe due to the fact that differences in RBFOX1 gene expression in brain samples are more pronounced than in LCLs.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2660F

Protein-protein interaction network for autism spectrum disorders. *R. Corominas[1], X. Yang[2,3], G.N. Lin[1], S. Kang[1], Y. Shen[2,3], L. Ghamsari[2,3], S.A. Wanamaker[2,3], S. Tam[2,3], M. Rodriguez[2,3], M. Broly[2,3], V. Romero[2,3], I. Lemmens[4], M. Tasan[5], J. Tavernier[4], F.P. Roth[5], M. Calderwood[2,3], J. Sebat[1], K. Salehi-Ashtiani[6], D.E. Hill[2,3], M. Vidal[2,3], T. Hao[2,3], L.M. Iakoucheva[1].* 1) Department of Psychiatry, University of California San Diego, La Jolla, CA., United States; 2) Center for Cancer Systems Biology (CCSB) and Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, MA., United States; 3) Department of Genetics, Harvard Medical School, Boston, MA., United States; 4) Department of Medical Protein Research, VIB, Ghent University, Ghent, Belgium; 5) Donnelly Centre for Cellular and Biomolecular Research, University of Toronto and Samuel Lunenfeld Research Institute, Mt. Sinai Hospital, Toronto, Ontario, Canada; 6) New York University Abu Dhabi, United Arab Emirates.

Autism spectrum disorder (ASD) is a psychiatric disorder with a strong genetic component. A large number of functionally diverse genes have been shown to play a role in ASD. We investigated ASD from a systems biology perspective by experimentally constructing a protein-protein interaction (PPI) network that involves 191 candidate genes. Using stringent yeast-two-hybrid system we screened 420 brain-expressed splice isoforms of these genes for interactions against human ORFeome (~15,000 clones) and built Autism-centered Interactome Network (ACIN). In addition, by screening splice isoforms for interactions against themselves, we built a unique autism "splice" interactome. High quality of ACIN was confirmed in an independent assay in mammalian system (MAPPIT) and further verified by integrating publically available annotations for interacting protein pairs. Significant enrichment of co-expressed, co-regulated and co-GO annotated pairs was observed in ACIN as compared to control. Moreover, ACIN preys (binding partners of ASD candidates) demonstrated increased interconnectivity and significant enrichment in CNV membership, suggesting their potential relevance to disease. Subsequent grouping of ACIN nodes based on CNV membership produced highly connected autism CNV-CNV network. For the first time, this network revealed direct functional connections between different autism-relevant CNVs. Additional analysis of recent exome sequencing data mapped to ACIN showed that ACIN preys were not enriched in genes hit by damaging *de novo* mutations. Integration of brain expression data (from *Voineagu et al, Nature, 2011*) with ACIN PPIs revealed that ACIN is significantly enriched in genes differentially expressed in ASD cases *vs.* controls (P=4.5×10-5). Analysis of isoform interaction network showed strong influence of alternative splicing on PPIs, introducing an additional layer of network complexity. Testing splice isoforms for interactions proved useful, as exemplified by *BZRAP1* isoforms that connects autism genes from functionally diverse pathways into a single component. In summary, our results corroborate autism relevance of previously known genes (eg.*FOXP2*, *STX1A*, *BZRAP1*) and identify new players in autism pathways that were previously unrelated or weakly related to ASD (eg.*PIAS2*, *NAPB*, *TCF4*). Moreover, the unique "splice" interactome provides a more complete view of connectivity among ASD genes, CNVs and pathways, representing a valuable resource for future studies.

2661W

Transcriptome analysis of brain-related diseases in whole blood by RNA sequencing. *S. de Jong[1], T. Wang[1], A.P.S. Ori[1], M. Wiedau-Pazos[2], E. Strengman[3], N. Tran[4], J. DeYoung[1], R.S. Kahn[5], M.P.M. Boks[5], S. Horvath[6,7], R.A. Ophoff[1,5,7].* 1) Center for Neurobehavioral Genetics, University of California Los Angeles, Los Angeles, CA; 2) Department of Neurology, David Geffen School of Medicine at UCLA, Los Angeles, California, USA; 3) Department of Medical Genetics, University Medical Center Utrecht, Utrecht, the Netherlands; 4) Informatics Center for Neurogenetics and Neurogenomics, University of California Los Angeles, Los Angeles, California, USA; 5) Department of Psychiatry, Rudolf Magnus Institute of Neuroscience, University Medical Center Utrecht, Utrecht, the Netherlands; 6) Department of Biostatistics, School of Public Health, University of California Los Angeles, Los Angeles, California, USA; 7) Department of Human Genetics, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, California, USA.

Neuropsychiatric and neurological disorders are prevalent in the population. They can severely disrupt daily life of patients and their families and bear a huge societal and economic burden. Genetic studies have identified a number of susceptibility loci but these do not nearly explain the high heritability estimates for these disorders, most likely due to extensive clinical and genetic heterogeneity. Efforts have been made to develop biomarkers using array-based gene expression profiles from peripheral blood of patients but results remain equivocal. In order to comprehensively study trait-specific gene expression profiles in blood we performed a RNA sequencing (RNA-seq) experiment with 192 subjects representing three different brain-related diseases. Transcriptome analysis by RNAseq allows for in-depth study of gene expression while the inclusion of patients with different disease types will provide insights into the specificity of these expression profiles distinguishing different brain-related disorders. We included four groups of subjects (each n=48): patients with diagnosis of (i) schizophrenia, (ii) severe bipolar disorder, (iii) amyotrophic lateral sclerosis, and (iv) healthy controls. RNA was collected from whole blood using PAXgene tubes and sequencing was performed using the Illumina TruSeq sample kits and HiSeq2000 technology. In addition to standard case-control comparisons, we will also present comparisons of gene expression signatures between these disorders. We use systems biology methods to reconstruct weighted co-expression networks and perform expression QTL analysis. Our study design with RNA-seq data of 192 subjects provides an unique opportunity for detailed analysis of gene-expression differences including allele-specific expression and alternative splicing in the context of psychiatric and neurologic diseases.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2662W

The role of GTF2IRD1 in the auditory pathology of Williams-Beuren syndrome. *CP. Canales[1], ACY. Wong[2], GD. Housley[2], PW. Gunning[1], EC. Hardeman[1], SJ. Palmer[1].* 1) Cellular and Genetic Medicine Unit, School of Medical Sciences, University of New South Wales, Sydney, Australia; 2) Translational Neuroscience Facility, School of Medical Sciences, University of New South Wales, Sydney, Australia.

Williams-Beuren Syndrome (WBS) is an autosomal dominant condition associated with multisystemic developmental abnormalities caused by a hemizygous deletion of a 1.55 Mb WBS critical region (WBSCR) on chromosome 7q11.23, involving up to 28 genes. The auditory impact has been widely studied because of reports of higher musical ability and phonophobia in WBS patients. Children are typically reported as having hyperacusis (high sensitivity to sound) or auditory allodynia (aversion to certain sounds not normally found aversive). However, high-frequency sensorineural hearing loss (SNHL) is common in individuals with WBS at all ages. In addition, patients who have “normal” hearing, as defined by behavioural thresholds, can show sub-clinical impairments or undetected cochlear pathology. Most reports suggest that although hearing capacity is reduced in WBS patients, the perception of sound is somehow amplified. Atypical smaller deletions of the WBSCR have implicated the genes *GTF2I* and *GTF2IRD1* as the prime candidates for the cause of specific cognitive and behavioural defects that result in the distinctive WBS personality profile. However, the SNHL has not yet been genetically mapped. In this study, we aimed to investigate the role of GTF2IRD1 in the causation of hearing loss using a targeted *Gtf2ird1* null mouse that blocks normal GTF2IRD1 protein production. GTF2IRD1 protein is a transcription factor that is broadly expressed during development and in adult neuronal tissues, but its cellular role is not completely understood. However, Gtf2ird1 null mice have shown some intriguing parallels with WBS, including craniofacial abnormalities and behavioural changes. We have examined hearing capacity by auditory brainstem response (ABR), which measures sound-evoked auditory neurotransmission from the cochlear nerve to the auditory midbrain, and by distortion product otoacoustic emissions (DPOAE) to analyse the amplifying function of the outer hair cells. Our results indicate that *Gtf2ird1* null mice have impaired cochlear sound transduction, as manifested by higher auditory thresholds in the DPOAE analysis and compromised auditory neurotransmission, as determined by the ABR analysis. Analysis of *Gtf2ird1* expression in the organ of corti has revealed transcription in the outer hair cells and very strong expression in the cochlear spiral ganglion cell bodies. These data indicate that GTF2IRD1 plays an important role in the WBS hearing phenotype.

2663T

Tricellulin is Required to Maintain the Barrier Function of the Reticular Lamina and Protect the Cochlear Hair Cells from Potassium Toxicity. *G. Nayak[1], I.A. Belyantseva[2], R. Yousaf[1], S. Edelmann[3], S. Lee[2], G. Sinha[3], M. Rafeeq[4], S.M. Jones[5], C.M. Van Itallie[6], J.M. Anderson[6], A. Forge[7], G.I. Frolenkov[3], S. Riazuddin[1, 8].* 1) Laboratory of Molecular Genetics, Division of Pediatric Otolaryngology Head & Neck Surgery, Cincinnati Children's Hospital Research Foundation, and the Department of Otolaryngology, College of Medicine, University of Cincinnati, Cincinnati, Ohio, Un; 2) Section on Human Genetics, Laboratory of Molecular Genetics, National Institute on Deafness and Other Communication Disorders, National Institutes of Health, Rockville, Maryland, United States of America; 3) Department of Physiology, University of Kentucky, Lexington, Kentucky, United States of America; 4) National Center of Excellence in Molecular Biology, University of the Punjab, Lahore, Pakistan; 5) Department of Special Education and Communication Disorders, University of Nebraska, Omaha, Nebraska, United States of America; 6) National, Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Maryland, United States of America; 7) Centre for Auditory Research, University College London, London, United Kingdom; 8) Division of Pediatric Ophthalmology, Cincinnati Children's Hospital Research Foundation, and the Department of Ophthalmology, College of Medicine, University of Cincinnati, Cincinnati, Ohio, United States of America.

Separation of compositionally distinct extracellular fluids is indispensable for epithelia that line the lumen of different organs and is mediated by various tight junctional complexes. In the inner ear, tight junctions between the epithelial cells lining the scala media are responsible for isolating the endolymph (high K^+, low Na^+) from the perilymph (low K^+, high Na^+) that fills the other cochlear compartments. Mutations in *TRIC*, that encodes a tricellular tight junction protein, lead to recessively inherited hearing loss in humans (*DFNB49*). To determine the mechanism of deafness due to loss of tricellulin, we developed and characterized the *Tric* knock-in mouse model. Mice homozygous for the knock-in mutation displayed early onset rapidly progressing hearing loss and were completely deaf by 4 weeks of age. Progression of hearing loss coincided with outer hair cell degeneration, followed by loss of inner hair cells. Although the stria vascularis and vestibular functions were intact, the tricellular tight junctions in the marginal cell layer and vestibular epithelia were structurally abnormal in the homozygous mutant mice. Potassium current measurements indicated that the outer hair cell maturation was normal in the homozygous mutants. Therefore, the organ of Corti degeneration in the *Tric* knock-in mouse suggests compromised barrier function of the reticular lamina, which leads to outer hair cell death due to K^+ leakage from the endolymph. Consistent with this hypothesis, the hair cell loss in tricellulin mutant mice was rescued *in vivo* when the endocochlear potential was eliminated by deleting *Pou3F4*. Finally, comprehensive phenotypic screening showed that the homozygous mutants have a broader phenotype besides hearing loss. Thus, the loss of tricellulin leads to syndromic deafness in mice.

2664F

Generation and characterization of *Plp1* dup, a new mouse model of Pelizaeus-Merzbacher Disease with tandem genomic duplication. *G. Hobson[1, 2, 3], K. Clark[1, 2], L. Sakowski[1, 2], K. Sperle[1], L. Banser[1], D. Bessert[4], R. Skoff[4].* 1) Nemours Biomedical Res, A I duPont Hosp Children, Wilmington, DE; 2) Dept Biology, Univ Delaware, Newark, DE; 3) Jefferson Med College, Thomas Jefferson University, Philadelphia, PA; 4) Dept Anatomy and Cell Biology, Wayne State University, Detroit, MI.

Pelizaeus-Merzbacher disease (PMD; MIM #312080) is an X-linked leukodystrophy that most commonly leads to life-long disability. The disorder arises from various lesions in the proteolipid protein 1 gene (*PLP1*; MIM #300401), which encodes proteolipid protein (PLP1), the most abundant protein in myelin. The most common mutation (50–75%) is a tandem, direct duplication of the genomic region encompassing *PLP1* and other flanking genes. Transgenic models carrying 4–7 extra copies of the *Plp1* gene on an autosome have been useful for understanding the consequences of overexpression of *Plp1* on myelination, but they do not recapitulate the genetics of the disease. We have developed a murine strain, *Plp1*dup, with a duplication analogous to human alleles. *Plp1*dup mice carry a stable tandem duplication of the genomic region on the X chromosome including the *Plp1* gene and 5 flanking genes. These mice have progressive CNS myelin degeneration and associated neurological abnormalities (motor skills, gait). The *Plp1*dup mouse will be a useful model for understanding the pathogenesis of PMD and for testing potential therapeutic strategies.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2665W

Gtf2i copy number affects the growth of neural progenitors in the developing mouse cortex. *H. Oh[1], L.R. Osborne[1,2,3].* 1) Institute of Medical Sciences, University of Toronto, Toronto, Ontario, Canada; 2) Department of Molecular Genetics, University of Toronto, Ontario, Canada; 3) Department of Medicine, University of Toronto, Toronto, Ontario, Canada.

Williams-Beuren syndrome (WBS) is a neurodevelopmental disorder caused by the deletion of 26 genes on chromosome 7q11.23. People with WBS exhibit an array of cognitive and behavioral features including intellectual disability, social disinhibition, anxiety, deficits in visuospatial construction and attention deficit hyperactivity disorder, but the neurobiological basis for these symptoms remains unknown. Our lab has generated mouse models with altered gene dosage of two candidate genes from the WBS deletion region, General Transcription Factor 2 I (Gtf2i) and GTF2I Repeat Domain containing protein 1 (Gtf2ird1). These models include mice with hemizygous deletion of Gtf2i (Gtf2i+/−) or of both genes (Gtf2i/Gtf2ird1+/−) and mice with duplication of Gtf2i (Gtf2i+/Dup). We have used these mice to determine whether altered expression of Gtf2i and/or Gtf2ird1 affects neural stem cell growth and neurogenesis, and potentially contributes to the neurological features of WBS. Neuronal precursor cells (NPCs) from the brain cortex of embryonic day 12.5 mice were dissected, dissociated, seeded at equal densities and cultured for between 1 and 3 days in vitro (d.i.v.). We then used cell-specific markers to analyze cell proliferation, apoptosis and differentiation. We initially used pooled litters which we validated using single embryo culture, where we were able to directly compare individual littermates of different genotypes. Our results showed that mice with a hemizygous deletion of Gtf2i and Gtf2ird1 had an overall reduction in the number of NPCs after 1 d.i.v compared to wildtype mice, whereas mice with an additional copy of Gtf2i had an increase in the number of precursors. Similar results were seen for differentiated neurons after 3 d.i.v. No change in either proliferation or apotosis was associated with any genotype. These results indicate that gene copy number of Gtf2i directly affects the growth of the neural stem cell population in the mid-gestation developing mouse cortex. We are now looking to see whether this affect continues into later developmental stages or translates into altered neuronal density in the post-natal or adult cortex. The growth and differentiation of the NPC pool into neurons and the specification of specific neuronal subtype identities are crucial in the proper development of the cortex, and we hypothesize that some of the neurological features of WBS may stem from impairments in these early stages of neuronal development.

2666T

***GPM6A* is duplicated in a patient with learning disability and behavioral anomalies and influences stress response and long-term memory in *Drosophila melanogaster*.** *C. Zweier[1], A. Gregor[1], J. Kramer[2], I. Schanze[1,3], A.B. Ekici[1], A. Schenck[2], A. Reis[1].* 1) Institute of Human Genetics, University Erlangen-Nuremberg, Erlangen, Germany; 2) Department of Human Genetics, Nijmegen Centre for Molecular Life Sciences, Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands; 3) Institute of Human Genetics, University Hospital Magdeburg, Magdeburg, Germany.

In a 10 year old patient with learning disability and a tested IQ of 75 we detected a de novo copy number gain of 655 kb containing only the *GPM6A* gene with an Affymetrix 6.0 SNP array. This aberration was not observed in 820 healthy controls. The girl additionally presented with behavioral anomalies such as provocative actions, repetitive questions, and problems in interaction with others. Apart from mild obesity and short metacarpalia IV and V she showed no specific dysmorphism or organic anomalies. Expression analysis in blood lymphocytes showed increased *GPM6A* expression in the patient compared to healthy controls. Glycoprotein M6a (GPM6A) is a transmembrane protein of the PLP/DM20 protein family whose expression is restricted to neurons. In rats GPM6A has been implicated in neurite outgrowth, neuronal differentiation and synapse formation. M6a has also been implicated in stress response in different animal models. So far, no association of *GPM6A* defects with cognitive disorders has been reported. Mutational screening of the gene in 284 patients with variable intellectual disability did not reveal any further alteration. We therefore employed *Drosophila melanogaster* as a model organism to further elucidate the role of GPM6A in stress response, synapse formation, and behavior and learning. Using the UAS-Gal4 system we analyzed the function of the Drosophila homologue m6 after tissue specific knockdown and overexpression with different m6 RNAi and m6 overexpressing lines. We found elevated expression levels of m6 in fly brains of L3 larvae and adult flies after heat shock stress treatment, which further increased with heat shock elongation. These findings indicate that m6 expression is also stress responsive in flies. While analyses of neuromuscular junctions in L3 larvae upon knockdown or overexpression of m6 did not reveal any significant alterations in synapse formation, behavioral testing using the courtship conditioning paradigm revealed impairment of long-term memory in both knockdown and overexpression conditions. Our findings in Drosophila indicate a role of m6 in neuronal function, especially in stress response and long-term memory, thus supporting a relevance of the*GPM6A* duplication in our patient for her cognitive and behavioral phenotype.

2667F

FASPS Associated Per3 Mutation Affects Mesenchymal Stem Cell Differentiation. *K. Kaasik[1], C.R Jones[1], Y.H Xu[1], S. Nishino[2], L.J Ptacek[1], Y.H Fu[1].* 1) Univeristy of California San Francisco, San Francisco, CA 1550 4th street, 2922, San Francisco, CA94143; 2) Stanford University.

Circadian clock and sleep homeostasis are tightly regulated. Here we identified double mutation in the circadian co-transcriptional factor Period 3 gene from the human subsets with Familiar Advanced Sleep Phase Syndrome (FASPS) and severe depression phenotype. The Per3 BAC transgenic mouse model recapitulates aspects of human FASPS phenotype. We found increased wakefulness and reduced non-rapid eye movement (REM) sleep during the light phase and quicker recovery of REM sleep after sleep deprivation. This phenotype was not observed in Per3 null mice and Per3 wt BAC mice. Depression-like behavior was released by lorazepam and imipramine treatment. We also observed that transgenic mice have reduced fat mass and increased lean mass compared them wt littermates. We found that fat cell formation from mesenchymal stem cell is inhibited in transgenic mouse model. Here we identified that circadian clock gene Per3 links metabolism to sleep homeostasis.

2668W

Characterization of a novel mitofusin 2 knock-in mouse model. *A. Rebelo[1], A. Strickland[1], F. Zhang[1], J. Price[1], J. Silva[1], R. Wen[1], B. Bolon[2], J. Vance[1], S. Zuchner[1].* 1) HIHG, University of Miami, Miami, FL; 2) College of Veterinary Medicine, Columbus, OH.

Charcot-Marie-Tooth disease (CMT) comprises a group of heterogeneous peripheral neuropathies affecting 1 in 2,500 individuals. While mutations in several genes cause axonal degeneration in CMT, mitofusin 2 (MFN2) variants account for 91% of the most severe cases of the disease; making it the single most important cause of inherited peripheral axonal degeneration. MFN2 is an integral mitochondrial outer membrane protein that plays a major role in mitochondrial fusion and motility, however, the mechanism by which dominant mutations in this protein lead to neurodegeneration is still unknown. In order to investigate the underlying pathophysiology, we have generated a MFN2 knock-in mouse strain containing the R94W mutation originally identified in several CMT families. We have performed behavioral, morphological, cell biological, and biochemical studies to characterize the consequences of this mutation. In mice, homozygous inheritance of this mutation leads to a low birth weight, followed by death within a day. On the other hand, heterozygous inheritance causes no gross behavioral deficits for the duration of the mouse life (up to 1 year of life observed). Heterozygous mice do not differ from their wildtype littermates based on an array of behavioral tests and morphological analyses. Interestingly, heterozygous and homozygous mouse embryonic fibroblasts (MEF) show fragmented instead of elongated mitochondrial morphology compared to wild type mice. Homozygous tissues show both decreased mitochondrial DNA and ATP amounts. The presence of molecular mitochondrial abnormalities in MFN2-R94W mice underlines the usefulness of this model for interrogating the pathophysiology of CMT. In contrast to the human phenotype, the lack of overt behavioral deficits in heterozygous mice is likely attributable to the shorter life span and axon length in rodents. However, molecular studies will benefit from cells and tissues harvested from these mice and information on pathway interactions can be obtained by breeding these mice with other mutant mouse models.

2669T

Triply mouse model of Tay-Sachs disease shows altered brain ganglioside pattern. *V. Seyrantepe, Z. Timur.* Molecular Biology and Genetics, Izmir Institute of Technology, Izmir, Turkey.

Tay-Sachs disease is a severe lysosomal storage disorder caused by mutations in the HEXA gene coding for α subunit of lysosomal β-hexosaminidase A enzyme, which converts GM2 to GM3 ganglioside. HexA-/- mice, depleted of β-hexosaminidase A enzyme, remain asymptomatic to 1 year of age, owing to the ability of these mice to degrade stored GM2 ganglioside via sialidase (s) (sialidases Neu1, Neu2, Neu3 or Neu4) into glycolipid GA2 which further processed by β-hexosaminidase B, thereby completely bypassing the HexA defect. Previously we showed that mouse model with targeted disruption of both HexA and Neu4 genes (HexA-/-Neu4-/-) have epileptic seizures and accumulating GM2 ganglioside in brain tissue. To elucidate whether lysosomal Neu1 sialidase can also contribute to GM2 ganglioside degradation, a triply mouse model has been generated by breeding HexA-/-Neu4-/- mouse model and hypomorph mouse with 8–10% lysosomal Neu1 sialidase activity. Thin layer chromatography analysis of brain tissue gangliosides from affected mice at age of 4–5 and 8–9 months old showed lysosomal Neu1 sialidase deficiency slightly affect brain ganglioside pattern in triply mice suggesting that a mouse model with targeted distruption of Neu1 gene are required to show the involvement of lysosomal Neu1 sialidase in glycolipid degradation in HexA-/- and HexA-/-Neu4-/- mice. In our study, increased expression levels of lysosomal Neu1 and plasma membrane-associated Neu3 sialidases in brain tissue of HexA-/- and HexA-/-Neu4-/- mice also suggest contribution of other sialidases to the metabolic bypass pathway.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2670F

Deficiency of *SPECC1L* down-regulates PI3K-AKT signaling in the pathogenesis of Oblique Facial Clefts. *N.R. Wilson[1], A.J. Olm-Shipman[1], L. Pitstick[2], B.C. Bjork[2], I. Saadi[1].* 1) Department of Anatomy and Cell Biology, University of Kansas Medical Center, Kansas City, KS; 2) Department of Biochemistry, Midwestern University, Downers Grove, IL.

Orofacial clefts are among the most frequent birth defects, affecting one child in eight hundred, in the U.S. alone. Although the majority of orofacial clefts are limited to the lip and palate, a small subset involve a cleft that extends from the oral cavity to the eye, termed Tessier type IV or Oblique Facial Clefts (ObFCs). Previously, we identified *SPECC1L* (sperm antigen with calponin homology and coiled-coil domains 1-like) as the first gene mutated in ObFCs. SPECC1L is a novel cytoskeletal `crosslinking' protein, which interacts with both the actin and microtubule cytoskeletons. Further, SPECC1L deficiency in *Drosophila*, zebrafish and mammalian cells results in migration and adhesion defects, revealing a potential mechanism underlying ObFCs (Saadi et al., *AJHG* 2011). A critical question is through what signaling pathway(s) SPECC1L regulates these two cytoskeletal systems to control cell migration and cell adhesion. To explore this, we performed expression microarrays and antibody arrays on SPECC1L-deficient cells. Our analyses reveal a down-regulation of the PI3K-AKT signaling pathway, which coordinates extracellular cues relevant to both migration and adhesion. These cues include growth factor responses and Integrin-ECM interactions that culminate in the remodeling of the actin cytoskeleton. Thus, our data suggest that SPECC1L is a novel, craniofacial-specific mediator of PI3K-AKT signaling. Interestingly, SPECC1L-deficient U2OS osteosarcoma cells show a drastic change in cell shape - from an epithelial polygonal to a fibroblast-like elongated shape - *only* upon reaching confluence. These deficient cells also show an increase in F-actin fluorescence. We propose this remarkable change in cell morphology upon cell-cell contact is PI3K-AKT mediated and bears importance for many developmental events critical in craniofacial morphogenesis, including induction of cranial neural crest (CNC) cells as well as coordinate movement of CNC-derived sheets of cells in the facial prominences that fail to fuse in ObFCs. We are now validating our findings *in vivo* and are modulating the PI3K-AKT pathway with small molecule agonists and antagonists to characterize further the function of SPECC1L in the pathogenesis of ObFCs.

2671W

Transgenic rescue of BBS phenotypes in Bbs4 null mice. *X. Chamling[1,2], S. Seo[3], K. Bugge[1,2], C.C. Searby[1,2], A.V. Drack[3], K. Rahmouni[4], V.C. Sheffield[1,2].* 1) Pediatrics, University of Iowa, Iowa City, IA; 2) Howard Hughes Medical Institute, University of Iowa, Iowa City, IA; 3) Department of Ophthalmology and Visual Sciences, University of Iowa, Iowa City, IA; 4) Department of Internal Medicine, University of Iowa, Iowa City, IA.

Bardet-Biedl syndrome (BBS) is a genetically heterogeneous autosomal recessive disorder characterized by obesity, retinal degeneration, polydactyly, hypogenitalism and renal defects. Recent findings have associated the etiology of the disease with cilia, and BBS proteins have been implicated in trafficking various ciliary cargo proteins. To date, 17 different genes have been reported for BBS among which BBS1 is the most common cause of the disease followed by BBS10, and BBS4. A murine model of BBS4 is known to phenocopy most of the human BBS phenotypes, and it is being used as a BBS disease model. To better understand the in vivo localization, cellular function, and interaction of BBS4 with other proteins, we generated a transgenic BBS4 mouse expressing the human BBS4 gene under control of the beta actin promoter. The transgene is expressed in various tissues including brain, eye, testis, heart, kidney, and adipose tissue. These mice were further bred to express the transgene in Bbs4 null mice, and their phenotype was characterized. Here we report that despite tissue specific variable expression of the transgene, human BBS4 was able to complement the deficiency of Bbs4 and rescue all the BBS phenotypes in the Bbs4 null mice. These results provide an encouraging prospective for gene therapy for BBS related phenotypes and potentially for other ciliopathies.

2672T

Overexpression of Rai1 in mouse forebrain neurons is enough to cause most PTLS-like phenotypes. *L. Cao[1], J. Molina[2], P. Carmona-Mora[3], A. Oyarzo[4], J. Young[1,3], K. Walz[1,3,5].*1) Dr. John T. Macdonald Foundation Department of Human Genetics, Miller school of medicine, university of Miami, Miami, FL, USA; 2) Centro de Estudios Científicos, CECS, Valdivia, Chile; 3) John P. Hussman Institute for Human Genomics, Miller School of Medicine, University of Miami, FL, USA; 4) Cellular Communication Laboratory, Program of Molecular & Cell Biology, ICBM, Medical Faculty, University of Chile, Santiago, Chile; 5) Department of Medicine, Miller School of Medicine, University of Miami, FL, USA.

PTLS (Potocki Lupski Syndrome) is a genomic disorder associated with a duplication in 17p11.2. The clinical features of PTLS include intellectual disability, autistic features and developmental deficits. A PTLS mouse model, Dp(11)17/+, was generated by chromosome engineering, harboring a genomic duplication syntenic to the PTLS region. Many PTLS phenotypes were recapitulated in this mouse model including underweight, low fat content, hyperactivity, impaired memory and learning ability and social deficiencies. Further investigations showed that most of these phenotypes can be rescued by correcting the RAI1 dosage, indicating that RAI1 dosage is directly associated with PTLS phenotypes. Since RAI1 is highly expressed in brain neurons and it seems to be responsible for the neurobehavioral phenotypes in the PTLS mouse model, we wanted to define the specific brain regions that are related to the abnormal phenotype due to RAI1 overexpression. Here we generated a transgenic mouse line carrying transgene encoding Rai1 controlled by a tTa responsive promoter. By crossing this mouse line with camkII-tTa mice which express tTa downstream of CamkII promoter, we obtained bi-transgenic mice (I-Rai1) which over-express Rai1 specifically in mouse forebrain neurons. Surprisingly, I-Rai1 mice showed most of the PTLS-like phenotypes including underweight, lower fat content, hyperactivity, dominant behavior and impaired learning and memory ability. Anxiety is not observed in I-Rai1 mice, indicating that other brain regions might be involved in the development of Rai1-related anxiety behavior. After fear conditioning, we observed less c-fos induction in I-Rai1 mouse hippocampus which might be responsible for the impaired contextual memory ability. Other stress induced immediate early genes (IEG) are also dysregulated in I-Rai1 mouse hippocampus such as Bdnf and Growth hormone. Cell fractionation experiment showed Rai1 protein is mainly bound to the chromatin in neurons, suggesting Rai1 might affect the IEG expression by interacting with chromatin structures. Our results clearly showed that several PTLS phenotypes (body weight, social, memory and learning ability) are centrally governed. Moreover, Rai1 dosage is critical in forebrain neurons, and it affects the regulation of stress induced gene expression. In summary we present here a phenotype-tissue/cell/function specific relationship that will be crucial for advancing in the development of PTLS therapeutic strategies.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2673F

A Zebrafish model for Richieri-Costa-Pereira Syndrome: knockdown of eif4a3 gene results in craniofacial development disorder. *L. Alvizi[1,2], A. Weiner[3], F. Favaro[1,2,4], R. Zechi-Ceide[4], A. Richieri-Costa[4], M. Guion-Almeida[4], N. Calcaterra[3], M. Passos-Bueno[1,2].* 1) Genetics and Evolutive Biology, University of São Paulo, São Paulo, Sao Paulo, Brazil; 2) Human Genome Research Center, Institute of Biosciences of University of São Paulo, São Paulo, SP, Brazil; 3) Instituto de Biología Molecular y Celular de Rosario (IBR), CONICET- Área Biología General, Departamento de Ciencias Biológicas, Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario, Rosario. Argentina; 4) Department of Clinical Genetics, Hospital of Rehabilitation of Craniofacial Anomalies (HRCA), University of São Paulo, SP, Brazil.

Purpose: Richieri-Costa-Pereira Syndrome (RCPS) is a rare acrofacial dysostosis characterized by Robin sequence, cleft mandible, short stature and limb defects. RCPS is inherited in an autosomic recesive fashion, presenting a high rate of consanguinity, increased male mortality and 25% recurrence in sibs among patient's families. Recently, our group has identified a mutation in EIF4A3 gene (Eukariotic Initiation Factor 4A3) in homozygosity in RCPS patients as a putative molecular cause underlying this disorder (Favaro et al., ASHG 2012 submitted). Once zebrafish modeling for human genetic diseases has been emerging as a very informative tool in the developmental field, our objective was to knockdown eif4a3 expression in zebrafish embryos in order to phenocopy RCPS and study the malformation of craniofacial structures during embryonic development. Methods: We injected zebrafish embryos at the two-cells stage with either an eif4a3 translation-blocking or a mismatch control morpholinos (GeneTools). Development was observed at the points of 4, 15, 24 and 48 hours post-fertilization (hpf) and 5 days post-fertilization (dpf). Expression of p53 and p21 were assessed through qRT-PCR using 24 hpf embryos to check morpholino off-target effects and p53 apoptotic pathway activation by p21. For a deeper phenotype characterization, whole-mount RNA in situ hybridization for myoD, sox9b and col2a1 expression were performed. Besides, cartilage and bone formation were analysed by alcian blue and calcein stainings. Finally, by acridine orange staining we detected the apoptosis level assessment. Results/Conclusion: At 24 hpf, 71% of embryos in average treated with eif4a3 translation-blocking morpholino presented severely affected as well as dysmorphic craniofacial structures and a mortality rate of 29%. Forty-eight hpf embryos were smaller in size than controls and also presented craniofacial abnormalities. Larvae at 5 dpf had no mandible development. p53 mediated apoptosis pathway was not overexpressed with morpholino treatment since p53 and p21 expression was not significantly distinct from negative controls, excluding an off-target effect from morpholino. Our results suggest that eif4a3 embryonic expression is essential for normal craniofacial development and its impairment leads to craniofacial dysmorphisms in zebrafish. Besides, eif4a3 knockdown in zebrafish appears to be a good model for future research in Richieri-Costa-Pereira Syndrome. CEPID/FAPESP/CONICET.

2674W

A glial origin for periventricular nodular heterotopia caused by impaired expression of Filamin-A. *a. carabalona[1,2,3], s. beguin[1,2,3], e. pallesi-pocachard[1,2,3], e. buhler[3,4], c. pellegrino[1,2,3], k. arnaud[1,2,3], p. hubert[5], m. ouahla[5], jp. siffroi[6], s. khantane[1,2,3], i. coupry[7], c. goizet[7], a. bernabe-gelot[8], a. represa[1,2,3], c. cardoso[1,2,3].* 1) NMED, Parc Scientifique de Luminy, Marseille, France; 2) Université de la Méditerranée, UMR S901 Aix-Marseille 2, Marseille, France; 3) Inserm Unité 901, Marseille, France; 4) Plateforme postgenomique INMED-INSERM, Parc Scientifique de Luminy, Marseille, France,; 5) Service de Réanimation Pédiatrique et de Néonatologie, Hôpital Necker Enfants Malades, Paris, France,; 6) Service de Génétique et d'Embryologie Médicales, Hôpital A. Trousseau, Paris, France,; 7) Laboratoire Maladies Rares: Génétique et Métabolisme (MRGM), Université Bordeaux, EA4576 Bordeaux, France; 8) Département de Neuropathologie, Service d'Anatomie et Cytologie Pathologiques, Hôpital A. Trousseau, Paris, France.

Periventricular nodular heterotopia (PH) is a human brain malformation caused by defective neuronal migration that results in ectopic neuronal nodules lining the lateral ventricles beneath a normal appearing cortex. Most affected patients have seizures and their cognitive level varies from normal to severely impaired. Mutations in the Filamin-A (or FLNA) gene are the main cause of PH, but the underlying pathological mechanism remains unknown. Although two FlnA knockout mouse strains have been generated, none of them showed the presence of ectopic nodules. To recapitulate the loss of FlnA function in the developing rat brain, we used an in utero RNA interference-mediated knockdown approach and successfully reproduced a PH phenotype in rats comparable with that observed in human patients. In FlnA-knockdown rats, we report that PH results from a disruption of the polarized radial glial scaffold in the ventricular zone altering progression of neural progenitors through the cell cycle and impairing migration of neurons into the cortical plate. Similar alterations of radial glia are observed in human PH brains of a 35-week fetus and a 3-month-old child, harboring distinct FLNA mutations not previously reported. Finally, juvenile FlnA-knockdown rats are highly susceptible to seizures, confirming the reliability of this novel animal model of PH. Our findings suggest that the disorganization of radial glia is the leading cause of PH pathogenesis associated with FLNA mutations.

2675T

Mining phenotype and disease models from the Mouse Genome Informatics Resource. *J.T. Eppig, S.M. Bello, H. Dene, A. Anagnostopoulos, H. Onda, R. Babiuk, M.N. Knowlton, M. Tomczuk, C.L. Smith.* Jackson Laboratory, Bar Harbor, ME.

The laboratory mouse is the premier model for studying human biology and disease. Large-scale forward and reverse mutagenesis projects are making available mutations in every gene. The phenotyping of mutant genotypes promises an explosion in analyzing gene function and in discovering and validating new models for human disease.

The Mouse Genome Informatics (MGI, www.informatics.jax.org) database is a treasure trove for biologists and clinicians seeking to correlate mouse and human phenotypes. MGI catalogs all mouse mutant alleles, with key molecular attributes, captures descriptions of phenotypes, and links mouse genotypes experimentally determined to be models of human disease with corresponding disease terms in OMIM. Standardization of nomenclature and application of bio-ontologies, including the Mouse Anatomical (MA) Dictionary and Mammalian Phenotype (MP) Ontology, ensure that data are consistently annotated, making robust searches possible.

Currently >740,000 alleles are cataloged in >20,000 genes. More than 32,000 alleles are propagated in live mice or are available as cryopreserved embryos or sperm; the rest existing only in ES cell lines. The amount of phenotypic data available is very rich, with nearly 44,000 unique genotypes curated with >223,000 MP terms. This substantial data set makes possible mining of MGI to look for phenotype associations (e.g., What phenotypes are found concordantly? What genotypes have similar phenotype profiles?). Such questions can uncover genes in a common pathway, implicate genes with particular functions, or identify relationships between diseases and their underlying genetic defects. User access to mouse phenotype and disease data is achieved in a number of ways from gene, allele, phenotype, or disease perspectives.

MGI will integrate data from the new International Mouse Phenotyping Consortium (IMPC) Centers, which are collecting systematic large-scale phenotyping data from knockouts and conditional-ready knockouts of every mouse gene. These new data, in context of existing phenotypic data, will be an unparalleled opportunity for mining phenotype-genotype-disease correlations and provide a platform for therapeutic discovery. Supported by NIH grant HG000330.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2676F

Molecular Genetics and Modeling of a Novel BBS5 Mutation. *M.H. Al-Hamed[1, 2], C.V. Lennep[2], F. Al-Fadhly[3], R.J. Simms[2], B.F. Meyer[1], J.A. Sayer[2].* 1) Genetics Department, King Faisal Specialist Hospital and Research Centre, P. O. Box 3354, Riyadh, Saudi; 2) Institute of Genetic Medicine, Newcastle University, Central Parkway, Newcastle upon Tyne, NE1 3BZ, United Kingdom; 3) Department of Pediatrics, Maternity & Children's Hospital, PO Box 6205,Al Madina Al Munawara, Saudi Arabia.

Bardet-Biedl Syndrome (BBS; OMIM 209900) is an autosomal recessive disorder clinically characterized primarily by obesity, progressive early-onset retinal degeneration, polydactyly, hypogenitalism, cognitive impairment and renal failure. BBS is a ciliopathy, with the multisystem syndrome secondary to defects in ciliary / basal body structure and function. To date, 15 BBS genes have been identified. We investigated a consanguineous family with 3 children presenting with BBS. Firstly, to identify the disease locus within this family, we carried out a genome wide linkage search using Affymetrix GeneChipR Human Mapping 250K Arrays. All three affected children shared one region of homozygosity on chromosome 2. This region of homozygosity contained the known BBS gene, BBS5, which was screened by direct sequencing of the PCR products of all coding regions (12 exons). In exon 12 of the BBS5 gene we found a novel mutation c.967insT (p.A323CfsX55) predicted to cause a frameshift which reaches a stop codon 168 nucleotides downstream. The mutation results in a predicted elongated peptide that has 378 amino acids instead of 341 in the wild type BBS5 protein. The three affected siblings were all homozygous for the mutation, whilst both parents and unaffected son were heterozygous for the mutation. Using zebrafish embryos, morpholino antisense oligonucleotides were used to modify the gene expression of BBS5. Morphant embryos displayed a typical ciliopathy phenotype with body axis curvature, cardiac abnormalities (including situs inversus) and pronephric duct dilatation. HEK293 were transfected with GFP-tagged mutant BBS5 (c.967insT) and wild type BBS5-GFP using Lipofectamine 2000. Mutant BBS5 failed to localize to the basal body in these cells. In conclusion we describe a family with clinical features of BBS, together with a novel mutation in BBS5. Modeling this disease in zebrafish mimics the human disease, and pathogencity of the novel BBS5 mutation is demonstrated by mislocalisation in renal epithelial cells.

2677W

Limb differences in muscle dystrophy and gene expression in a murine model of MDC1A suggest possible targets for therapy. *J.S. Benjamin, R. Marx, R.D. Cohn.* McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins School of Medicine.

Merosin-deficient congenital muscular dystrophy (MDC1A) is an autosomal recessive congenital muscular dystrophy caused by mutations in the *LAMA2* subunit of the lamanin heterotrimeric extracellular matrix protein. The disease presents neonatally or shortly after birth, shows a range of phenotypes (including general muscle weakness, a delay of normal motor milestones and hypotonia), and currently has no known cures or effective therapies. Several mouse models for the disease exist including laboratory induced mutants in Lama2 as well as a few spontaneous lines. One such spontaneous line, the Dy2j in C57BL/6, contains a deletion in *Lama2* that causes a partial loss of function leading to muscle weakness several weeks after birth. We present an interesting phenotypic finding that while the hind limbs of these mice show weakness and paralysis after several weeks, the forelimb show near normal function until 12 months of age. Histological analysis combined with measurement of the area of fibrosis demonstrate that the hind limb tibialis anterior muscle shows approximately five-fold increase in area of fibrosis compared to the forelimb triceps muscle. Since the forelimb muscle seems to be protected and instead shows a milder muscular dystrophy phenotype when compared to other tissues, we hypothesize that genes differentially expressed between the triceps and tibialis anterior muscles could be responsible. To assess these differences, we ran an Affymetrix Gene 1.0 ST mouse expression array on 3 month old wild type and Dy2j mice for both tibialis anterior and triceps. Interestingly, among the genes most up regulated in the triceps compared to tibialis anterior, we see genes involved with polyamine biosynthesis/metabolism (which has been implicated in cell proliferation and migration) including *Amd1*, *Smox*, and *Zmynd17*, as well as *Spns2* (involved in myocardial precursor migration). Additionally, in the genes most increased in the tibialis anterior compared to triceps we see an increase in *Mgp*, a negative regulator of cell growth. Further functional studies on these genes and other candidates will be necessary to determine if they can play a role in protecting muscle from dystrophy, and can become viable targets for therapy.

2678T

***FREM1* Deficiency Causes a Syndromic Form of Congenital Diaphragmatic Hernia in Humans and Mice.** *T.F. Beck[1], O. Shchelochkov[2], Z. Yu[1], B. Kim[1], H. Zaveri[1], B. Lee[1,3], D.W. Stockton[4], M. Justice[1], D.A. Scott[1,5].* 1) Molec & Human Gen, Baylor College Med, Houston, TX; 2) Department of Pediatrics, The University of Iowa, Iowa City, IA; 3) Howard Hughes Medical Institute, Houston, TX; 4) Department of Pediatrics, Wayne State University, Detroit, MI; 5) Department of Molecular Physiology and Biophysics, Baylor College of Medicine, Houston, TX.

Congenital diaphragmatic hernia (CDH) is a life-threatening birth defect that affects approximately one in 4,000 newborns and is a feature of several genetic syndromes. Autosomal recessive mutations in *FREM1* have been shown to cause eye defects and renal anomalies as part of Bifid Nose and Anorectal and Renal malformations (BNAR) and Manitoba-Oculo-Tricho-Anal (MOTA) syndromes, but have not been previously implicated in the development of CDH. However, we recently identified a female patient with a left-sided CDH and features suggestive of BNAR. Copy number and sequence analysis of her *FREM1* gene revealed two mutations, a maternally-inherited, approximately ~86 kb deletion affecting only*FREM1*, and a paternally-inherited point mutation affecting the splice donor site of exon 28. This suggests that the spectrum of phenotypes associated with autosomal recessive *FREM1*mutations includes CDH. We also note that *FREM1* is located in a region of chromosome 9 that is recurrently deleted in individuals with CDH. Further evidence for the role of *FREM1* in the development of CDH comes from our work on a novel mouse strain, termed ***eye2***, which carries a homozygous truncating mutation in *Frem1*. *Frem1*$^{eye2/eye2}$ mice develop eye defects and renal agenesis similar to those seen in BNAR and MOTA syndromes but also develop anterior CDH and lung segmentation defects. CDH in these mice occurs in the muscularized portion of the diaphragm located directly behind the sternum. We confirmed that *Frem1* is expressed in the anterior portion of the developing diaphragm and found that *Frem1*$^{eye2/eye2}$ embryos had decreased levels of cell proliferation in their diaphragms when compared to wild type embryos. The penetrance of CDH in *Frem1*$^{eye2/eye2}$ mice was found to be highly dependent on strain background suggesting the existence of genetic modifiers. We conclude that *FREM1* plays a critical role in the development of the eye, kidney, lung, and diaphragm and that FREM1 deficiency can predispose to the development of a syndromic form of CDH in both humans and mice.

2679F

Characterization of visual impairment in a Wfs1 mouse model of Wolfram syndrome. *C. Delettre[1], D. Bonnet-Wersinger[1], Y. Tanizawa[3], G. Lenaers[1], C. Hamel[2].* 1) INSERM U1051, Montpellier, France, PhD; 2) INSERM U1051, Montpellier, France, MD; 3) Department of Bio-Signal Analysis, Yamaguchi University Graduate School of Medicine, Ube, Japan, MD.

Wolfram syndrome is a childhood onset rare genetic disease (1/180,000) featuring diabetes mellitus and optic neuropathy unavoidably progressing towards legal blindness before the age of 20. A Wfs1-/- mouse model has been generated (Ishihara et al, 2004), showing pancreatic beta cell atrophy. Nothing is known about the visual function of Wfs1-/- mouse. We have studied the visual impairment of these mice by electrophysiology and histopathology. Electroretinogram testing (ERG, retinal function) and visual evoked potentials (VEPs, visual pathway) were performed in Wfs1-/- and Wfs1+/+ mice at 3, 6, 9, and 12 months of age. Total nuclei and retinal ganglion cell (RGC) populations of nerve fiber layer were quantified from Brn3a immunolabeling of retinal sections. RGC axonal loss was also quantified by electron microscopy in transversal optic nerve sections. Grating acuity and contrast sensitivity were measured based on the optomotor tracking responses on a virtual apparatus (OptoMotry; Cerebral Mechanics, Lethbride, Alberta, Canada). Both a- and b-wave ERG amplitudes were slightly reduced at 12 months. In contrast, VEPs showed progressive decrease of N+P amplitudes in Wfs1-/- by 30, 40 and 50% at 6, 9 and 12 months, respectively, as compared with controls. Brn3a positive RGC and total nuclei in RGC layer were not significantly lost in Wfs1-/- genotype. Transmission electron microscopy analysis of 10 month-old Wfs1-/- mice determined little reduction in axonal density. Frequency and contrast thresholds of optokinetic tracking reflex remained normal in Wfs1-/- mice. Progressive VEPs alteration with minimal cell loss suggests functional defect of the signal conduction in the optic pathway. However, visual acuity is preserved and immunohistology didn't show any alteration of RGC layer in the retina. We conclude of a mild phenotype in the exon2 Wfs1-/- mice. Compensation, specific murine phenotype or truncated Wfs1 proteins can be assumed to explain this unaltered visual phenotype.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2680W

Gene therapy for X-linked retinitis pigmentosa in a knockout mouse model of *RP2*. *S. Mookherjee[1], L. Li[2], A.K. Ghosh[3], N. Khan[3], P. Colosi[1], H. Khanna[2], A. Swaroop[1].* 1) Neurobiology - Neurodegeneration & Repair laboratory, National Eye Institute, National Institutes of Health, Bethesda, MD; 2) Department of Ophthalmology, University of Massachusetts Medical School, Worcester, MA; 3) Department of Ophthalmology and Visual Sciences, University of Michigan, Ann Arbor, MI.

X-linked retinitis pigmentosa (XLRP) is a relatively severe form of retinal degeneration that is characterized by early onset night blindness, progressive loss of peripheral vision and pigmentary retinopathy. Mutations in the *RP2* gene account for 10–20% of XLRP. Though the precise function of this ubiquitously expressed protein is still poorly understood, RP2 may play a role in pericetriolar vesicle transport between Golgi and primary cilium. No specific treatment is available for the patients with XLRP. Gene replacement therapy presents a viable option but, to date, no animal models have been available for efficacy studies. Here we report the generation and characterization of a conditional *Rp2* knockout model that is being used to evaluate gene transfer strategies for the treatment of XLRP disease. We generated an $Rp2^{flox}$ mouse by introducing *loxp* sites flanking exon 2. The $Rp2^{flox}$ mouse was then bred to CAG-Cre (a ubiquitous Cre expressing transgenic line) mice and male progeny was analyzed further for photoreceptor development and function. Immunoblot and immunofluorescence analysis of the mutant retina did not reveal detectable expression of the RP2 protein. Electroretinography (ERG) analysis of $Rp2^{CAG\text{-}Cre/lox}$ mice revealed a progressive loss of cone function starting at around 5 months of age followed by decrease in rod function by 6 months of age. We also detected reduction in the thickness of outer nuclear layer of the retina starting at approximately 4 months of age. Further analysis of retinal whole mounts at 4 months of age showed a decrease in total cone population as well as shortened outer segments and swollen soma, as revealed by staining with peanut agglutinin (PNA). Staining of cryosections of the retina of 4 month old $Rp2^{CAG\text{-}Cre/lox}$ mice showed reduction of M- and S-opsin staining as well as mislocalization of S-opsin in cones. We have performed unilateral subretinal injection of an AAV8 sc RK RP2 vector at 2 months of age. The vector genome has a self-complementary structure and is composed of the rhodopsin kinase promoter, a CMV/b-globin intron, the RP2 cDNA and the beta globin polyadenylation site. Treated mice are being evaluated for therapeutic effect on retinal structure and visual function. Our studies provide the framework for elucidating the mechanism of XLRP-RP2 pathogenesis and development of gene-based therapy for this blinding disease.

2681T

Altered splicing of BIN1 muscle-specific exon in human and Great Danes with progressive centronuclear myopathies. *J. Bohm[1], N. Vasli[1], M. Maurer[2], B.S. Cowling[1], W. Kress[3], G.D. Shelton[4], U. Schara[5], J. Anderson[6], R. Herrmann[7], J. Weis[8], L. Tiret[2], J. Laporte[1].* 1) Department of Translational Medicine and Neurogenetics, IGBMC, Illkirch, France; 2) Génétique Fonctionnelle et Médicale, Ecole Nationale Vétérinaire d'Alfort, Maisons-Alfort, France; 3) Department of Human Genetics, Julius-Maximilian University, Würzburg, Germany; 4) Department of Pathology, University of California at San Diego, La Jolla, USA; 5) Department of Paediatric Neurology, University of Essen, Essen, Germany; 6) Institute of Comparative Medicine, University of Glasgow Veterinary School, Glasgow; 7) Department of Paediatrics I, University Hospital Essen, Essen, Germany; 8) The Institute for Neuropathology, Medical Faculty, RWTH Aachen University, Aachen, Germany.

Excitation-contraction coupling (ECC) occurs at the skeletal muscle triad and converts an electrical stimulus into mechanical muscle work. The triad is composed of the sarcoplasmic reticulum and a plasma membrane invagination, called the T-tubule. One of the key factors in muscular membrane remodeling is amphiphysin 2 (BIN1), located at the triad. We identified and characterized BIN1 mutations affecting the splicing of the muscle-specific exon 11 and involving a rapidly progressing myopathy in human and dogs. The impact of both mutations on the AG acceptor splice site was verified by mRNA analysis. The canine muscle disorder is known as IMGD (Inherited Myopathy in Great Danes) and typically starts before 10 months of age and most of the dogs are euthanized before 18 months of age due to severe debilitating muscle weakness. The human patients presented with a disease onset at 3.5 years and a rapidly progressive and fatal myopathy. Mutations in the ubiquitous BIN1 exons have previously been associated with recessive centronuclear myopathy (CNM) in patients with a clinically milder phenotype than described here. In addition, mis-splicing of BIN1 exon 11 was associated to myotonic dystrophies. Human and canine muscle histopathology displayed prominent nuclear internalization, atrophy, and unusual rippled sarcolemma. Ultrastructural analysis confirmed nuclear internalization, mitochondrial accumulations around the nuclei and myofibrillar disarray. Immunolabelling revealed a defective structure of the triad, and myotubes transfected with the exon 11 containing isoform showed massive tubulation, whereas the construct without exon 11 did not have this effect. Taken together, our data suggest that the amphiphysin 2 muscle isoform plays an important role in triad buildup and/or maintenance and that defective excitation-contraction coupling is a primary cause of the progressive human and dog diseases. Our study is the first genetic proof of the direct implication of BIN1 exon 11 in the muscle-specific function of amphiphysin 2. We provide the evidence that an alteration of the physiological function of the membrane-deforming properties of amphiphysin 2 and its alternative splicing-dependant activity is a common mechanism for canine and human myopathies. The identification and characterization of a spontaneous canine model represents a faithful model for preclinical trials. This is of particular interest, as BIN1 null mice are perinatally lethal.

2682F

Up-regulation of androgen-binding protein is associated with testicular phenotype in $Smarcal1^{del/del}$ mice. *A. Baradaran-Heravi, C.F. Boerkoel.* Department of Medical Genetics, Child and Family Research Institute, University of British Columbia, Vancouver, BC, Canada.

Schimke immune-osseous dysplasia (SIOD) is a fatal childhood disorder of spondylo-epiphyseal dysplasia and short stature, renal failure, T-cell deficiency, and small testes. SIOD is caused by bi-allelic mutations of *SMARCAL1*, which encodes a DNA annealing helicase and is involved in DNA repair, replication, recombination and transcription. Similar to SIOD patients $Smarcal1^{del/del}$ male mice have small testes characterized by degeneration of seminiferous tubules and a relative increase in interstitial cells. We hypothesized that $Smarcal1^{del/del}$ testis phenotype arises cell autonomously from defective proliferation of germ cells. Using immunohistochemistry and immunofluorescence, we found that Smarcal1 was highly expressed in Sertoli and Leydig cells but not in germ cells. This was confirmed by RT-PCR on the isolated germ cells and Sertoli cells. This finding makes a major role for Smarcal1 in DNA replication or recombination of sperm precursors unlikely. Based on the phenotypic similarity of $Smarcal1^{del/del}$ testes to the ones of aromatase-deficient and androgen-binding protein (ABP) over-expression mice, we looked for expression levels of *Cyp19* (encodes aromatase) and *Shbg* (encodes ABP) genes within $Smarcal1^{del/del}$ testes. We found that similar to ABP over-expression testes, both of these genes were up-regulated in $Smarcal1^{del/del}$ testes particularly around puberty. We are currently investigating whether the pathology in $Smarcal1^{del/del}$ testes is associated with overexpression of ABP. Such a mechanism would fit with SMARCAL1 homologues functioning as modulators of gene expression and disease arising when their deficiency results in alteration of expression beyond a threshold.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2683W

A Mouse Model of *mut*$^{-}$ Methylmalonic Acidemia (MMA). *J.S. Sénac[1], V.H. Aswani[2], J.R. Sysol[1], I. Manoli[1], C.P. Venditti[1].* 1) Genetics and Molecular Biology Branch, National Institute of Health, Bethesda, MD; 2) Marshfield Clinic, Department of Internal Medicine & Pediatrics, Marshfield, WI.

Methylmalonic acidemia (MMA) is caused by mutations of the mitochondrial gene methylmalonyl-CoA (MUT). We have previously studied MMA using a mouse model mimicking the most severe form of the disease (*mut*^{0}) and characterized by a complete absence of Mut. *Mut*$^{-/-}$ mice are fragile, display lethality and subsequently present limited use to study the induction, control, and rapid treatment of some clinical manifestations, such as metabolic strokes, pancreatitis and kidney disease. To overcome these difficulties, we developed tissue specific transgenic mice but these animals restrict the study of MMA to selected organs. In order to circumvent these barriers, we developed a new mouse model with a milder enzymatic phenotype (*mut*$^{-}$). Among the wide spectrum of *mut*$^{-}$ mutations, the mechanism of enzymatic impairment has been defined only for a few, such as p.G717V, documented to be a putative Km mutant. Following expression and kinetic assay of the homologous mouse mutant (G715V) and the wild type (3C) MUT enzymes in yeast, we generated transgenic mice that ubiquitously express the mutant cDNA (*Mut*$^{-/-}$;Tg$^{INS\text{-}CBA\text{-}G715V}$) and transgenic mice that express the wild type cDNA control (*Mut*$^{-/-}$;Tg$^{INS\text{-}CBA\text{-}3C}$). *Mut*$^{-/-}$;Tg$^{INS\text{-}CBA\text{-}G715V}$ mice were born in Mendelian proportions, can be maintained regular diet, and achieve 80–90% of control littermate weight through the first year of life. The p.G715V Mut enzyme is expressed in liver, kidney, brain, and muscle but has a limited activity as measured by the oxidation of 1-^{13}C propionate into $^{13}CO_2$. The *Mut*$^{-/-}$;Tg$^{INS\text{-}CBA\text{-}G715V}$ mice metabolized 46.8±4.3% of the label in 25 min, compared to 63.1±3.1% in *Mut*$^{+/-}$ and 10±1.1% in *Mut*$^{-/-}$. Baseline plasma MMA levels (μM) were 159.4±53.6 in transgenic mice, compared to <5 in controls, which is similar to levels observed in G717V MMA patients. While other MMA mouse models are susceptible to death on an unrestricted diet, *Mut*$^{-/-}$;Tg$^{INS\text{-}CBA\text{-}G715V}$ can tolerate a regular diet and survive a high protein diet challenge despite a significant increase in plasma MMA level, rapid weight loss, and formation of megamitochondria in the liver and kidney. Finally, these mouse model fail to respond to hydroxocobalamin treatment. This novel murine model mimics physiologic and phenotypic observed in mild MMA patients. These mice provide a new inducible platform to study the pathophysiology and treatment of clinical syndromes associated with MMA, such as metabolic strokes, pancreatitis and chronic kidney disease.

2684T

Allelic heterogeneity contributes to variable ocular dysgenesis caused by *Col4a1* and *Col4a2* mutations in mice. *M. Mao[1], DS. Kuo[1], C. Labelle-Dumais[1], M. Jeanne[1], J. Favor[2], DB. Gould[1].* 1) Departments of Ophthalmology and Anatomy, Institute for Human Genetics, UCSF School of Medicine, San Francisco, CA, 94143, USA; 2) Institute of Human Genetics, Helmholtz Zentrum München, Neuherberg, Germany, D-85764.

Type IV collagen alpha1 (COL4A1) and alpha2 (COL4A2) are major components of the basement membrane. Mutations in *COL4A1* and *COL4A2* cause a broad spectrum of disorders including cerebrovascular diseases, ocular dysgenesis, renal and muscle defects. Not all patients with *COL4A1* and *COL4A2* mutations have the same disease, and genetic context and allelic heterogeneity are possible explanations for the phenotypic variability. We have previously shown that a splice site mutation, *Col4a1*$^{\Delta ex41/+}$ causes ocular dysgenesis including anterior segment dysgenesis and optic nerve hypoplasia in mice. In addition, we also found that genetic context can modify the severity of both diseases. The purpose of this study was to test if allelic heterogeneity also contributes to a variable ocular dysgenesis. To this end, we bred eight different *Col4a1* mutations and one *Col4a2* mutation into the same genetic background (C57BL/6J) and compared the severity of ocular dysgenesis between these mutations. We found that all mutations caused ocular dysgenesis; however, the pathology differed in penetrance and severity. Molecular analysis revealed variable biosynthetic consequences of these mutations. Most mutations, but not all, showed increased intracellular and decreased extracellular levels of COL4A1 and COL4A2 proteins. Our results demonstrate phenotypic and molecular heterogeneity among *Col4a1* and *Col4a2* mutations. Whether the phenotypic heterogeneity reflects a mechanistic heterogeneity remains to be tested.

2685F

Tmprss3 loss of function impairs cochlear inner hair cell BK channel properties. *B. Delprat[1,2], L. Molina[3], L. Fasquelle[1,2], R. Nouvian[1,2], N. Salvetat[3], H. Scott[4], M. Guipponi[5], F. Molina[3], J.L. Puel[1,2].* 1) Inserm U 1051, Institut des Neurosciences, Hôpital Saint Eloi, 34091 Montpellier, France; 2) Université de Montpellier, 34091 Montpellier, France; 3) SysDiag UMR 3145 CNRS / Bio-Rad, Cap Delta / Parc Euromédecine, 1682 rue de la Valsière, CS 61003, 34184 Montpellier Cedex 4, France; 4) Department of Molecular Pathology, the Institute of Medical and Veterinary Science, SA Pathology and The Centre for Cancer Biology, Box 14 Rundle Mall Post Office, Adelaide, SA 5000 and the School of Medicine University of Adelaide, SA, 5005, Australia; 5) Department of Genetic Medicine and Development, University of Geneva Medical School and University Hospitals of Geneva, 1211 Geneva 4, Switzerland.

Mutations in TMPRSS3 cause human hereditary deafness DFNB8/10. Initial characterization of Tmprss3Y260X-mutant mice reveals that the protein has a critical role in hair cell survival at the onset of hearing. In order to unravel the role of Tmprss3 in cochlear hair cell physiology, we characterized the altered protein network of proteins mutant mice compared to wild-type animals using 2D gels followed by mass spectrometry. This proteomic approach led to the identification of four putative altered pathways highlighted by four hub proteins: potassium BK channels, 14-3-3 β, 14-3-3 ζ and PSD95. Using immunohistochemistry and patch-clamp analysis of inner hair cells (IHCs), we demonstrated that Tmprss3 is essential for the expression of BK channels at the IHC plasma membrane and that the protein modulates the biophysical properties of the remaining BK channels. In conclusion, our data demonstrate, for the first time, that a type II serine protease, Tmprss3, is able to modulate potassium BK channels.

2686W

Interaction with p97/VCP alters the intracellular itinerary of the copper-transporting ATPase, ATP7A. *L. Yi, S.G. Kaler.* Molecular Medicine Program, NICHD/NIH, Bethesda, MD.

ATP7A is a P-type ATPase that regulates cellular copper homeostasis via activity at the trans-Golgi network (TGN) and plasma membrane (PM). Mutations in ATP7A lead to Menkes disease or its variants, occipital horn syndrome, or adult-onset distal motor neuropathy, a newly discovered condition for which the precise mechanism is obscure. Previously, we characterized the two known distal motor neuropathy mutations, T994I and P1386S, and demonstrated subtle defects in intracellular trafficking resulting in increased plasma membrane (PM) localization, as well as abnormal interaction between ATP7A-T994I and p97/VCP, a TGN-resident protein mutated in two autosomal dominant conditions featuring motor neuron degeneration: 1) inclusion body myopathy and 2) amyotrophic lateral sclerosis (Lou Gehrig's disease). Using a series of protein truncation and biochemical methods in combination with cell biological techniques, we identified the p97/VCP and ATP7A interacting binding sites: the N-terminal globular domain of p97/VCP, and two regions in the cytosolic portion of ATP7A. Binding of p97VCP to the first ATP7A binding site, in the middle of the wild type protein near the Actuator domain, was associated with formation of cytosolic ATP7A protein aggregates on confocal microscopy. The second site, found only in the T994I mutant protein, is located in the carboxyl-half of the protein between the 6th and 7th transmembrane segments of ATP7A (amino acid residues 1108–1356). This covert p97/VCP binding site is evidently exposed by changes in ATP7A conformation associated with the nearby T994I mutation. Deletion of the C-terminal tail of ATP7A (residues 1416–1450) results in PM retention of the WT, but not the T994I mutant protein, indicating that the aberrant p97/VCP interaction impairs ATP7A vesicular trafficking. These findings help illuminate the mechanism underlying ATP7A-related distal motor neuropathy and establish a link between p97/VCP and genetically distinct forms of motor neuron degeneration.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2687T

JNK phosphorylation is responsible for the greater osteogenic potential of periosteum derived fibroblasts harboring the FGFR2 S252W mutation (Apert Syndrome). *R. Atique[1], E. Yeh[1], H. Matsushita[2], N. Alonso[2], MR. Passos-Bueno[1].* 1) Human Genome Research Center-University of São Paulo, São Paulo, Brazil; 2) Department of Plastic Surgery, School of Medicine, University of São Paulo, São Paulo, Brazil.

Apert syndrome is a syndromic craniosynostosis characterized by premature fusion of the coronal sutures and symmetric syndactyly of the hand and feet. It is caused by one of two possible gain of function mutations in the FGFR2 gene (S252W and P253R). The mutated FGFR2 loses specificity for its ligands leading to its overactivation. Gene expression array studies have shown that one of the main differentially expressed genes in Apert syndrome fibroblast cultures is DUSP2, whose protein is a phosphatase with specificity to the MAPK family of kinases, which includes ERK and JNK. Alterations in JNK phosphorylation have been previously linked to osteogenic differentiation. In this study we aim to evaluate the role of JNK in the augmented osteogenic diferential potential of fibroblasts from the periosteum covering the coronal suture region and if there is any difference in the phosphorylation levels of JNK by activation of the FGFR2 S252W receptor when compared to the wild-type. We first assessed the difference in JNK phosphorylation (P-JNK) levels between FGFR2 S252W fibroblasts and controls, we found no statistically significant differences. We hypothesized that culture condition may have interfered in these results. To test this hypothesis, we compared the JNK expression in starving and non-starving conditions. Starvation increased JNK expression significantly more in S252W cells (1500 %) than in control cells (300%), showing that the JNK signaling regulation is disturbed in the S252W fibroblasts. We then evaluated, through Elisa assay, if the activation of FGFR2 by FGF2 treatment altered JNK phosphorylation. The FGF2 treatment increased the JNK phosphorylation levels in S252W fibroblasts by 25% and decreased them in controls by 15%. Lastly we evaluated the effects of JNK inhibition on osteogenic differentiation by inhibiting the kinase activity of JNK. The complete inhibition of JNK activity reversed the phenotype leading the osteogenic differentiation levels in FGFR2 S252W fibroblasts similar to controls (P<0,05) . In conclusion, we demonstrated that JNK is involved with the greater osteogenic potential in fibroblasts harboring the Apert syndrome mutation (FGFR2 S252W).

2688F

Novel p.Leu795Pro INSR mutation with decreased autophosphorylation in a patient with Donohue syndrome. *T. Hovnik[1,2], K. Trebus{{ak}}ak Podkrajs{{ak}}ek[2], J. Kovac{{ak}}[1,2], N. Bratanic{{ak}}[1], T. Battelino[1].* 1) Department of Pediatric Endocrinology, Diabetes and Metabolic Diseases, University Children's Hospital, UMC, Ljubljana, Bohoric{{ak}}eva 20, 1000 Ljubljana, Slovenia; 2) Center for Medical Genetics, University Children's Hospital, UMC, Ljubljana, Vrazov trg 1, 1000 Ljubljana, Slovenia.

Donohue syndrome (leprechaunism; OMIM *246200) is a rare, recessively inherited disorder of extreme insulin resistance due to mutation in the insulin receptor gene (INSR). The pathogenesis of insulin resistance is due to a cellular defect in insulin binding, receptor autophosphorylation and tyrosine kinase activity on insulin-stimulated biological activity. We report a patient with pronounced clinical picture of leprechaunism who developed severe progressive hypertrophic obstructive cardiomyopathy (HOCM) and renal tubular dysfunction which improved on continuous subcutaneous infusion of rhIGF-I. To investigate the molecular basis of insulin resistance, INSR gene molecular analysis and insulin receptor (IR) autophosphorylation on cultured fibroblasts were performed. Sequencing analysis of the 22 exons and intron-exon junctions revealed a homozygous novel missense mutation c.2465T>C located in exon 12, leading to a substitution of the leucine to proline at amino acid position 795 (p.Leu795Pro) at the conserved site in the extracellular domain of the IR β subunit. Insulin-induced autophosphorylation in the patient was reduced by 87.3% when compared to control, and by 86% when compared to his mother's fibroblasts. Functional characterization showed that novel INSR gene mutation in insulin receptor β subunit is causing a syndrome of extreme insulin resistance due to decreased insulin receptor autophosphorylation rather than insulin binding impairment.

2689W

Post-transcriptional regulation of DMD by its 3'UTR. *A. Larsen, M. Howard.* Human Genetics, University of Utah, Salt Lake City, UT.

Mutations in the DMD gene can cause either the severe Duchenne Muscular Dystrophy (DMD) or the milder Becker Muscular Dystrophy (BMD). Mutations that lead to the milder Becker phenotype are generally in-frame deletions that lead to a partially functional dystrophin protein. This observation has led to potential therapeutic strategies, such as exon-skipping, that could be used to treat Duchenne patients by expressing a mini-dystrophin. Although transcriptional regulation of dystrophin has been well characterized, the post-transcriptional regulation of DMD expression is not well understood. Here, we report the identification of highly conserved regions in the DMD 3'UTR that regulate mRNA stability and translation in differentiated C2C12 myotubes. To investigate the role the DMD 3'UTR has on gene expression, we generated Renilla Luciferase constructs containing the full-length DMD 3'UTR and several deletions within the DMD 3'UTR. Using the dual-luciferase reporter assay, we show that deleting an ~80 bp region in the 3'UTR results in ~50% decrease in protein levels with no significant decrease in steady-state RNA levels suggesting that this region acts as a translational enhancer. We also used qPCR to measure RNA levels of C2C12 myoblasts and myotubes treated with Actinomycin D and show that deleting the conserved regions of the DMD 3'UTR results in a modest decrease in mRNA stability.

2690T

NPHP10 (SDCCAG8) interacts with components of the multi-aminoacyl-tRNA synthetase complex. *K. Weihbrecht[1,2], M. Humbert[1,2], V. Sheffield[1,3], S. Seo[2].* 1) Dept. of Pediatrics, Univ. of Iowa, Iowa City, IA; 2) Dept. of Ophthalmology and Visual Sciences, Univ. of Iowa, Iowa City, IA; 3) Howard Hughes Medical Institute.

Nephronophthisis (NPHP) is a recessive kidney disorder that is the leading cause of early onset, end-stage renal failure. Many proteins mutated in cystic kidney disease have been shown to localize to the primary cilia and centrosomes, providing a coalescing mechanism for NPHP-related ciliopathies (NPHP-RC). Aside from renal failure and kidney cysts, retinal degeneration and dysplasia or degeneration of the cerebellum are also seen in many NPHP-RCs. SDCCAG8 is a nephronophthisis gene (NPHP10), with patients exhibiting retinal and renal abnormalities, obesity, and learning disabilities. Mutations in SDCCAG8 were also found in several BBS patients, making SDCCAG8 the 16th BBS gene (BBS16). However, little is known about the molecular functions of NPHP10 and how loss of NPHP10 function leads to the observed phenotypes. Our goal was to gain insight into the function of NPHP10 by determining its interactors. Here, we show that NPHP10 interacts with components of the multi-aminoacyl tRNA synthetase complex (MSC), including 8 out of 9 aminoacyl tRNA synthetases (ARS) as well as aminoacyl-tRNA-synthetase-complex interacting multifunctional protein 2 (AIMP2). We performed tandem affinity purification using stably transfected HEK293T cells expressing FLAG- and S-tagged NPHP10 and isolated its associated proteins. We further determined that among the MSC components, NPHP10 directly interacts with AIMP2. Sucrose gradient ultracentrifugation of WT mouse tissues showed that NPHP10 is associated with the MSC in the kidney, brain, and eye, all affected tissues in NPHP10 patients. Lastly, we show that disruption of the MSC through AIMP2 knockdown significantly reduces ciliogenesis. Altogether, these findings suggest that NPHP10 functions through an MSC-mediated mechanism or vice versa, and that the MSC is an important component for ciliary function.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2691F

Effect of the mutations Gly250Val and Leu451Val on the expression of the gene HEXA: disease causing or pseudo-mutations. *L. Peleg[1,3], R. Svirski[3], Y. Lerental[2], M. Karpati[1], A.J. Simon[2].* 1) Genetic Inst, Sheba Medical Ctr, Ramat Gan, 52621, Israel; 2) Cancer Research Center, Sheba Medical Ctr, Ramat Gan, 52621, Israel; 3) Human Molecular Genetics and Biochemistry, Sackler School of Medicine, Tel Aviv University,Tel Aviv, 69978, Israel.

Defects in the alpha-subunit of beta-hexosaminidase A (HexA) are responsible for the various forms of the autosomal recessive Tay-Sachs disease (TSD). The disease is relatively common among Ashkenazi and Moroccan Jews (carrier rate1:29 and 1:110 respectively) and an elevated frequency (1:140) was also found among Iraqi Jews (IJ) (pan-ethnic frequency - 1:280). Two mutations had been identified among the IJ: Gly250Val and Leu451Val (G749T and C1351G, respectively). A reduced HexA activity, in the TSD carrier's range, was found in all heterozygotes and they were identified in 75% of the IJ carriers. Yet, neither homozygotes nor compound heterozygotes with any disease causing mutation have been diagnosed so far. The purpose of this study was to examine the effect of the IJ mutations on the expression and function of the HEXA gene. Using site directed mutagenesis the IJ mutations and a known disease causing mutation (Arg137X) were introduced to pCMV6-vector containing the entire human HexA-alpha polypeptide. The plasmids were sequenced to ensure the presence of those mutations. The level of expression was examined in COS-7 cell cultures by enzyme activity and Western blot analyses. The enzymatic activity was measured with the fluorescence substrate 4-methylumbelliferyl-sulfo-N-acetyl-beta-glucosaminide (4MUGS) (specific to alpha subunit). In cultures baring the Arg137X mutation, as expected, no activity was found. In the case of Leu451Val and Gly250Val, 13% and 1% respectively, of the normal activity (no mutated insert) was measured. For western blot analysis pEGFP fused to alpha chains was used. Analysis of the bands showed a reduction of 95% and 41% of the alpha protein in the presence of Arg137X and Gly250Val mutations (respectively), while only about 12% reduction in the Leu451Val band (not significant). The results of this study might help genetic counseling for prenatal diagnosis of TSD cases. The pregnancies with compound heterozygotes fetuses were as yet terminated.

2692W

Identification and biochemical characterization of a novel mutation in DDX11 causing Warsaw Breakage Syndrome. *J. Capo-chichi[1], S. Bhati[2], J.A. Sommers[2], T. Yammine[3], E. Chouery[3], L. Patry[1], G.A. Rouleau[5], M.E. Samuels[1], F.F. Hamdan[1], J.L. Michaud[1], R.M Brosh Jr[2], A. Mégarbane[3,4], Z. Kibar[1,6].* 1) Centre de Recherche du CHU Sainte-Justine, Montréal, Québec, Canada; 2) 2Laboratory of Molecular Gerontology, National Institute on Aging, NIH, NIH Biomedical Research Center, 251 Bayview Blvd, Baltimore, MD 21224, USA; 3) Unité de Génétique Médicale et laboratoire associé Inserm UMR_S910, Université Saint Joseph, Beirut, Lebanon; 4) Institut Jérôme Lejeune, Paris, France; 5) Center of Excellence in Neuroscience of Université de Montréal, CHUM Notre-Dame Hospital Research Center and the Department of Medicine, Montréal, Canada; 6) Department of Obstetrics and Gynecology, Université de Montréal, Montréal, Canada.

Mutations in the gene encoding the iron-sulfur-containing DNA helicase DDX11 (ChlR1) were recently identified as a cause of a new recessive cohesinopathy Warsaw Breakage Syndrome (WABS), in a single patient with severe microcephaly, pre-and postnatal growth retardation, and abnormal skin pigmentation. Here, using homozygosity mapping in a Lebanese consanguineous family followed by exome sequencing, we identified a novel homozygous mutation (c.788G>A [p.R263Q]) in DDX11 in 3 affected siblings with severe intellectual disability (ID) and many of the congenital abnormalities reported in the WABS original case. Cultured lymphocytes from the patients showed increased drug-induced chromosomal breakage, as found in WABS. Biochemical studies of purified recombinant DDX11 indicated that the p.R263Q mutation impaired DDX11 helicase activity by perturbing its DNA binding and DNA-dependent ATP hydrolysis. Our findings thus confirm the involvement of DDX11 in WABS, expand its phenotypical spectrum, and provide novel insight into the structural requirement for DDX11 activity.

2693T

Development of a cell-based reporter assay for the analysis of regulatory interactions between FGF23/KLOTHO/FGFR1, small inhibitors, and downstream targets. *S. Diener[1], B. Lorenz-Depiereux[1], T.M. Strom[1,2].* 1) Institute of Human Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Bavaria, Germany; 2) Klinikum Rechts der Isar der Technischen Universität München, Institute of Human Genetics, Munich, Bavaria, Germany.

The analysis of rare genetic disorders affecting phosphate homeostasis led to the identification of several proteins essential for the renal regulation of phosphate homeostasis: PHEX (XLH [MIM 307800]), FGF23 (ADHR [MIM 193100]), SLC34A3 (HHRH [MIM 241530]), DMP1 (ARHR1 [MIM 241520]), ENPP1 (ARHR2 [MIM613312]), GALNT3 (FTC [MIM 211900]), and KLOTHO (FTC [MIM 211900]). A key regulator of phosphate homeostasis is the fibroblast growth factor 23 (FGF23). It is mainly secreted from osteocytes, circulates in the blood, and binds to receptor heterodimers composed of FGF receptor 1 (FGFR1) and KLOTHO in the kidney. FGF23 activates KLOTHO/FGFR1 to inhibit renal phosphate reabsorption and to suppress 1,25-dihydroxyvitamin D3 synthesis. As a key signaling pathway mitogen-activated protein kinase (MAPK) pathway is employed. To analyze regulatory interactions between FGF23/KLOTHO/FGFR1, small inhibitory compounds, and further downstream targets, we have developed FGF23-inducible HEK293 cells that stably express KLOTHO (HEK293-KL). Stable cell clones were picked, expanded, and expression of KLOTHO was confirmed by Western blot analysis. By investigating the activation of MAPK pathway we could show that HEK293-KL cells are FGF23-inducible. Moreover, we could inhibit the induction with FGF23 by the use of two small inhibitory molecules: (1) SU5402, an inhibitor of FGFR1, and (2) U0126, an inhibitor of MAPK pathway. Taken together, we have established a potent cell-based reporter assay, which can now be used to investigate FGF23/KLOTHO/FGFR1 receptor signaling and receptor complex inhibition in more detail. By RNA-Seq, which refers to high-throughput sequencing technologies, we will analyze transcriptome-wide changes in expression levels of genes in HEK293-KL cells after the induction with FGF23 in comparison to uninduced cells. Thereby, we will try to identify novel downstream targets which may be candidates for regulatory compounds involved in phosphate homeostasis.

2694F

The most activating calcium-sensing receptor mutation is located in the sixth transmembrane domain and can be corrected with chemical treatments. *AS. Lia-Baldini[1], A. Nizou[1], C. Magdelaine[2], F. Sturtz[1,2], B. Funalot[1,2], A. Liendhard[3].* 1) EA6309 - Maintenance Myélinique et Neuropathies Périphériques, Université de Limoges, Limoges, France; 2) Laboratoire de Biochimie, Centre Hospitalo-Universitaire de Limoges, Limoges, France; 3) Service de Pédiatrie, Centre Hospitalo-Universitaire de Limoges, Limoges, France.

The calcium-sensing receptor (CASR) has an important role in calcium homoeostasis by controlling PTH secretion and renal calcium handling. Gain-of-function mutations of the CASR gene have been identified in patients with sporadic or familial autosomal dominant hypocalcaemia (ADH), while inactivating mutations of the CASR gene cause familial hypocalciuric hypercalcaemia (FHH). We report here two cases of novel CASR mutations localized on the same amino-acid but leading either to ADH or to FHH. The first patient was an 11-year-old girl suffering from ADH, who developed nephrocalcinosis when she was 5 years old. The second patient was a 30-year-old lady presenting a mild hypercalcaemia. Genetics studies were performed by PCR amplification of CASR coding exons and direct sequencing of PCR products. Using the MAPK assay system and transient transfection of wild-type and mutated CASR into COS-7, we studied the ability of these mutated receptors to react to extracellular Ca2+ and to the negative allosteric CASR modulator, NPS2143. We found two heterozygous missense mutations localized in the sixth transmembrane domain of CASR. In vitro analysis of functional properties of the first mutated receptor to react to extracellular calcium did not reveal the classical expected leftward shift in the concentration-response curve for the mutant 1 compared to wild-type receptor, but a constitutive effect of this mutant. The mutated receptor exhibited almost 70% wild-type receptor activity even in the absence of extracellular calcium. We showed that this strong activity can be reduced with NPS2143 treatment. For mutant 2, a typical rightward shift compared to the wild-type receptor was observed. This finding, which has never been described to our knowledge, suggests the importance of this amino acid for the activation of the CASR protein and its critical physiological role in CASR signaling.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2695W

Dysregulated pathways in Cornelia de Lange syndrome cells mutated in core cohesin genes. *A. Musio[1], L. Mannini[1], A. Gimigliano[2], I. Bianchi[2], M. A. Deardorff[3], F. Cucco[1], I. D. Krantz[3], L. Bini[2].* 1) Istituto di Ricerca Genetica e Biomedica, CNR, Pisa, Italy; 2) Functional Proteomics Laboratory, Department of Biotechnologies, University of Siena, Siena, Italy; 3) Division of Genetics, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA.

Cohesin is an evolutionary conserved multiprotein complex responsible for correct chromosome segregation during the cell cycle. Beyond this structural function, increasing data argues non-canonical roles for cohesin, including gene expression regulation. Mutations in cohesin and regulatory cohesin genes have been identified in human disorders, collectively called cohesinopathies, with Cornelia de Lange syndrome being the most frequent. However, the etiopathogenetic mechanisms underlying Cornelia de Lange syndrome are still poorly understood. To define biochemical pathways that are affected in Cornelia de Lange syndrome we analyzed the proteomic profile of Cornelia de Lange syndrome cell lines carrying mutations in the core cohesin genes, SMC1A and SMC3. Dysregulated protein expression was found in Cornelia de Lange syndrome probands compared to controls. The proteomics analysis was able to discriminate among probands harboring mutations in the different domains of the SMC proteins. Differentially expressed protein spots were identified by mass spectrometry and were shown to be involved in important biological processes such as metabolism, cytoskeleton organization, antioxidant and detoxification, cellular and protein fate and RNA processing. The c-MYC gene represents a divergent hub and lies at the center of these pathways and is down-regulated in Cornelia de Lange syndrome cell lines. This study allowed us to highlight, for the first time, specific biochemical pathways that are affected in Cornelia de Lange syndrome providing plausible causal evidence for some of the phenotypic features seen in Cornelia de Lange syndrome. This work was supported by a grant from Tuscany Region to AM.

2696T

Conformations of Slipped-DNA Junctions Determine Correct or Mutagenic Repair of CAG/CTG Repeats. *M.M. Slean[1, 2], K. Reddy[1,2], B. Wu[3], K. Nichol Edamura[2], M. Kekis[1], F. Nelissen[3], R. Aspers[3], M. Tessari[3], O.D. Scharer[4], S.S. Wijmenga[3], C.E. Pearson[1,2].* 1) Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada; 2) Program of Genetics and Genome Biology, Hospital for Sick Children, Toronto, Ontario, Canada; 3) Department of Biophysical Chemistry, Radboud University Nijmegen, Institute of Molecules and Materials, Nijmegen, The Netherlands; 4) Program of Pharmacological Sciences and Chemistry, Stony Brook University, Stony Brook, New York, USA.

Slipped-strand DNAs, formed by out-of-register mispairing of repeat units on complementary strands, are transient intermediates in repeat length mutations causing at least 30 neurodegenerative diseases. Expansions of (CTG)•(CAG) repeated DNAs is the mutagenic cause of 14 neurological diseases, likely arising through the formation and processing of slipped-strand DNAs. The three-way slipped-DNA junction, at which the excess repeats slip-out from the duplex, is a poorly understood feature common to these mutagenic intermediates. Here, we reveal that slipped-junctions can assume various interconverting conformations where the strand opposite the slip-out is either fully-base paired, or has one- or two-unpaired nucleotides. Junction conformation can alter correct nick-directed repair levels. Surprisingly, certain junction conformations are aberrantly repaired to expansion mutations: Mis-direction of repair to the non-nicked strand opposite the slip-out leads to integration of the excess slipped-out repeats rather than their excision. Thus, slipped-junction structure can determine whether repair attempts lead to correction or expansion mutations.

2697F

High incidence of *SMAD3* mutations in thoracic aortic aneurysm and/or dissection patients. *H. Morisaki[1, 2], I. Yamanaka[1], A. Yoshida[1, 2], R. Sultana[1], H. Tanaka[3], Y. Iba[3], H. Sasaki[3], H. Matsuda[3], K. Minatoya[3], T. Kosho[4], N. Okamoto[5], A. Kawame[6], T. Morisaki[1, 2, 7].* 1) Dept Bioscience & Genetics, NCVC Res Inst, Suita, Osaka, Japan; 2) Dept Medical Genetics, NCVC, Suita, Osaka, Japan; 3) Dept Cardiovascular Surgery, NCVC, Suita, Osaka, Japan; 4) Dept Medical Genetics, Shinshu U Sch Med, Matsumoto, Nagano, Japan; 5) Dept Medical Genetics, Osaka Med Ctr & Res Inst Maternal & Child Health, Izumi, Osaka, Japan; 6) Nagano Childres's Hosp, Azumino, Nagano, Japan; 7) Dept Mol Pathophysiol, Osaka U Grad Sch Pharm Sci, Suita, Osaka, Japan.

SMAD3, a direct mediator of transcriptional activation by the TGFβ receptor, was identified to be the candidate responsible for aneurysm-osteoarthritis syndrome (AOS), new syndrome featuring thoracic aortic aneurysm and/or dissection (TAAD), arterial tortuosity, mild craniofacial features, skeletal and cutaneous anomalies, and early-onset osteoarthritis. Clinical features of AOS often overlap those of Loeys-Dietz syndrome caused by mutation in *TGFBR* genes, while osteoarthritis is a key feature observed in more than 90% of the reported cases. We screened*SMAD3* gene mutations in 194 Japanese index cases without mutation in *FBN1*, *TGFBR1*, *TGFBR2* and *ACTA2*, consisting of 88 cases suspected of connective tissue disorder with some distinct extravascular features (MFS/LDS-like) and 106 cases with young-onset or familial TAAD without these features (TAAD-only). Nine novel *SMAD3* mutations were identified in 4 cases (4.5%) in MFS/LDS-like and 5 cases (4.7%) in TAAD-only. All mutations were thought to be pathogenic since they were either nonsense-, frameshift-, splicing-mutations or missense-mutations in highly conserved residues with well characterized structural importance. Family history was positive in 4 index cases. Total of 12 patients including family members were examined by clinical geneticists, cardiologists and radiologists. Only 4 presented orthopedic problems, including 2 with a history of bone fractures and other 2 with excessive osteoporosis in the bone X-ray. Mild hypertelorism was observed in two but bifid uvula or cleft palate was not observed in any. Skeletal findings were not distinct except for one family, whose member presented mild scoliosis or dolichostenomelia. Thoracic aortic aneurysm/dissection was observed in all of adult cases, but Valsalva sinus was escaped in 2 cases. Arterial tortuosity was observed only in 2 elderly cases. Aneurysms in branch or intracranial arteries were present in 2 cases. Dural ectasia was observed in 4 cases. In summary, we found *SMAD3* mutations in about 5% not only in MFS/LDS-like group but also in TAAD-only group. The figure was much higher than the previous reports, though it may be due to ascertainment bias since our study group was recruited more broadly in sporadic cases. Because the clinical spectrum of patients with *SMAD3* mutations were very broad, we propose that screening of *SMAD3* gene is important in patients with young or familial TAAD even without early-onset joint abnormalities.

2698W

Clinical Exome Sequencing Leads to the Diagnosis of Timothy Syndrome in a Patient with Prolonged QT, Dysmorphic Features, Seizures, and Intellectual Disability. *W. Zeng[1], K.D. Gonzalez[1], X. Li[1], H.M. Lu[1], H. Lu[1], E.M. Miller[2], S. Ware[2], J.J Wei[1].* 1) Ambry Genetics, Aliso Viejo, CA; 2) The Heart Institute, Cincinnati Children's Hospital Medical Center, Cincinnati, OH.

Early results from clinical exome sequencing have demonstrated the clinical utility of the technology for diagnostic purposes. Exome sequencing was performed on a 2 year-old male with prolonged QT, hypotonia, joint hypermobility, dysmorphic features, cryptorchidism, seizures, developmental delay and bilateral clinodactyly. The family history was negative for similar phenotypes. Extensive molecular and biochemical testing was uninformative. The patient was heterozygous for a novel missense mutation (c. 3497T>C; p.I1166T) in exon 27 of the *CACNA1C*gene. Co-segregation analysis showed that his mother and father did not carry the mutation, indicating a *de novo* mutation occurrence. The amino acid is completely conserved throughout vertebrates. Mutations in *CACNA1C* lead to Timothy syndrome (TS), a rare arrhythmia disorder with fewer than 20 reported cases associated with congenital heart disease, syndactyly, immune deficiency, and autism. TS was listed among the differential diagnoses for the proband and previous sequencing for the common mutation(s) in TS was negative. The majority of reported mutations in the *CACNA1C* gene have been *de novo*. In addition, the proband's phenotype overlaps with those typically seen in Timothy Syndrome: a rate-corrected QT interval of between 480 ms and 700 ms, patent ductus arteriosos, facial anomalies including a flat nasal bridge and a round face, and neurological symptoms including seizures, intellectual disability, and hypotonia. Overall, diagnostic exome sequencing established a molecular diagnosis for a patient in whom traditional testing methods proved unhelpful. Additionally, the findings further expand the knowledge of the clinical spectrum of patients with mutations in the *CACNA1C* gene.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2699T

Mutations in calmodulin cause ventricular tachycardia and sudden cardiac death. *M. Nyegaard[1,7], M.T. Overgaard[2,7], M.T. Søndergaard[2], M. Vranas[1], E.R. Behr[3], L.L. Hildebrandt[2], J. Lund[2], P.L. Hedley[4], A.J. Camm[3], G. Wettrell[5], I. Fosdal[6], M. Christiansen[4], A.D. Børglum[1].* 1) Department of Biomedicine, The Bartholin Building, Aarhus University, DK-8000 Aarhus, Denmark; 2) Department of Biotechnology, Chemistry and Environmental Engineering, Aalborg University, DK-9000 Aalborg, Denmark; 3) The Division of Cardiac and Vascular Sciences, St George's University of London, SW17 0RE London, UK; 4) The Department of Clinical Biochemistry and Immunology, Statens Serum Institut, DK-2300 Copenhagen, Denmark; 5) The Division of Paediatric Cardiology, SE-221 85 Lund University Hospital, Sweden; 6) Department of Paediatrics, Visby Hospital, SE-621 55 Gotland, Sweden; 7) MN and MTO contributed equally.

Catecholaminergic polymorphic ventricular tachycardia (CPVT) is a devastating inherited disorder characterized by episodic syncope and/or sudden cardiac arrest during exercise or acute emotion in individuals without structural cardiac abnormalities. Although rare, CPVT is suspected to cause a substantial part of sudden cardiac deaths in young individuals. Mutations in the cardiac sarcoplasmic calcium channel *RYR2* gene have been identified as causative in approximately half of all dominantly inherited CPVT cases. Applying a genome-wide linkage analysis in a large Swedish family with a severe dominantly inherited form of CPVT-like arrhythmias, we mapped the disease locus to chromosome 14q31–32. Sequencing the *CALM1* gene encoding calmodulin, the quintessential intracellular calcium sensor which exhibits extraordinary evolutionary conservation, revealed a heterozygous missense mutation (Asn53Ile) segregating with the disease. A second, de novo, missense mutation (Asn97Ser) was identified in a CPVT patient of Iraqi origin by screening a collection of 61 arrhythmia samples negative for RYR2 mutations. Both mutations demonstrated compromised calcium binding and the p.Asn97Ser mutation displayed an aberrant interaction with the RYR2 calmodulin binding domain peptide at low calcium concentrations. We conclude that calmodulin mutations can cause severe cardiac arrhythmia. As a consequence of our findings, the calmodulin genes are candidates for genetic screening of patients with idiopathic ventricular tachycardia and unexplained sudden cardiac death.

2700F

Role of SP1-binding site polymorphism of COLIAI gene and a sequence variation 713-8 del C of TGFβ1 polymorphism in osteoporosis in thalassemia major patients. *S. Agarwal, K. Singh, R. Kumar.* Genetics, Sanjay Gandhi Inst Med Sci, Lucknow, India.

Background: Osteoporosis represents an important cause of morbidity in β-thalassemia major though its pathogenesis has not been completely elucidated. Genetic factors play an important role in the pathogenesis of osteoporosis and several candidate gene polymorphisms have been implicated in the regulation of this process. G→T polymorphism in the regulatory region of the collagen type I alpha 1 (COLIAI) gene and deletion of a C in the intron sequence eight base prior to exon 5(713-8 del C) has been targeted with osteoporosis. Aim: To find out the distribution of COLIAI Sp1(G-T) polymorphism and TGF-β1 (713-8delC) sequence variation and its relationship with bone mineral density (BMD) in thalassemia major patients of North India. Material and Methods: One hundred five beta thalassemia major patients attending IPD at Department of Genetics, SGPGIMS, Lucknow for blood transfusion were recruited for the study. TGF beta (713-8delC) gene and COLIA1 Sp1(G-T) polymorphism were analysed by PCR-restriction fragment length polymorphism (PCR-RFLP). Bone mineral density (BMD) was measured by Dual Energy X ray Densitometry (DEXA). Result : Z score of BMD of lumbar spine and hips were -2.39±1.09 and -2.37±0.86. The study indicated 19.8% of β -thalassemia patients as homozygous for wild type G/G (SS), 35.8% as heterozygous for G/T (Ss) and 43.4% as homozygous mutant T/T type (ss). A significant association of COLIA1 gene with Z-score of BMD at hip (p=0.047) and at lumbar spine (p=0.001) region was observed. Out of 105 studied cases, 57 patients had a one base deletion in intronic sequence 8 (713-8delC). However, no association of 713-8delC variant was found with Z score of BMD at lumbar Spine (P value=0.061) and Hips (P value=0.773). We observed Cc genotype of Tgfβ1 as a risk factor (odds ratio=3.3) in low bone density at lumbar spine. Conclusion: Our results raise the possibility that genotyping at the Sp1 site could be of clinical value in identifying chances of osteoporosis (low bone density) at hip and spine regions and patients with ss genotype had lower BMD as compared to SS genotype. While, TGF β sequence variation did not influence the BMD. The sequence variation, 713-8delC, in the TGF-β1 gene is more frequent in thalassemia patients with low bone density as compared to patients with normal bone density (odds ratio=3.3). Hence COLIA1 can be used as genetic marker for accessing osteoporosis in thalassemia patients.

2701W

Loss-of-function mutations in IGSF1 cause a novel X-linked syndrome of TSH deficiency and macroorchidism. *M. Breuning[1], Y. Sun[1], B. Bak[2], N. Schoenmaker[3], A.S.P. van Trotsenburg[4], P. Voshol[3], E. Cambridge[5], J. White[5], P. le Tisser[6,7], S. Mousavy Gharavy[7], J. Martinez-Barbera[7], W. Oostdijk[1], L. Persani[8,12], P. Beck-Peccoz[9,12], T. Davis[10], A. Hokken-Koelega[11], C. Ruivenkamp[1], J. Laros[1], M. Kriek[1], S. Kant[1], J. den Dunnen[1], R. Hennekam[4], K. Chatterjee[3], M. Dattani[7,13], J. Wit[1], D. Bernard[2].* 1) Leiden University Medical Center, Leiden, the Netherlands; 2) McGill University, Montréal, Canada; 3) University of Cambridge, Cambridge, UK; 4) Academic Medical Center, Amsterdam, the Netherlands; 5) The Wellcome Trust Sanger Institute, Hinxton, UK; 6) National Institute for Medical Research, London, UK; 7) UCL institute of Child Health, London, UK; 8) IRCCS intituto Auxologico Italiano, Milan, Italy; 9) Fondazione IRCCS Ca'Granda, MIlan, Italy; 10) The University of Western Australia, Australia; 11) ErasmusMC-Sophia Children's Hospital, Rotterdam, the Netherlands; 12) Università degli Studi di Milano, Milan, Italy; 13) Great Ormond Street Hospital for Children, London, UK.

Congenital central hypothyroidism occurs as component of multiple pituitary deficiencies or as isolated thyroid-stimulating hormone (TSH) deficiency. Deep X-exome sequencing in two male cousins with TSH deficiency, macroorchidism, and delayed testosterone production identified an inframe deletion in IGSF1. Contemporaneously, whole exome sequencing in two male siblings from a second family with congenital TSH deficiency revealed a nonsense mutation in the same gene. Analyses of nine additional families with TSH deficiency, macroorchidism and variable prolactin deficiency, uncovered a further six IGSF1 mutations (including nonsense, missense, frameshift) and two whole gene deletions. The phenotypes of the families highly resemble each other. IGSF1 encodes an integral membrane glycoprotein highly expressed in anterior pituitary gland. Disease-associated IGSF1 mutations impaired plasma membrane trafficking of the IGSF1 protein. Igsf1 deletion in male mice caused significant decreases in intra-pituitary TSH, circulating thyroxine, pituitary TRH receptor expression, and increased body weight and fat mass. These data suggest that loss of function mutations in X-linked IGSF1 cause lifelong TSH deficiency, macroorchidism, and variable prolactin deficiency.

2702T

Hyperphagia, leptin, and brain-derived neurotrophic factor in subjects with Alström syndrome and BMI-Z matched controls. *M.D. Hicks[1], J.D. Marshall[2], P. Maffei[3], A.E. Hanish[1,4], L.A. Hunter[1], S.M. Brady[1], N.M. Sedaka[1], R. Sherafat Kazemzadeh[1], J.W. Tsao[5], G. Milan[3], J. Naggert[2], J.A. Yanovski[1], J.C. Han[1].* 1) Section on Growth and Obesity, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD 20892 USA; 2) The Jackson Laboratory, Bar Harbor, ME 04609 USA; 3) Department of Medicine, Internal Medicine, University Hospital of Padua, Italy; 4) National Institute of Nursing Research, National Institutes of Health, Bethesda, MD 20892 USA; 5) Uniformed Services University of the Health Sciences, Bethesda, MD 20814.

Background: Alström syndrome (AS) is a rare autosomal recessive disorder caused by mutations in ALMS1. AS is a pleiotropic condition, with a wide range of clinical features, including early onset obesity. Recent studies of a mouse model of AS suggest that the Alms1 protein may play an important role in the basal bodies of primary cilia in hypothalamic neurons [1]. Leptin and its downstream mediator, brain-derived neurotrophic factor (BDNF), act within the hypothalamus to decrease appetite [2]. Studies assessing hyperphagia and the leptin pathway have not been previously reported in patients with AS compared with similarly obese, non-syndromic control subjects. Methods: Subjects with AS were matched 1:2 with healthy controls by age, sex, and body mass index Z-score (BMI-Z). Fasting serum leptin and BDNF for all subjects were measured by commercial ELISA. Parents of children ages 4–17 years also completed a hyperphagia questionnaire previously validated for Prader-Willi syndrome [3]. Skewed data were normalized by log transformation. ANCOVAs compared hyperphagia score, leptin, and BDNF (adjusted for age, sex, and BMI-Z). Results: Subjects with AS [n=35] and controls [n=70] were comparable for age (mean± SD for AS vs. controls: 17.5 ± 9.3 vs. 17.9 ± 9.9y, p=0.87), sex (% male: 40% vs. 43%, p=0.84), BMI (34.0 ± 10.7 vs. 31.9 ± 7.6 kg/m2, p=0.25), and BMI-Z (2.14 ± 0.60 vs. 2.07 ± 0.56, p=0.54). Subjects with AS compared with controls had significantly higher hyperphagia score (adjusted mean ± SEM: 24.3 ± 1.1 [n=17] vs. 18.1 ± 1.1 [n=14], p=0.04) and significantly higher leptin (39.5 ± 1.1 [n=32] vs. 23.5 ± 1.1 [n=59] ng/mL, p<0.001). BDNF was similar between groups (18.0 ± 1.2 [n=31] vs. 18.7 ± 0.9 [n=58] ng/mL, p= 0.66), but BDNF/leptin ratio was lower in AS compared with controls (0.5 ± 0.3 [n=30] vs. 1.3 ± 0.2 [n=58], p=0.03). Conclusions: Subjects with AS have greater hyperphagia scores and higher leptin concentrations than expected for their body size, suggesting a dysfunction in leptin-mediated appetite regulation. Negative feedback regulation by leptin on adipocyte-derived leptin production has been described [4]. Thus, the observations that subjects with AS have leptin that is elevated out of proportion with body size and have lower BDNF/leptin ratios are consistent with a potential role for ALMS1 in energy homeostasis within the hypothalamus and that a defect in the leptin signaling pathway may contribute to the pathophysiology of obesity in AS.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2703F

The Euro-WABB Registry: differences in molecular genetic confirmation between monogenic Wolfram, Alström, and Bardet-Biedl syndromes. *A. Farmer[1], S. Ayme[2], P. Maffei[3], S. Mccafferty[4], R. Sinnott[4], W. Mlynarski[5], V. Nunes[6], V. Paquis[7], K. Parkinson[8], V. Tillman[9], T. Barrett[10].* 1) Wellcome Trust Clinical Research Facility, Birmingham Children's Hospital, Birmingham, West Midlands, United Kingdom; 2) National Institue of Health and Medical Research (INSERM), INSERM SC11, 102 Rue Didot, Paris, France FR - 75014; 3) Università degli Studi di Padova, Dipartimento di Scienze Mediche e Chirurgiche, Via VIII Febbraio 2 - 35122, Padova, Italy, IT - 35122; 4) University of Glasgow, University Avenue, Glasgow, UK - G12 8QQ, UK; 5) Medical University of Lodz, 4 Kosciuszki Ave, Lodz, PL-90-419, Poland; 6) Fundació Institut Investigació Biomèdica de Bellvitge, Gran via Hospitalet, 199, Hospitalet de Llobregat (Barcelona), ES - 08907, Spain; 7) Centre National de la Recherche Scientifique, CNRS Délégation Côte d'Azur - 250 rue Albert Einstein (LBPG UMR 6267 - INSERM U998 - UNS Faculté de Médecine, 28 avenue de Valombrose 06107 Nice Cedex 2), Valbonne, FR - 06560, France; 8) 49 Southfield Avenue, Paignton, South Devon, TQ3 1LH, UK; 9) Department of Paediatrics, University of Tartu, Lunini 6, Tartu, EE - 501014, Estonia; 10) School of Clinical and Experimental Medicine, College of Medicine and Dentistry, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK.

Objectives. We aimed to develop a registry for the rare genetic diseases Wolfram (WS), Alstrom (AS), Bardet Biedl (BBS) and other diabetes syndromes, containing clinical, genetic diagnostic and outcome data. The purpose is to establish the natural history of these diseases; to assess clinical management; to characterize cohorts for future clinical trials; and to establish genotype phenotype relations. This abstract describes the first 50 patients recruited. Methods. Patients with a confirmed diagnosis (clinical or genetic) were recruited from both within and beyond Europe by their physicians. Information was collected for 42 `core' data fields, reached by consensus to differentiate between syndromes. We analysed prevalence of core clinical symptoms including obesity and diabetes. Results. The age range was 2–44 yrs (interquartile range 6–20yrs). There were 15 patients with WS (median age 18yrs (range 9–44yrs), 16 with AS (14 yrs (2–30yrs), 17 with BBS (8yrs (4–16), 1 with Wolcott-Rollison and 1 with vision and hearing impairment of unknown cause. The prevalences of molecular genetic confirmation were: WS 14/15; AS 14/16; BBS 12/17;p<0.01 for WS and AS combined vs BBS. The prevalences of diabetes and median ages of onset were: WS (14/15; 6yrs); AS (5/16; 13yrs); BBS (2/17; 10 yrs); p<0.01 for ages of onset WS vs AS and BBS combined). Conclusions. The core dataset captured sufficient data to differentiate between diabetes syndromes. Mutations in WFS1 and ALMS1 genes were identified in >95% of patients with WS and AS, but in BBS only 70% of patients had mutations in known BBS genes. The prevalence of diabetes is low in AS and BBS during childhood. Further patient recruitment and longitudinal data collection will use a consensus extended dataset of 400 fields to accurately characterize the phenotypes.

2704W

Biallelic *RFX6* inactivation can cause intestinal atresia with childhood rather than neonatal onset diabetes. *F.H. Sansbury[1, 2], R. Caswell[2], H. Lango Allen[2], S.E. Flanagan[2], B. Kirel[3], S. Ellard[2, 4], A.T. Hattersley[2, 4].* 1) Peninsula Clinical Genetics Service, Royal Devon & Exeter NHS Foundation Trust, Royal Devon & Exeter Hospital (Heavitree), Gladstone Road, Exeter, Devon EX1 2ED, UK; 2) Peninsula College of Medicine and Dentistry, Universities of Exeter and Plymouth, PMS Building, Royal Devon & Exeter Hospital (Wonford), Barrack Road, Exeter, Devon EX2 5DW, UK; 3) Eskisehir Osmangazi University, Faculty of Medicine, Department of Pediatrics, 26480 Eskisehir, Turkey; 4) Royal Devon & Exeter NHS Foundation Trust, Royal Devon & Exeter Hospital (Wonford), Barrack Road, Exeter, Devon EX2 5DW, UK.

Biallelic *RFX6* inactivation is a known cause of intestinal atresia, hepatobiliary abnormalities and NDM (neonatal diabetes mellitus). All the 7 cases to date had NDM, defined as diabetes diagnosed in the first 6 months. We report the first cases without NDM: two double first cousins who presented with intestinal atresia at birth but did not develop diabetes until ages 3 and 6 years.

The proband (born 32/40, 1.65 kg, 42nd centile) was born with duodenal atresia, jejunal web and Meckel's diverticulum. An identical twin with duodenal atresia died aged 1 month. Case 1 then had recurrent gastrointestinal bleeding due to erosive gastritis and duodenal ulceration. She presented aged 3 years with diabetic ketoacidosis. She requires 0.35 units / kg / day of long acting insulin, plus only occasional bolus insulin.

Her cousin (born 34/40, 1.7 kg, 7th centile) was born with duodenal atresia and mid gut malrotation. He has long standing anaemia, and was diagnosed with diabetes aged 6 years.

Both children tested negative for pancreatic auto-antibodies GAD and IA2. Genetic testing showed that both children are compound heterozygotes for the *RFX6* nonsense mutations c.2596C>T (p.Arg866X) and c.2176C>T (p.Arg726X).

This result significantly widens the diabetic phenotype of *RFX6* mutations. Sadly many patients with intestinal atresia die in infancy, before diabetes might present. Clinicians should consider biallelic *RFX6* inactivation as a potential cause of intestinal atresia.

2705T

Identification of a de novo mutation in the thyroid hormone receptor β gene (TRβ;) in a Colombian family with RTH. *M.C. Lozano[1], L. Mejia[2,3], P. Duran[4,5], C. Lattig[1].* 1) Universidad de Los Andes, Bogotá, Colombia; 2) Clínica Valle del Lili, Cali, Colombia; 3) Universidad Libre, Cali, Colombia; 4) Fundación Hospital Cardioinfantil-Instituto de Cardiologia, Bogota, Colombia; 5) Laboratorio de Investigacion Hormonal, Bogota, Colombia.

Resistance to thyroid hormone is an autosomal dominant disorder which affects 1 in every 40,000 births. This disorder is characterized by reduced soft tissue responsiveness to thyroid hormone, with increased levels of T4 and T3 and no inhibition of non-suppressed thyroid hormone (TSH), due to mutations present in the thyroid hormone receptor β gene (TRβ), particularly in its T3 binding domain. The clinical phenotype varies both between different families and between affected family members. Individuals with RTH can have variable resistance in different tissues, as a consequence of mixed features of hypo-and hyperthyroidism. The most common symptoms are goiter, tachycardia, attention deficit disorder, hearing impairment and delayed bone growth. In this study, we performed clinical and molecular characterization of a male patient, who was diagnosed with RTH at 15 months of age. He harbors a new mutation in exon 10 of the TRβ gene, which consists of a deletion of a cytosine at nucleotide 1318 and results in a frameshift that produces a stop codon at position 442. The mutation was not found in neither his parents nor his two healthy sisters, indicating a de novo mutational event. In Colombia, there is only one previously reported mutation associated with RTH (Salguero et. al., 2010), being the present study a second report. It is noteworthy that mutations found in Colombia are unique to our country, which should be taken into account for proper genetic counseling.

2706F

Use of lineage markers to predict founder effect of E180Splice mutation in GHR gene causing Laron syndrome: a witness of Jewish exodus. *F.T. Goncalves[1], C. Fridman[1], E.M. Pinto[10], Z. Laron[2], J. Guevara-Aguirre[3], F.G. Carsola[4], D. Damiani[5], T.S. Lins[6], O. Shevah[2], V. Hwa[7], A.L. Rosembloom[8], R.G. Rosenfeld[9], U.P. Arnhold[10], A.A.L. Jorge[10].* 1) Legal Medicine, Ethics and Occupational Health, Medical School, University of Sao Paulo, Sao Paulo, Brazil; 2) Schneider Childrens' Med Ctr, Petah-Tikva, Israel; 3) Inst for Endocrinology Metabolism and Reproduction, Quito, Ecuador; 4) University of Chile, Santiago, Chile; 5) Instituto da Criança, Medical School, University of Sao Paulo, Sao Paulo, Brazil; 6) Inst Materno Infantil de Pernambuco, Recife, Brazil; 7) Oregon Hlth and Scis University, Portland, OR; 8) University of Florida Coll of Med, Gainesville, FL; 9) Lucile Packard Foundation for Children's Hlth, Palo Alto, CA; 10) Department of Endocrinology, Medical School, University of Sao Paulo, Sao Paulo, Brazil.

Laron syndrome (LS) is a genetic disorder caused by mutations in the Growth Hormone Receptor (GHR) gene leading to severe GH insensitivity. The most prevalent GHR mutation is E180splice which was found only in isolated population of Spanish descent from the Andes (Ecuador); recently it was also found in Israel, Brazil and Chile. A common origin of the E180splice mutation is postulated and it is supported by the presence of common introns haplotype. Herein, to confirm this hypothesis, we analyzed polymorphic markers surround GHR gene (markers from chromosome 5 short arm, intragenic SNPs and GHR exon 3), paternal lineage origin analyzing Y-chromosome STR and maternal lineage investigating mitochondrial DNA haplogroups. We genotyped 20 LS patients from Ecuador, Israel, Brazil and Chile which were homozygous for E180splice mutation and their first degree relatives heterozygous for mutation and compared to 42 normal individuals and 5 LS patients with other mutations. An identical haplotype was identified in all but one of the patients carrying the E180splice mutation: D5S2082: 192/192; D5S665: 150/150; D5S2087: 246/246; rs6179 G/G rs6180 C/C, IVS9 haplotype-I and GHRfl. One patient differed from others by D5S2082 (168/192). Control individuals and LS patients with other mutation showed different haplotypes. This result corroborated that the E180splice-associated haplotype found in all families has not independent origins ($p < 10-8$) and they have a common ancestor. Half of patients showed paternal origin belonging to haplogroup R1b (found in Portugal and Spanish) and 46.6% belonged to haplogroups J and E (typical in the Middle East and in Eastern European Jews). mtDNA analysis showed that 29.4% of patients belonged to Amerindian haplogroup B, 17.6% had European haplogroup J (from Balkan Peninsula), 35.3% presented African haplotypes and the Israeli patient showed subhaplogroup R0 found in Arabia. Jews were expelled from Spain in 1492, but many had converted to the Christian faith. These New Christians emigrated from the Iberian Peninsula to the New World. The founder effect of E180Splice mutation may be explained by this historic event, which we reinforced with strong evidence of co-segregation between several polymorphisms, the germline E180splice mutation and paternal lineage in LS patients from different countries, indicating this mutation originated from a single common ancestor. Supported: FAPESP, LIM25-HCFMUSP, LIM40-HCFMUSP.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2707W

Genetics of congenital hypogonadotropic hypogonadism in Denmark. *J. Tommiska[1,2], P. Christiansen[3], N. Jørgensen[3], J. Lawaetz[3], A. Juul[3], T. Raivio[1,2].* 1) Institute of Biomedicine, University of Helsinki, Helsinki, Finland; 2) Children's Hospital, Helsinki University Central Hospital, Helsinki, Finland; 3) Department of Growth and Reproduction, Rigshospitalet, Faculty of Health Sciences. University of Copenhagen, Denmark.

Congenital hypogonadotropic hypogonadism (HH) is a rare disorder characterized by incomplete or absent puberty caused by the lack or deficient number of hypothalamic gonadotropin-releasing hormone (GnRH) neurons, disturbed secretion or action of GnRH, or both. When HH presents with deficient sense of smell (anosmia), the condition is termed Kallmann syndrome (KS). The phenotypic heterogeneity in patients with KS and normosmic HH (nHH) varies from genital hypoplasia and absent puberty to reversal of HH later in life. In the majority of cases, the molecular genetic cause remains unresolved. We examined the genetic background and phenotypic features of HH in Denmark. Thirty eight male patients with congenital HH were screened for mutations in *KAL1*, *FGFR1*, *FGF8*, *PROK2*, *PROKR2*, *GNRHR*, and *KISS1R*. Three KS patients had mutations in *KAL1* (C105VfsX12, C53X, ex5-8del). Five patients had mutations in*FGFR1* (V86M, R209C, R646W, A512V, and a synonymous change c.1614C>T, p.I538I, predicted to affect splicing); they all had severe HH (micropenis and cryptorchidism), and 3 of them also had cleft lip/palate, tooth agenesis, syndactylia, or hearing loss. One KS patient had a *PROK2* mutation (p.I55fsX1) previously identified in autosomal recessive KS, but in a heterozygous state. Among those with nHH, one proband had a homozygous R262Q mutation in *GNRHR*. He had recovered from HH before the age of 32, was without Rx for 30 yrs, and at the age of 60 yrs, presented with signs of late-onset hypogonadism. Three heterozygous R262Q mutations and one Q106R mutation were also found in *GNRHR*, but also in patients with KS, so it is not likely that these mutations contribute to the phenotype, although di/oligogenic inheritance has been suggested in HH. In conclusion, (i) the frequency of *KAL1* and *FGFR1* mutations in this patient series from the largest tertiary centre in Denmark was similar to that recently described in Finland; (ii) the phenotype of the patient with a homozygous *GNRHR* R262Q mutation displayed fascinating age-dependent variation; (iii) although the mutations in *TAC3*, *TACR3*, *KISS1*, *GNRH1*, and *CHD7* have not yet been screened, given the rarity of them in HH patients in all populations, it seems likely that the majority of the HH patients will remain without a molecular genetic diagnosis also in Denmark, implying the existence of hitherto undescribed genes underlying the condition.

2708T

A rare gain-of-function mutation in an inhibitory upstream open reading frame in CDKN1B 5' untranslated region causes MEN4 phenotype. *G. Occhi[1], D. Regazzo[1], G. Trivellin[1,2], F. Boaretto[3], D. Ciato[1], S. Ferasin[1], F. Cetani[4], E. Pardi[4], M. Korbonits[2], N. Pellegata[5], G. Opocher[3], F. Mantero[1], C. Scaroni[1].* 1) Department of Medicine, University of Padova, Padova, Padova, Italy; 2) Department of Endocrinology, Barts and the London School of Medicine, Queen Mary University of London, London, United Kingdom; 3) Veneto Institute of Oncology, Familial Cancer Clinic, Padova, Italy; 4) Department of Endocrinology and Metabolism, Section of Endocrinology and Bone Metabolism, University of Pisa, Pisa, Italy; 5) Institute of Pathology, Helmholtz Zentrum München German Research Center for Environmental Health, Neuherberg, Germany.

The CDKN1B gene encodes for the cyclin-dependent kinase inhibitor p27KIP1 that is a key player in many cell processes including proliferation and differentiation. Impaired p27KIP1 expression and/or localization is often observed in tumor cells, further confirming its central role in regulating the cell cycle. Recently, germline mutations in CDKN1B have been associated with the inherited multiple endocrine neoplasia syndrome type 4 (MEN4, MIM610755). MEN4 is a rare autosomal dominant syndrome characterized by varying combinations of tumors that affect at least two endocrine organs. Aim of present study was to determine the possible causes of multiple endocrine tumors in 24 MEN1-like patients by analyzing the entire CDKN1B gene for point mutations and large rearrangements in germline DNA. A c.-456_-453delCCTT was identified in a regulatory upstream ORF (uORF) in the 5'UTR, in a 62 year old acromegalic patient with a pancreatic endocrine neoplasm. The 4-bp deletion causes the shift of the uORF termination codon with the consequent lengthening of the uORF encoded peptide and the drastic shortening of the intercistronic space. Our functional studies clearly prove a negative influence of such deletion on the translation reinitiation at the CDKN1B starting site. LOH analysis from the pancreatic lesion revealed no loss of the second allele, confirming that p27KIP1 may act as haploinsufficient tumor suppressor. Importantly, our in-silico analyses showed a very high level of inter-species conservation of the uORF within the CDKN1B gene (both at nucleotide and protein level), strongly supporting a physiological role of the uORF-encoded peptide in modulating p27KIP1 translation. Although regulatory uORFs are very common in human and mouse genes, being observed in about half of cases, CDKN1B gene represents the fourth case for which an uORF affecting mutation have been clearly associated to human diseases. In addition, our findings demonstrate for the first time that, besides the already described degradation by the ubiquitin/proteasome pathway and non-covalent sequestration, p27KIP1 activity can also be modulated by an uORF, and that mutations in this region may have deleterious consequences.

2709F

Next generation sequencing to identify cryptic ABCC8 and HADH splice site mutations in patients with congenital hyperinsulinaemic hypoglycaemia. *S. Ellard[1,2], R. Caswell[2], M. Weedon[2], K. Hussain[3], S.E. Flanagan[2].* 1) Molec Gen, Royal Devon & Exeter Hosp, Exeter, United Kingdom; 2) Peninsula Medical School, University of Exeter, United Kingdom; 3) Institute of Child Health, University College London, United Kingdom.

Background Inactivating ABCC8 or KCNJ11 mutations are found by Sanger sequencing in 80% of patients with focal or diffuse forms of diazoxide unresponsive hyperinsulinaemic hypoglycaemia (HH). Some patients with a normal result have partial gene deletions. The mutational mechanism for focal HH suggests all patients will have a ABCC8 or KCNJ11 mutation. Sequence analysis of the exons, intron/exon boundaries, minimal promoter and alternative transcripts previously detected mutations in only 3/6 consanguineous probands with diazoxide-responsive HH who were homozygous across HADH. Objectives To use next generation sequencing (NGS) to search for non-coding ABCC8 mutations in 2 probands with focal HH and non-coding HADH mutations in the 3 consanguineous probands linked to HADH. Methods Long range PCR was used to amplify 117kb genomic DNA encompassing ABCC8/KCNJ11 or 94kb encompassing HADH. The products were sequenced on an Illumina GAIIx. Additional patients were tested by Sanger sequencing for novel mutations, including 23 diazoxide-unresponsive cases (for ABCC8) and 47 consanguineous diazoxide-responsive cases (for HADH). Results We identified a paternal ABCC8 intronic A>G variant in the 2 probands with focal HH and a homozygous HADH G>T mutation in the 3 consanguineous probands. Sanger sequencing detected the ABCC8 mutation in a further 3 and the HADH mutation in 5 additional consanguineous probands. In silico analysis predicts that both variants create cryptic splice donor sites which is likely to result in the incorporation of an out of frame pseudoexon. RNA studies are in progress. Conclusion We identified a novel intronic ABCC8 mutation in 5/25 (20%) probands with diazoxide unresponsive HH and a homozygous intronic HADH mutation in 8/50 (16%) consanguineous probands without a genetic diagnosis. This study highlights the importance of non-coding variants in the aetiology of HH and demonstrates the utility of next generation sequencing to identify these mutations.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2710W

ABCC8 and _KCNJ11_ gene mutational analysis in Slovenian patients with neonatal diabetes and congenital hyperinsulinism. *K. Trebusak Podkrajsek[1], M. Avbelj[1], T. Hovnik[1], N. Bratina[1], N. Bratanic[1], T. Battelino[1,2].* 1) University Children's Hospital, University Medical Centre Ljubljana, Ljubljana, Slovenia; 2) Medical Faculty, University of Ljubljana, Slovenia.

Sulfonylurea receptor 1 (SUR1) encoded by *ABCC8* gene, and inward rectifier KIR6.2 encoded by *KCNJ11* gene, constitute ATP-sensitive K+ channel in pancreatic β-cells responsible for regulation of glucose-induced insulin secretion. Loss of function and activating mutations in respective genes are associated with hyperinsulinism of infancy (HI) and neonatal diabetes (NDM) with/without associated neurological disorder, respectively. The possible implication of *ABCC8* and *KCNJ11* gene mutations in Slovenian patients with the two rare congenital disorders in insulin regulation was evaluated. Five patients with NDM and 3 patients with HI were screened for mutations in *ABCC8* and *KCNJ11* genes by direct sequencing of the whole reading frame and exon-intron boundaries. A novel homozygous *ABCC8* gene mutation p.Gly316Trp was identified in one patient with permanent NDM requiring insulin therapy since the age 2,5 months and otherwise normal development. At the age of 19 years he had well controlled diabetes using only glimepiride. In one of the HI patients a heterozygous *ABCC8* p.Ser1386Phe mutation was identified, inherited from the father diagnosed with type 2 diabetes. The patient had hyperinsulinemic hypoglycemia since neonatal period, developed transient diabetes with insulinopenia at the age of 25 years and suffers from epilepsy, mild to moderate mental retardation and psychiatric disorder marked with aggressiveness and agitation. No other mutations in either ABCC8 or KCNJ11 were detected in this small cohort of patients. The mutations in *ABCC8* and *KCNJ11* do not appear to be a frequent cause of congenital disorders in insulin regulation in Slovenian patients. A novel p.Gly316Trp mutation in *ABCC8* gene in patient with NDM was identified. While p.Ser1386Phe *ABCC8* mutant has been reported previously as an autosomal dominant cause of HI, incomplete penetrance of the disease in our case suggests the heterozygous mutation alone is not sufficient to result in a severe phenotype.

2711T

Severe familial growth hormone deficiency due to defective U12-dependent processing of pituitary developmental genes. *L.A. Perez-Jurado[1], R. Flores[1], G.A. Martos-Moreno[2], I. Cuscó[1], J. Argente[2].* 1) Unitat de Genetica, DCEX, Universitat Pompeu Fabra & Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Barcelona, Spain; 2) Department of Endocrinology, Hospital Infantil Universitario Niño Jesús, Department of Pediatrics, Universidad Autónoma de Madrid & Centro de Investigación Biomédica en Red Fisiopatología de la Obesidad y Nutrición (CIBEROBN), Madrid, Spain.

The minor spliceosome, a ribonucleoprotein complex of proteins and small nuclear RNAs, catalyzes U12-dependent RNA splicing of a small proportion (<1%) of human genes. Mutations in the U4atac snRNA component of the minor spliceosome have been recently found in the early lethal malformation syndrome microcephalic osteodysplastic primordial dwarfism type 1 (MOPD1). We have studied a family with three sisters affected with severe postnatal growth retardation (height around -6 SDS at diagnosis) and otherwise normal development. Growth hormone (GH) levels were undetectable after standard stimuli, while prolactin levels were in the low-normal range and the remaining pituitary hormones were normal. Brain MRI scans revealed hypoplasia of the anterior hypophysis in all three sisters and the therapeutic response to GH replacement has been excellent to date, suggesting a diagnosis of familial (autosomal recessive) isolated GH deficiency with associated pituitary hypoplasia. Mutational and segregation analyses discarded involvement of all known genes of the GH axis. Exome analysis revealed biallelic mutations in the *RNPC3* gene, coding for the 65 kDa protein of the minor splicesome that binds to the 3'-stem-loop of m(7)G-capped U12 snRNA and regulates pre-mRNA minor splicing. We selected candidate genes undergoing U12 processing for expression analysis in lymphoblastoid cell lines of affected individuals. Defective U12 transcripts with the alternative utilization of U2-dependent processing sites were observed in two genes *ARPC5L* and *MAPK8*, predicting the generation of shorter proteins in both cases. *ARPC5L* encodes the subunit 5-like of the actin related protein 2/3 complex involved in regulation of actin polymerization and formation of branched actin networks. *MAPK8* encodes a ubiquitously expressed serine/threonine-protein kinase involved in various developmental processes such as cell proliferation, differentiation, migration, transformation and programmed cell death. Interestingly, *ARPC5L* is differentially expressed in the developing pituitary, being an excellent candidate for the tissue-specific phenotype observed in this family. Our findings provide novel insights into the molecular basis of familial growth hormone deficiency and demonstrate a crucial role of the minor spliceosome component RNPC3 in pituitary development and growth regulation.

2712F

Impaired energy homeostasis and arrested maturation underlie neutrophil dysfunction in glycogen storage disease type Ib. *H.S. Jun[1], D.A. Weinstein[2], Y.M. Lee[1], B.C. Mansfield[1], J.Y. Chou[1].* 1) SCD, PDEGEN, NICHD, NIH, Bethesda, MD; 2) Dept Pediatrics, Glycogen Storage Disease Program, Division of Pediatric Endocrinology, University of Florida, Gainesville, FL.

Glycogen storage disease type Ib (GSD-Ib) is caused by a deficiency in a ubiquitously expressed glucose-6-phosphate (G6P) transporter (G6PT) that translocates G6P from the cytoplasm into the lumen of the endoplasmic reticulum. Inside the lumen, G6P is hydrolyzed by a glucose-6-phosphatase (G6Pase) to glucose and phosphate. A functional G6PT/G6Pase-α complex maintains interprandial blood glucose homeostasis while a functional G6PT/G6Pase-β complex maintains neutrophil homeostasis and function. In GSD-Ib patients the non-functional G6PT/G6Pase complexes results in impaired glucose homeostasis, neutropenia, and neutrophil dysfunction. The loss of glucose homeostasis in GSD-Ib is well understood and recent studies reveal that inactivation of either G6PT or G6Pase-β results in enhanced neutrophil apoptosis that leads to neutropenia seen in both disorders. However, the underlying cause of neutrophil dysfunction remains unknown. The bone marrow in some GSD-Ib patients exhibits neutrophil maturation arrest which may contribute to neutrophil dysfunction observed clinically. We have recently shown that impaired energy homeostasis underlies neutrophil dysfunction in a related disease, G6Pase-β deficiency. In this study, we show that the levels of immature blood in the majority (12 of 13) of GSD-Ib patients are significantly higher than those in normal subjects and neutrophils from all 13 patients are dysfunctional. Since G6PT and G6Pase-β are functionally coupled, we hypothesized that as with G6Pase-β deficiency, impaired neutrophil energy homeostasis might also underlie neutrophil dysfunction in GSD-Ib. Indeed, neutrophils from GSD-Ib patients have a lower glucose uptake, and lower levels of G6P, lactate, and ATP than neutrophils from normal subjects. In parallel, the expression and membrane translocation of NADPH oxidase subunit $p47^{phox}$ are down-regulated. Together, the results show that GSD-Ib neutrophils exhibit impaired energy homeostasis. Previous studies suggested that G-CSF failed to delay apoptosis of cultured GSD-Ib neutrophils. We now show that G-CSF improves the survival of cultured neutrophils from both GSD-Ib and control subjects although GSD-Ib neutrophils still exhibit enhanced apoptosis. The results establish that the G6PT-mediated G6P/glucose cycling is essential for neutrophil survival and energy homeostasis and a deficiency leads to neutropenia and neutrophil dysfunction.

2713W

A Novel Mutation in _PSTPIP1_ is Responsible For PAPA Syndrome in a Jordanian Child. *H. El-Shanti[1,3], F. Al-Amri[2], B. Fathalla[2], S. Al-Dosari[1], M. Al-Mutawa[1], M. Kambouris[1,4].* 1) Shafallah Med Gen Cen, Doha, Qatar; 2) Hamad Medical Corporation, Doha, Qatar; 3) University of Iowa, Pediatrics, Iowa City, IA, USA; 4) 3Yale University School of Medicine, Genetics, New Haven, CT, USA.

Pyogenic arthritis, pyoderma gangrenosum, and acne (PAPA) syndrome (OMIM #604410) is an autosomal dominant auto-inflammatory disease caused by mutations in the proline-serine-threonine phosphatase-interacting protein 1 *PSTPIP1*. The produced protein is a cytoskeleton-associated adaptor protein that modulates T-cell activation, cytoskeletal organization, and IL-1βrelease. The only two mutations described so far, A230T and E250Q, have been found in patients and families, but other variations were described without clear relation with the phenotype. The two mutations are thought to disrupt the binding of PSTPIP1 with PTP-PEST, a regulatory phosphatase, and increase its avidity for pyrin in the cytosol, thereby dysregulating IL-1βproduction. PAPA syndrome typically presents with recurrent sterile, erosive arthritis in childhood, occurring spontaneously or after minor trauma, occasionally resulting in significant joint destruction. By puberty, joint problems tend to subside and cutaneous symptoms increase. Cutaneous manifestations include pathergy, frequently with abscesses at the sites of injections, severe cystic acne, and recurrent nonhealing sterile ulcers, often diagnosed as pyoderma gangrenosum. We describe a 4 year old boy with recurrent sterile arthritis and history of skin abscesses, but no pyoderma gangrenosum. The sequencing of the coding region of *PSTPIP1* and flanking intronic regions revealed two variations, p.Gln219His (Q219H) and p.Asp246Asn (D246N). Both variants are predicted to be probably damaging by Polyphen, but as tolerated by SIFT. The first is present in the father who is completely asymptomatic, but the second is a de novo variation only present in the child. We anticipate that the mutation p.Asp246Asn is the cause of the symptoms in this child since it falls within the coiled coil domain that harbors all the previously described mutations, a fact that casts suspicion that the PTP-PEST interaction is mediated through this domain. Since the E250Q and A230T variants of PSTPIP1 were shown to severely abrogate binding to PTP-PEST in yeast two hybrid and co-immunoprecipitation experiments, we will assess its function experimentally using similar techniques.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2714T

Prevalence of ααα anti3.7 triplication in beta-thalassemia carriers in south of Iran. *G. Shariati[1,2], M. Hamid[3], A.H. Saberi[1,2], B. Sharififard[1], B. Keikhaei[4], M. Pedram[4].* 1) Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran; 2) Narges Medical Genetics & PND Laboratory, No. 18, East Mihan Ave, Kianpars, Ahvaz, Iran; 3) Molecular Medicine Dept., Biotechnology Research Center, Pasteur Institute of Iran, Tehran, Iran; 4) Research Center for Thalassemia and Hemoglobinopathy, Ahvaz Jundishapur University of Medical Sciences, Ahvaz-Iran.

There are no any studies investigated alpha globin gene triplication in beta-thalassemia carrier in Irania population, which was the aim of this study. The prevalence of alpha gene triplication in beta thalassemia carriers was studied in 2600 unrelated individuals in south of Iran, Ahvaz city. The association of heterozygous beta thalassemia with gene triplication was found in 30 patients (1.15%). The present result emphasizes that there is need to investigate systematically the ααα anti3.7 triplication in all hemoglobinopathies carriers in south of Iran.

2715F

Mutational analysis of Korean patients with Diamond-Blackfan anemia. *H. Chae[1,3], J. Park[1,3], M. Kim[1,3], A. Kwon[3], S. Lee[1,3], Y. Kim[1,3], J.W. Lee[2], N.G. Chung[2], B. Cho[2], D.C. Jeong[2].* 1) Department of Laboratory Medicine, The Catholic University of korea, Seoul, Korea; 2) Department of Pediatrics, The Catholic University of Korea, Seoul, Korea; 3) Catholic Genetic Laboratory Center, The Catholic University of Korea, Seroul, Korea.

Background Diamond-Blackfan anemia (DBA) is a congenital bone marrow failure syndrome characterized by hypoproliferative anemia, associated physical malformations, and a predisposition to cancer. DBA has been associated with mutations and deletions in the large and small ribosomal protein (RP) genes, and genetic aberrations are detected in approximately 50%–60% of patients. Methods Nine Korean DBA patients were screened for mutations in 8 known DBA genes, RPS19, RPS24, RPS17, RPS10, RPS26, RPL35A, RPL5, and RPL11, using the direct sequencing method. For mutation negative cases, array-CGH was performed to identify copy number variations. Genotype-phenotype correlations were evaluated. Results Mutations in RPS19, RPS26 and RPS17 were detected in 4, 2 and 1 patients respectively, and among RPS19 mutations 2 were novel (c.26T>A, c.357-2A>G). No mutations were detected in RPS24, RPS10, RPL35A, RPL5, and RPL11. No deletions involving the known DBA gene regions were identified on array-CGH. Genotype-phenotype correlation of RPS19 aberrations was not evident with respect to malformation status and treatment response. Conclusion Mutations in known DBA genes accounted for 78% of Korean DBA patients, and 44% were RPS19 mutations. We also report 2 novel RPS19 mutations. Our data suggests that Korean DBA patients are more susceptible to RP gene mutation than the Japanese or western DBA cohort.

2716W

Apoptosis is responsible for neutropenia in Cohen syndrome. *L. Duplomb[1,2], L. Faivre[1,2,3], A. Hammann[4], G. Jego[4], N. Droin[5], E. Lopez[1], B. Aral[1,2], N. Gigot[1,2], S. El Chehadeh[1,3], F. Huet[1,2], P. Callier[1,2], J. Thevenon[1,2], V. Carmignac[1], C. Capron[6], MA. Gougerot-Pocidalo[7], F. Girodon[2], JN. Bastie[2,4], L. Delva[4], C. Thauvin-Robinet[1,2,3].* 1) Génétique des Anomalies du Développement, Dijon, France; 2) CHU, Dijon, France; 3) Centre de Génétique et Centre de Référence “Anomalies du Développement et Syndromes Malformatifs”; 4) Université de Bourgogne, Inserm UMR 866, Dijon, France; 5) Institut Gustave Roussy, Inserm UMR 1009, Villejuif, France; 6) Institut Cochin, Inserm UMR 1016, Paris, France; 7) Univeristé Paris VII, Inserm UMR 773, Paris, France.

Cohen syndrome (CS) is a rare autosomal recessive disorder, characterized by typical facial appearance, microcephaly, truncal obesity, retinal dystrophy and intermittent neutropenia. The gene responsible for Cohen syndrome is VPS13B (COH1) which encodes a large protein of 4022 amino acids. Recently, VPS13B has been characterized as a peripheral membrane protein of the Golgi complex, with a crucial role for Golgi integrity. Neutropenia is one of major characteristics of CS but this neutropenia is usually intermittent and does not lead to severe infections. However, the exact role of the protein as well as the mechanisms involved in CS neutropenia are still unknown. The mechanisms responsible for neutropenia in other syndromic or non syndromic congenital neutropenia secondary to diverse genes (LEF-1, ELA2, HAX1, AK2,G6PC3...) include maturation arrest in neutrophil differentiation or enhanced apoptosis of mature neutrophils associated or not with endoplasmic reticulum stress. We first showed that CS neutrophils display normal generation of reactive oxygen species, chemotaxis, phagocytosis and morphology. However, when observed by transmission electron microscopy, the Golgi apparatus of neutrophil precursors from CS patients showed a disrupted morphology . In order to understand the mechanisms leading to neutropenia in CS patients, we have isolated CD16+ neutrophils from CS patients and healthy donors and showed by Annexin V-propidium iodide staining that CS neutrophils had an exaggerate apoptosis rate compared to healthy donors after 4h of culture. Expression of related endoplasmic reticulum stress-proteins, Bip and CHOP, were normal in CS neutrophils. For most of the patients, GM-CSF had no protector effect on CS neutrophils despite normal GM-CSF serum level, GM-CSF receptor mRNA expression and GM-CSF signaling (phospho-erk). We also showed that VPS13B mRNA expression dramatically decreased with the spontaneous apoptosis of neutrophils from healthy donors whereas mRNA expressions of 4 other genes (ELA2, HPN3, MPO and LYZ) were stable, confirming the correlation between VPS13B and apoptosis. Our work suggests that CS neutropenia is due to an increased apoptosis rate of circulating neutrophils and these results can be a clue for understanding other features of CS.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2717T

A novel pig model of SCID; A tool for cancer and immune research. *N.M. Ellinwood[1,2], E.H. Waide[1], D.M. Thekkoot[1], N. Boddicker[1], R.R.R. Rowland[3], E.M. Snella[1], J.K. Jens[1], C.R. Wyatt[3], J.W. Ross[1], D. Troyer[4], C.K. Tuggle[1], J.C.M. Dekkers[1].* 1) Dept Animal Sci, Iowa State Univ, Ames, IA; 2) Dept Vet Clin Sci, Iowa State Univ, Ames, IA; 3) Dept of Diag Med and Pathobio, Kansas State Univ; 4) Dept of Anat and Physio, Kansas State Univ.

Severe combined immunodeficiency (SCID) is an inherited trait identified in humans, horses and dogs, and artificially created in rodents. SCID mice have been used extensively to study immunity and cancer progression and treatment. We recently reported the first discovery of an autosomal recessive SCID in pigs based on the postmortem of 4 related piglets, which showed severely deficient immune systems, later characterized as being devoid of B and T cells. Objectives herein were to demonstrate the utility of this model in biomedical research, identify the genetic basis of the defect, and show that bone marrow transfer (BMT) was curative. Standard protocols were used for subcutaneous injection of human cancer cells into ears of SCID and normal pigs, and immunohistochemistry to detect human proteins. Genotyping was performed using an Illumina Porcine Beadchip. SNP association analysis was performed using the dfam option of PLINK. Homozygous affected piglets were injected intravenously with unfractionated bone marrow cells harvested from euthanized fully MHC-matched littermates at 16 days of age. Matings of known carriers produced 6 litters with confirmed SCID piglets at approximately 25% frequency. Affected piglets (n=3) were unable to prevent growth of human melanoma (A375SM) and pancreatic carcinoma (PANC-1) cells; human origin was verified by human mitochondrial protein staining of tumors. To identify the genomic region associated with this defect, 20 affected piglets, 50 unaffected littermates, the 6 parents, and 96 normal ancestors were genotyped. Association testing identified a 5.6 Mb region associated with the defect (P = 2.7×10- 7). Sequencing of the region is underway to identify the causative mutation. BMT was conducted on 3 SCID piglets. One developed graft versus host disease at 10 days post-transplant and was euthanized due to immunosuppressive therapy complications. The remaining two transplants have been clinically normal and free of any clincal infectious disease. We have documented engraftment in the recipients based on flow cytometry, and recipients are now 4 months of age. These novel findings are important as SCID pigs will be a valuable large animal model for basic and translational infectious disease, immune, and cancer research. The BMT treated SCID pigs will ensure efficient production of additional SCID animals.

2718F

Polymorphism in the HMOX-1 gene is associated with high levels of HbF in Brazilian patients with sickle cell anemia. *M.B. Melo[1], G.P. Gil[1], G. Ananina[1], M.B. Oliveira[1], M.J. Silva[1], M.N.N. Santos[2], M.A.C. Bezerra[3], B.L.D. Hatzlhofer[3], A.S. Araujo[3], F.F. Costa[4].* 1) CBMEG, Univ Campinas, Campinas SP, Brazil; 2) Department of Pathology, Univ Campinas, Campinas SP, Brazil; 3) Fundação Hemope, Recife, PE, Brazil; 4) Hemocentro, Univ Campinas, Campinas SP, Brazil.

Despite the fact that SCA is caused by a single mutation, the clinical course of these patients is extremely variable; the severity of the symptoms can vary from mild to very severe. The increase of fetal hemoglobin (HbF) has an important role in modulating the clinical status of sickle cell anemia (SCA), once it reduces the tendency of Hb polymerization inside the red cell, exerting an inhibitory role on the sickling process and reducing the severity of clinical manifestations. This phenomenon may result from rare mutations inside the β-globin gene cluster, from mutations in the promoter of the γ-globin gene or from some loci not linked to the β-globin locus. Studies have shown that oxidative stress is very elevated in patients with sickle cell anemia. The increased oxidative stress in SCA occurs due to the generation of free radicals during some events such as deoxygenation of red blood cells, ischemia-reperfusion and vaso-occlusive events. The goal of this study was to investigate the association of three polymorphisms involved in the oxidative stress pathway with the levels of HbF in patients with sickle cell anemia in a Brazilian population. One hundred and nine patients with sickle cell anemia were recruited for peripheral blood collection and for genomic DNA extraction. The epidemiological and clinical data including levels of HbF, sex and age were evaluated. Three polymorphisms, C242T and -930 A/G in the Cytochrome b (-245) Alpha Subunit (CYBA) gene and -413 A/T in the Heme Oxigenase-1 (HMOX1) gene were genotyped through direct sequencing. For statistical analysis the linear regression model from the transformed values of HbF was used. The HbF levels were not significantly associated with sex and age. When the levels of HbF were compared with the polymorphisms, the C242T and -930 A/G did not show significant association, but the TT genotype of the -413 A/T HMOX1 polymorphism showed statistically significant association with increased levels of HbF (p-value = 0.0307). We observed an association between the TT genotype of the -413 A/T polymorphism, present in the promoter of the HMOX1 gene with increased levels of HbF, which indicates the presence of a new marker related to HbF levels in SCA Brazilian patients.

2719W

Beta-Thalassemia in The Aegan Region Of Turkey: Identification of A Novel Deletion Mutation. *F. Ozkinay[1,2], H. Onay[2], E. Karaca[2], E. Arslan[2], A. Ece[2], B. Erturk[2], M. Tekin[2], O. Cogulu[1,2], Y. Aydinok[1], C. Vergin[3].* 1) Dept Pediatrics, Ege Univ, Izmir, Turkey; 2) Dept Medica Genetics, Ege Univ, Izmir, Turkey; 3) Dept Pediatrics, Dr. Behcet Uz Hospital, Izmir, Turkey.

OBJECTIVE: To investigate the characteristics of the couples counseled for beta thalassemia and to investigate molecular spectrum of beta thalassemia in the the Aegan Region of Turkey METHODS: Medical records of 438 beta thalassemia related consultations provided at a reference hospital in the Aegan Region of Turkey between 2004–2011 were investigated retrospectively. Between 2004–2011, total 1410 (1166 postnatal and 244 prenatal) molecular tests for beta thalassemia were performed at this center. There were 964 unrelated alleles among the group tested. Either a polymerase chain reaction (PCR)-based reverse dot-blot including 22 mutations or beta globin gene sequencing methods were performed to detect mutations. RESULTS: In 238 of 438 beta thalassemia related consultations, couples were identified to be carriers during premarital screenings. In 46, couples were found to be thalassemia carriers during pregnancy. In 85, couples had a previous child with thalassemia major. Majority of remaining cases were unmarried carriers. Two hundred forty four prenatal diagnosis 244 pnatal diagnosis were performed for thalassemia between 2004–2011 and 60 fetuses were found to be affected. Except 2 cases all couples chose to terminate the pregnancy. Molecular analysis of 964 unrelated chromosomes revealed 28 different genetic mutations in beta-thalassaemia gene among which 7 mutations accounted for 78% of the total beta-thalassaemia genes identified. The most common mutation was IVS1-110 with a frequency of 40% and followed by IVS-1-1 (8%), Codon 8 (7%), IVS-2.745 (7%), IVS2.1 (7%), IVS1.6 (5%), Codon 39 (4%) and Codon 6 (3%). A novel deletion mutation, c.206-212delTCGGTGC was detected in one of the carriers. This study provides important genetic information on beta thalassemia and is useful for genetic counseling in this region.

2720T

Development of a Diagnostic Panel for Genetic Immunodeficiencies using Next Generation Sequencing. *E.M. Coonrod[1], J. Durtschi[1], A. Kumanovics[1,2], R.L. Margraf[1], H.R. Hill[1,2,3,4], K.V. Voelkerding[1,2].* 1) R& D, ARUP Laboratories, Salt Lake City, UT; 2) Department of Pathology, University of Utah School of Medicine, Salt Lake City, UT; 3) Department of Medicine, University of Utah School of Medicine, Salt Lake City, UT; 4) Department of Pediatrics, University of Utah School of Medicine, Salt Lake City, UT.

Common Variable Immunodeficiency (CVID), X-linked and Autosomal Recessive Agammaglobulinemias, Hyper-IgM syndrome, and Severe Combined Immunodeficiency (SCID) are genetic immunodeficiencies that predispose to recurrent infections that can be severe and life threatening. Diagnosis can be costly and time consuming because of genetic and phenotypic heterogeneity of these disorders and because other conditions must first be excluded. There are currently 35 genes with variants known to cause these genetic disorders, however, in the case of CVID, only ~15% of cases are associated with a variant in one of the known genes. This suggests that causative variants in additional genes will likely be discovered. An improvement in molecular diagnostic testing for these disorders would incorporate lower costs with a faster turn-around time and the flexibility to add targets as additional genes are discovered. Here we compare custom capture methods for a 35 gene panel to an exome sequencing approach with the goal of developing an accurate and flexible molecular diagnostic test for genetic immunodeficiencies. Custom capture designs with Agilent's HaloPlex and Ion Torrent's AmpliSeq platforms were analyzed in silico. The custom capture designs varied in their theoretical capture of the genes of interest. For example, the TNFRSF13b (TACI) gene, which causes 8–12% of CVID cases, is predicted to be captured at 100% and 74.6% based on the Haloplex and Ion Torrent designs, respectively. Exome data from samples sequenced at our facility were used to determine the capture efficiency and resulting sequencing read coverage for the set of 35 genes along with an additional ~1,050 genes with known roles in immune system function that could be added to the panel in the future. Exomes were prepared using Nimblegen's SeqCap EZ Exome v3.0 kit and one exome was run per lane on the Illumina HiSeq 2000. Initial analyses show that 99.4% of coding sequences contain capture probes within 100 base pairs of the target region. The mean read coverage of the coding sequences in these data sets was 113 fold and 96% of the coding regions were covered by a median of 15 reads. Additional analyses will compare the in silico custom capture designs with our exome sequencing data to determine the most cost effective, comprehensive, and flexible approach to developing a molecular diagnostic test for genetic immunodeficiencies.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2721F

Spectrum of gene mutations in Chinese patients with refractory human herpes viruses associated disease. *H. Liu, F. Wang, Y. Zhang, Y. Wang, W. Teng, L. Guo, C. Tong, T. Wu, Y. Zhao, J. Gu, D. Lu.* Molecular Diagnosis Center, Beijing Daopei Hospital, Beijing, China.

We aimed to investigate the PRF1, UNC13D, STX11, STXBP2, SH2D1A and XIAP gene mutations in Chinese refractory herpes viruses (HHV) associated disease patients. Hemophagocytic Lymphohistocytosis (HLH), X-linked Lymphoproliferative Disease (XLP), chronic active EBV infection (CAEBV) and lymphoma patients were diagnosed according to international criteria. HHV1-8 was screened by PCR. Full length of coding region and splicing donor and receptor of all the 6 genes were sequenced and analyzed. We find that: 1) 43 patients infected with EBV, one of which only EBV positive in tonsils and lymph nodes biopsy specimens but negative in peripheral blood. 3 patients infected with HHV7 and 2 with CMV, 1 with HHV6, 1 with HSV1 and 1 with HSV2. 3 patients were simultaneously positive for 2 or 3 kinds of HHV viruses. 2) 10 cases carrying PRF1 mutations, the median age was 6.5(1–19) years. Totally 10 kinds of PRF1 mutation were detected. All of the 8 biallelic mutations carriers were all diagnosed as FHL2, and the other 2 monoallelic mutation carrier were both diagnosed as CAEBV. 3) 18 cases carrying UNC13D mutation, the median age was 6(0.8–14) years. There were 5 cases carrying biallelic mutations and 13 carrying monoallelic mutations. Totally 19 kinds of UNC13D mutation were detected. 7 cases were diagnosed as FHL3, 8 were diagnosed as CAEBV, 2 were diagnosed as Hodgkin lymphoma (HL). 4) Only one 10 years old female patient carrying a monoallelic STX11 F281 mutation and was diagnosed as EBV associated liver T-cell lymphoma. 5) 11 cases carrying STXBP2 mutations. The median age was 2(1–19) years. Totally 11 kinds of mutations were detected. The 3 biallelic mutations carriers were all diagnosed as FHL5, and the other 8 monoallelic mutations carrier were all diagnosed as CAEBV. 6) 5 male patients carrying SH2D1A mutations and diagnosed as XLP1. 7) 3 male patients carrying XIAP mutations and diagnosed as XLP2. By analysis the spectrum of disease and gene mutations in 47 Chinese patients, we reported that NK cell dysfunction mutations are important reason in Chinese refractory HHV associated disease, including FHL, XLP, CAEBV and lymphoma. The most common infected virus is EBV, and then HHV7, CMV. The most common mutated gene is UNC13D, and then STXBP2, PRF1.

2722W

Exome sequencing identifies missense IRF2BP2 mutation in a family with autosomal dominant common variable immunodeficiency (CVID). *L. Tian[1], D. Li[1], M. Keller[2], K. Cecilia[1], E. Frackelton[1], F. Otieno[1], H. Hakonarson[1,3,6], J.S. Orange[4,5,6].* 1) Center for Applied Genomics, The Children's Hospital of Philadelphia, Philadelphia, PA, USA; 2) Division of Immunology, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA; 3) Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA; 4) Center for Human Immunobiology, Texas Children's Hospital, Houston, Texas, USA; 5) Department of Pediatrics, Baylor College of Medicine, Houston, Texas, USA; 6) These authors contributed equally.

Common variable immunodeficiency (CVID) is a clinically and genetically heterogeneous group of disorders characterized by hypogammaglobulinemia, recurrent infections, and an inability to mount specific antibody responses to antigen. CVID can be multifactorial, or inherited in an autosomal dominant or autosomal recessive manner. Using whole-exome sequencing, we searched for the pathogenic mutation in a family with autosomal dominant CVID. The proband is a young woman who experienced recurrent sinopulmonary infections beginning in early adolescence. Her family history is notable for common variable immunodeficiency in both the proband's father as well as her older brother. Her brother was also diagnosed with type I diabetes. She has no history of autoimmune disease. Four family members were sequenced, of which three are affected (father and 2 children) and one is unaffected (mother). Genomic DNA was isolated from blood samples using standard methods. Whole-exomes were captured using the Agilent SureSelect Human All Exon 50Mb kit and subsequently sequenced with Illumina HiSeq2000 to an effective mean coverage of 80 fold. Paired-end reads (2×101 bp) were aligned to the reference human genome (UCSC hg19) using Burrows-Wheeler alignment (BWA). SNPs and indels were identified using the Genome Analysis Tool Kit. ANNOVAR and SnpEff were used to annotate the variants. We selected variants present in the affected and missing from the unaffected individuals. We narrowed our search by excluding variants that are homozygous in all patients since we are searching for a dominant mutation. The variants were filtered against the 1000 Genomes Project, and 5400 exomes from NHLBI GO Exome Sequencing Project. A missense mutation in IRF2BP2 was identified to be the most probable disease causing mutation. Sanger sequencing of DNA from all available individuals were performed and the IRF2BP2 mutation co-segregated with the disease phenotype. IRF2BP2 encodes an interferon regulatory factor-2 (IRF2) binding protein that interacts with the C-terminal transcriptional repression domain of IRF2. Detailed functional analysis of the IRF2BP2 mutation will be presented at the conference.

2723T

C-terminal frameshift truncation and extension mutation in POU3F4. *W. Park[1], B. Choi[2], D. Kim[1], A. Kim[2].* 1) Biomedical Sciences, Seoul National University College of Medicine, Seoul, South Korea; 2) Department of Otolaryngology, Seoul National University Bundang Hospital, Seongnam.

X-linked nonsyndromic hearing loss can be caused by various mutations of the POU domain, class 3, transcription factor 4 (POU3F4) gene. In Korean patients, we found five novel missense and frameshift truncation and extension mutations. Two missense mutations (p.Thr211Met and p.Gln229Arg) in the POU-specific domain could disturb transcriptional activity in in vitro assay and structure simulation. Two frameshift extension mutations (p.Thr354GlnfsX115 and p.X362ArgextX113) were located outside of POU domain and nuclear localization signal at the C-terminus. Those mutant POU3F4 proteins showed low expression levels, but could be rescued by MG132, a proteasome inhibitor in vitro. Mutant POU3F4 proteins were also localized to the cytoplasm and did not have transcriptional activity. Frameshift mutation (p.Leu317PhefsX12) in POU3F4 leads to the truncation of the C-terminal 44 amino acids including POU domain and nuclear localization signal. This frameshift truncation mutant protein was localized in nucleus and cytoplasm. This frameshift mutant was predicted to encode a protein that would not bind properly to DNA, which was confirmed by a luciferase assay even in presence of MG132. Our results suggested that frameshift truncation and extension mutations in the C-terminus of POU3F4 gene could lead to cytoplasmic localization and subsequent proteosomal degradation due to the structural aberration, which caused transcriptional inactivity, and thus nonsyndromic hearing loss.

2724F

Dominant allele of *DFNM1* suppresses *DFNB26* associated recessive hearing loss. *R. Yousaf[1,2], Z.M. Ahmed[1,3], G. Nayak[1], R.J. Morell[4], S. Sumanas[5], E.R. Wilcox[4], S. Riazuddin[6], T.B. Friedman[4], S. Riazuddin[1,2,3].* 1) Division of Pediatric Otolaryngology Head & Neck Surgery, Cincinnati Children's Hospital Research Foundation, and Department of Otolaryngology, College of Medicine, University of Cincinnati, Ohio 45229, USA; 2) National Center of Excellence in Molecular Biology, University of the Punjab, Lahore, 53700, Pakistan; 3) Division of Pediatric Ophthalmology, Cincinnati Children's Hospital Medical Center, and Department of Ophthalmology, College of Medicine, University of Cincinnati, Ohio 45229, USA; 4) Laboratory of Molecular Genetics, National Institute on Deafness and Other Communication Disorders, Rockville, Maryland 20850, USA; 5) Division of Developmental Biology, Cincinnati Children's Hospital Research Foundation, and Department of Pediatrics, University of Cincinnati, Ohio 45229, USA; 6) Allama Iqbal Medical Research Center, Allama Iqbal Medical College, Lahore, 53700, Pakistan.

We previously reported mapping of nonsyndromc deafness locus *DFNB26*, in a large consanguineous family with eight affected individuals. Interestingly, seven individuals in this family, despite being homozygous for the *DFNB26* linked haplotype, had normal hearing, thus were non-penetrant for the *DFNB26* mutation (Riazuddin et al, 2000, Nature Genetics, 26(4), 431–434). We will discuss in detail the identification of *DFNB26* and *DFNM1* genes at the ASHG 2012 Annual Meeting. The *DFNB26* missense mutation significantly impaired phosphoinositide binding ability and plasma membrane localization of the mutant protein. We used zebrafish embryos to study the effect of *dfnb26* gene knockdown, followed by rescue experiments with wild type and mutant human *DFNB26* and *DFNM1* mRNAs. In the mouse inner ear, *Dfnb26* and *Dfnm1* encoded proteins are localized in the neurons, which innervate both inner and outer hair cells. In lymphoblastoid cells of individuals homozygous for *DFNB26* mutation, several genes of the HGF signaling pathway are dysregulated. Since, the *DFNB26* gene is involved in a wide variety of signaling pathways, deciphering the mechanism of suppression of the hearing loss phenotype by a dominant modifier allele of *DFNM1* may have future clinical relevance beyond addressing *DFNB26* hearing loss.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2725W

Phenome-wide, cross-species comparisons identify novel disease genes with high accuracy. *D.P. Smedley[1], C.J. Mungall[2], S.C. Doelken[3,4], S. Köehler[3,5], B.J. Ruef[6], M. Westerfield[6], P.N. Robinson[3,4,5], S.E. Lewis[2], W.C. Skarnes[1], members of the Mouse Genetics Project.* 1) Wellcome Trust Sanger Institute, Cambridge, Cambridgeshire, United Kingdom; 2) Genomics Division, Lawrence Berkeley National Laboratory, Berkeley, CA; 3) Institute for Medical and Human Genetics, Charité-Universitätsmedizin, Berlin, Germany; 4) Max Planck Institute for Molecular Genetics, Berlin, Germany; 5) Berlin Center for Regenerative Therapies (BCRT), Charité-Universitätsmedizin, Berlin, Germany; 6) ZFIN University of Oregon, Eugene, OR.

Despite numerous recent successes in identifying causative mutations for genetic disorders, an associated gene has yet to be identified for many diseases and for the others, rarer molecular associations remain to be discovered. Systematic sequencing of genes in linkage regions or exome sequencing approaches are routinely applied to the problem but in both cases 10–100's of candidate genes have to be analysed. Here we describe how semantic phenotypic comparisons of human disease and model organism phenotypic data can be used for prioritization and validation of these disease sequencing efforts. Despite the incomplete coverage of mouse and zebrafish phenotype data, up to 73% of known disease genes can be rediscovered and at stringent levels where on average single candidates are identified for each disease locus, 70% of predictions are correct and 26% of known genes recalled. To uncover novel disease genes, we applied our approach to diseases with a mapped locus but no known molecular association as reported by OMIM and identified candidates for 72 and 22 records using mouse and zebrafish data respectively. We then pursued a validation strategy of reviewing the literature for recently published associations, contacting researchers in each disease area to establish if they have tested our predictions and pursuing collaborative sequencing studies. We have now collected evidence for almost half the predictions based on mouse phenotypic similarities and genomic location. 25% have been confirmed as associated mutations have now been detected in the predicted gene with another 31% having very strong evidence for association. 28% have been screened without success for coding variation but the possibility of non-coding or structural rearrangements remains and, finally, the remaining 14% have been excluded as a mutation in another gene has been found. In conclusion, we find that cross-species phenotypic comparisons are an invaluable tool for the identification of both disease genes and existing animal models for biological characterisation of disease mechanism. This approach will become increasingly powerful as high quality phenotypic annotations of mouse and zebrafish models are generated for all protein coding genes through the efforts of the International Mouse Phenotyping Consortium (www.mousephenotype.org) and Zebrafish Mutation Project (http://www.sanger.ac.uk/Projects/D_rerio/zmp). Our latest predictions are available at http://www.sanger.ac.uk/mouseportal.

2726T

Mutation Screening in *CYP1B1* in Glaucoma Cases from North India. *A. Bhardwaj, V. Vanita.* Department of Human Genetics, Guru Nanak Dev University, Amritsar, Punjab, India.

Purpose: We aimed to analyze the disease causing mutations in *CYP1B1* in glaucoma patients from North India. Material and Methods: The present study included 142 glaucoma patients recruited from Dr. Daljit Singh Eye Hospital, Amritsar. Out of these, 109 cases were with primary glaucoma, of which 60 were of primary open angle glaucoma (POAG), 37 primary angle closure glaucoma (PACG) and 12 primary congenital glaucoma (PCG) and rest 33 cases had secondary glaucoma. Detailed ophthalmic evaluation in these cases included visual acuity, tonometry, gonioscopy, slit lamp examination, fundus examination, optical coherence tomography (OCT) and pachymetry. Measurements regarding intraocular pressure, corneal thickness, cup:disk (C/D) ratio, and refractory error were also recorded. Mutation screening in the candidate gene *CYP1B1* was performed by bi-directional sequence analysis of the amplified products. Results: A previously known mutation i.e. p.Arg368His in exon 3 of *CYP1B1* was identified in patients with PACG. In addition, three previously reported polymorphisms (p.Val432Leu, p.Asp449Asp, p.Asn453Ser) were also detected in cases with juvenile open angle glaucoma, primary congenital glaucoma and in angle closure glaucoma phenotypes. Conclusions: We report p.Arg368His as recurrent mutation in *CYP1B1* in our analyzed PACG patients. Three identified polymorphisms in present study have also been previously reported for identical phenotypes. These findings suggest a strong phenotype-genotype correlation of different types of glaucoma with identified substitutions in *CYP1B1*.

2727F

Screening of SOX10 and MITF regulatory regions in Waardenburg syndrome type 2. *N. Bondurand[1, 2], V. Baral[1, 2], A. Chaoui[1, 2], Y. Watanabe[1, 2], M. Goossens[1, 2, 3], T. Attie-Bitach[4], S. Marlin[5], V. Pingault[1, 2, 3].* 1) Dept Genetics, Inserm U955, IMRB, Creteil, France; 2) Université Paris Est, Faculté de Médecine, Creteil, France; 3) AP-HP, Hôpital H. Mondor-A. Chenevier, Service de biochimie et génétique, Créteil, France; 4) INSERM U781, Université Paris Descartes,, Hôpital Necker-Enfants Malades, Paris, France; 5) Service de Génétique, Centre de référence (Surdités génétiques), INSERM U587, Hôpital Armand Trousseau, France.

Waardenburg syndrome (WS) is a rare auditory-pigmentary disorder that exhibits varying combinations of sensorineural hearing loss and pigmentation defects. Four subtypes are clinically defined based on the presence or absence of additional symptoms. WS type II (WS2) can result from mutations within the MITF or SOX10 genes; however, 70% of WS2 cases remain unexplained at the molecular level, suggesting that other genes might be involved and/or that mutations within the known genes escaped previous screenings. The recent identification of a deletion encompassing three of the SOX10 regulatory elements in a patient presenting with another WS subtype, WS4, defined by its association with Hirschsprung's disease, led us to search for deletions and point mutations within the MITF and SOX10 regulatory elements in 28 yet unexplained WS2 cases. Two nucleotide variations were identified: one in close proximity to the MITF distal enhancer (MDE) and one within the U1 SOX10 enhancer. Despite the fact that each of these variations could potentially create or alter transcription factor binding sites, the results of in vitro functional analysis as well as familial segregation analysis argued in favour of a non-causative effect of the variations identified. Altogether, our results still leave about 70% of WS2 cases unexplained at the molecular level, suggesting that mutations within regulatory elements of WS genes are not a major cause of this neurocristopathy.

2728W

Utilization of gene mapping and candidate gene mutation screening for diagnosing clinically equivocal conditions. *V. Chini[1], D. Stambouli[2], F. M. Nedelea[2], G. A. Filipescu[3], D. Mina[4], H. El-Shanti[1, 4], M. Kambouris[1, 5].* 1) Molecular Genetics, Shafallah Medical Genetics Center, Doha, Qatar; 2) Cytogenomic Medical Laboratory, Bucharest, Romania; 3) Elias University Emergency Hospital, Bucharest, Romania; 4) University of Iowa, Pediatrics, Iowa City, IA, USA; 5) Yale University, School of Medicine, Genetics, New Haven CT, USA.

A pregnant female requested prenatal diagnosis for a congenital and complex eye disease segregating in her family. The three-generation pedigree of Romanian ethnic origin was suggestive of an X-linked inheritance, due to exclusively affected males and no father to son transmission. Affected individuals had bilateral optic nerve atrophy, microphthalmia, nystagmus, leukocoria, cataract, retinal detachment, eye tumors reported as retinoblastomas, moderate mental and motor retardation, and seizures. All efforts to obtain the detailed medical records of affected individuals were fruitless. The disease locus was mapped utilizing 78 microsatellite markers that span the X-chromosome at ≈2 Mb intervals. Affected individuals share a ≈10Mb region between DXS1056 and GATA160B08 at Xp11.23–11.4. Candidate genes in this linkage interval included BCOR and NDP. Mutation screening identified a c.C267A p.F89L mutation in the NPD gene in all affected individuals, previously described in a single unrelated Dutch family and speculated as causing Norrie disease (Nikopoulos et al, Human Mutation 31:656–666, 2010) In retrospect clinical symptomatology fits the Norrie disease phenotype. The reported retinoblastomas were most likely pseudogliomas characteristic of Norrie disease. Detection of the c.C267A p.F89L mutation in a second unrelated family confirms the pathogenic nature of the mutation for Norrie disease. Utilization of gene mapping through linkage analyses and candidate gene screening previously utilized exclusively for research applications, were applied at a diagnostic setting and were essential in deciphering the offending molecular defect and diagnosing a disease which due to lack of medical records and poor and misleading clinical history would have no chance of being diagnosed correctly. Clinical diagnosis and mutation identification were essential prerequisites for providing proper genetic counseling and prenatal diagnosis in this family.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2729T

Search for a new gene responsible for non-syndromic craniosynostosis. *S. El Chehadeh-Djebbar[1], F. Di Rocco[2], C. Collet[3], M. Le Merrer[3], J. Puechberty[4], G. Lefort[4], B. Aral[5], S. Lambert[6], N. Gigot[5], M. Payet[6], C. Ragon[6], N. Marle[6], A-L. Mosca-Boidron[6], J-R. Teyssier[5], F. Mugneret[6], C. Thauvin-Robinet[1], P. Sarda[4], L. Faivre[1], P. Callier[6].* 1) Centre de Génétique et Centre de Référence Anomalies du Développement et Syndromes Malformatifs, Hôpital d'Enfants, CHU de Dijon, France; 2) Centre de référence des dysostoses craniofaciales, Hopital Necker, Paris, France; 3) Centre de référence des maladies osseuses constitutionnelles, Hopital Necker, Paris, France; 4) Laboratoire de Cytogénétique, CHU Arnaud de Villeneuve, Montpellier, France; 5) Laboratoire de Génétique Moléculaire, Plateau Technique de Biologie, CHU de Dijon, France; 6) Laboratoire de Cytogénétique, Plateau Technique de Biologie, CHU de Dijon, France.

Background: Craniosynostoses are rare heterogeneous condition characterized by premature fusion of one or more of the cranial sutures and comprise two main groups, non-syndromic craniosynostoses, the most frequent (85%), and syndromic craniosynostoses (15%). In certain cases, the latter can be associated with a microdeletion on the short arm of chromosome 7 (CRS1 region, located in 7p21.1), which is responsible for Saethre-Chotzen syndrome. A chromosome 7p14.3–15.1 CRS2 region has been suspected in a small number of patients with craniosynostosis and who had no anomaly in 7p21.1. Objectives: To confirm the existence of the CRS2 locus, and to determine the implication of the PLEKHA8 gene in the etiopathogenesis of craniosynostoses. Method: This study was conducted in two steps. The first comprised a clinical and molecular cytogenetic characterization of two patients with syndromic craniosynostosis and a 7p deletion containing the CRS2 region, as well as two relatives with non-syndromic craniosynostosis who carried a familial balanced translocation involving the 7p15.1 region. The translocation breakpoints were cloned. The second step consisted in sequencing the candidate gene interrupted by the translocation in a cohort of 37 patients with familial (14) and non-familial (23) non-syndromic scaphocephaly, negative for TWIST, FGFR1, FGFR2 and FGFR3 sequencing. Results: We confirmed and further defined the existence of a CRS2 3.6 Mb locus in 7p. The PLEKHA8 (OMIM 608639) gene was interrupted by the 7p translocation breakpoints. PLEKHA8 is involved in Trans Golgi Network to plasma membrane transport and in the formation of post-Golgi constitutive carriers. Interestingly, the PLEKHM1 gene, from the same family, is implicated in osteopetrosis, with a putative critical function in vesicular transport in the osteoclast. However, direct sequencing of this gene in a cohort of 37 patients with non-syndromic craniosynostosis revealed no pathogenic variants. Conclusions: The results of this study, associated with data in the literature, confirm the existence of a second candidate region located in 7p and called CRS2 in etiopathogenesis of craniosynostoses. However, we did not confirm the involvement of the PLEKHA8 gene.

2730F

Mutation screening in *CNGA1* in an Autosomal Recessive Retinitis Pigmentosa (ARRP) family. *S. Goyal, V. Vanita.* Department of Human Genetics, Guru Nanak Dev University, Amritsar, Punjab, India.

Purpose: The aim of present study was to perform mutation screening in gene for Cyclic nucleotide gated channel alpha 1 (*CNGA1*) in an ARRP family of Indian origin. Material and Methods: We came across a large four generation ARRP family at the Dr. Daljit Singh Eye Hospital, Amritsar, with five affected individuals and parents of the proband being first cousins. Age of onset of the disease in this family was from 16 to 22 years. Ophthalmic examinations that included visual acuity, fundus testing and fundus photography after pupil dilation confirmed the status of the five individuals being affected by non-syndromic typical RP. The *CNGA1* that encodes the alpha subunit of the rod photoreceptor cyclic nucleotide gated channel (CNG) and mediates visual transduction in retinal rods and which is linked with ARRP, was selected for mutation screening. Bi-directional sequence analysis of amplified products of all the exons (III, IV, V, VI and VII, IX, XA, XB, XC, XD, XE) and exon-intron boundaries of *CNGA1* was performed in this family. Results: No disease linked mutation was detected in *CNGA1* in the analyzed ARRP family. Conclusions: No any disease linked mutation could be detected in *CNGA1* for non-syndromic typical RP in the tested autosomal recessive family. Present findings suggest the role of some other gene for retinitis pigmentosa in present family.

2731W

A novel mutation of the Nicastrin (NCSTN) gene in a large Indian family with autosomal dominant Hidradenitis suppurativa (or Acne inversa). *D. Jhala[1], U. Ratnamala[2], T.Y Mehta[3], F.M.S Al-Ali[4], M. Raveendrababu[1], S.S Chettiar[1], M.V Rao[1], S. Nair[5], J.J Sheth[6], M.R Kuracha[7], S.K Nath[8], U. Radhakrishna[9,10].* 1) Department of Zoology, School of Sciences, Gujarat University, Ahmedabad, Gujarat, India; 2) Department of Pharmacology, Creighton University, Omaha, NE, USA; 3) Samarpan Medical & Research Organization on Skin, Modasa, India; 4) Dermatology Department, Rashid Hospital, Department of Health and Medical Services (DOHMS), Dubai, United Arab Emirates; 5) Department of Fetal Medicine, Lifeline Genetics and Research Centre, Lifeline Hospital, Adoor, Kerala, India; 6) Foundation for Research in Genetics and Endocrinology, Ahmedabad, India; 7) Department of Surgery, Creighton University, Omaha, NE, USA; 8) Arthritis and Immunology Research Program, Oklahoma Medical Research Foundation, Oklahoma City, US; 9) Department of Surgery-Transplant, Nebraska Medical Center Omaha, NE, USA; 10) Green Cross Blood Bank, Ahmedabad, India.

Hidradenitis Suppurativa (HS), also known as `Acne Inversa' is a chronic-inflammatory and scarring disease that usually presents after puberty. It is characterized by swollen or inflamed lesions on the groin, anogenital parts, buttocks, thighs, scalp, neck, nipples and armpits, with subsequent scarring and chronic seepage. The prevalence of HS appears to be 1% of the general population and is 4–5 times greater in women than in men. Genetic predisposition plays a key role in HS and considerable progress has been made recently in identifying genes that are responsible for HS predisposition. Recently, mutations of NCSTN, PSEN1 and PSENEN genes on chromosomal regions 1q23.3, 19q13.12 and 14q24.2, respectively, have been found in sporadic and familial HS patients, suggesting that these genes play an important role in HS development. Here, we report a novel mutation of NCSTN in a large seven-generation Indian family (N = 211) with an isolated non-syndromic HS in which 31 members affected by HS with different phenotypic expression including typical characteristics of HS. The other findings were four women with hirsutism and two deaths due to squamous cell carcinoma, in this family. Mutation search in exons and splice junctions of above three candidate genes identified a frameshift mutation, c.687insCC (p.Cys230ProfsX31), which is predicted to terminate at codon 261 after the addition of 31 novel amino acids within exon 6 of NCSTN. It is known that NCSTN is subunit of the gamma-secretase complex, an endoprotease complex that catalyzes the intramembrane cleavage of integral membrane proteins such as Notch receptors and APP (beta-amyloid precursor protein). We did not detect a pathogenic or predisposing variant in any of the three candidate genes in another multigenerational HS family of Arab origin analyzed, demonstrating genetic heterogeneity of HS. Identification of present mutation expands NCSTN gene mutation spectrum associated with HS phenotype.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2732T

A Systematic Approach to Identifying Disease Genes in Mullerian Aplasia. *HG. Kim[1], HK. Nishimoto[1], R. Ullmann[2], JA. Phillips III[3], Y. Shen[4], RH. Reindollar[5], HH. Ropers[2], VM. Kalscheuer[2], LC. Layman[1].* 1) Section of Reproductive Endocrinology, Infertility, & Genetics, Dept OB/GYN, Georgia Health Sciences University, Augusta, GA; 2) Department of Human Molecular Genetics, Max Planck Institute for Molecular Genetics, Ihnestrasse 63–73, 14195 Berlin, Germany; 3) Dept of Pediatrics, Vanderbilt University, Nashville, TN 37212, USA; 4) Dept of Neurology, Children's Hospital, Harvard Medical School, Waltham, MA 02453, USA; 5) Dept Ob/Gyn, Dartmouth College, Lebanon, NH 03756, USA;.

Developmental abnormalities of the uterus and vagina occur in ~7–10% of women and result in a wide range of reproductive dysfunction. Mullerian aplasia (MA), also known as congenital absence of the uterus and vagina (CAUV) and Mayer-Rokitansky-Kuster-Hauser (MRKH) syndrome, represents the most drastic developmental aberration of the Mullerian system since reproduction is essentially arrested. MA occurs in 10% of women with primary amenorrhea, but associated anomalies (renal agenesis, skeletal abnormalities, cardiac anomalies, and deafness) may also be present. The molecular basis for most MA remains unknown as WNT4 point mutations have only been found in a subset of MA patients. A critical barrier to progress in the identification of new MA genes is the lack of realistic candidate genes and familial cases. Molecular mapping of chromosome translocation breakpoints in isolated patients offers an opportunity to identify novel disease genes that are disrupted or dysregulated by the chromosome rearrangement. Even a single balanced chromosome translocation has the potential to lead to the localization and cloning of the disease-causing gene through molecular analysis of the breakpoints. Small Copy Number Variation (CNV) also may pinpoint genomic regions containing only a couple of positional candidate genes. Here we applied positional cloning through both breakpoint mapping in a balanced translocation and array CGH for the identification of small CNVs in MA patients to discover new causative genes in MA. Additionally, we adopted whole exome sequencing in trios of MA families to identify de novo mutations in new MA genes. Our cohort of MA patients with chromosomal abnormalities includes: 1) a de novo balanced translocation 46,XX,t(3;16)(p21;p13.3)dn; 2) a 1.3 Mb deletion at 17q12; and 3) 9 trios (MA patient and healthy parents), as well as additional MA patients. Utilizing these three parallel techniques, we took advantage of the opportunity to locate disease genes of Mullerian aplasia by using our unique cohort of MA patients. We have identified four candidate genes (HNF1B, CTNNA3, KIAA1680, and CHSY3), all of which have intragenic duplications or deletions. Our long-term goal is to identify new disease genes and the biochemical pathways important in human Mullerian development.

2733F

Linkage analysis in an Autosomal Recessive Retinitis Pigmentosa (ARRP) Family. *Vanita. Kumar, S. Saini.* Department of Human Genetics, Guru Nanak Dev University, Amritsar, Punjab, India.

Purpose: The aim of present study was to perform linkage analysis in an ARRP family of Indian origin to identify the disease linked gene. Material and Methods: We came across an ARRP family at the Dr. Daljit Singh Eye Hospital, Amritsar, with three affected siblings and their parents being first degree cousins. Ophthalmic examinations that included visual acuity, fundus testing and fundus photography after pupil dilation confirmed three sibs to be affected with non-syndromic retinitis pigmentosa. Linkage analysis was carried out using over 110 highly polymorphic fluorescently labeled microsatellite markers for already known candidate gene regions for autosomal recessive as well as autosomal dominant RP. Genotyping was carried out on a DNA sequencer ABI-377, data collected and analyzed by GENESCAN version 3.1.2, and genotyping was done using GENOTYPER 2.5.1 software. 2-point and multipoint linkage analysis were performed by means of the program packages linkage v.5.2 and Genehunter v.2.1. Results: Linkage analysis was carried out with nine genomic DNA samples from three affected and six unaffected members including parents and four of their unaffected sibs. Two-point lod score values of less than -2.0 at theta = 0.001 were obtained with the analyzed markers at already known candidate gene loci thus excluding these to be linked with ARRP in this family. Multipoint lod score analysis further confirmed the 2-point lod score data. Conclusions: We excluded all the candidate gene loci known to be linked with autosomal recessive as well as autosomal dominant RP in an ARRP family of Indian origin having three members affected with non-syndromic RP. These findings indicate the role of a novel gene for ARRP in present analyzed family. Further, these findings highlight the inherent genetic heterogeneity for non-syndromic RP.

2734W

Homozygosity Mapping and Mutation Profile of the *MYO7A* Gene in Saudi patients with Usher Syndrome. *K. Ramzan[1], M. Al-Owain[2], K. Taibah[3], G. Bin-Khamis[4], R. Allam[1], A. Al-Mostafa[1], S. Al-Hazza[5], F. Imtiaz[1].* 1) Department of Genetics, King Faisal Specialist Hospital & Research Centre, Riyadh, Riyadh, Saudi Arabia; 2) Department of Medical Genetics, KFSHRC, Riyadh, Saudi Arabia; 3) ENT Medical Centre, KFSHRC, Riyadh, Saudi Arabia; 4) Department of Otolaryngology, KFSHRC, Riyadh, Saudi Arabia; 5) Department of Ophthalmology, KFSHRC, Riyadh, Saudi Arabia.

Usher syndrome deafness-blindness disorder, is a clinically and genetically heterogeneous disorder. Three clinical subtypes of Usher syndrome have been identified, with mutations in different set of genes. Usher syndrome type 1 (USH1) is the most severe form of Usher syndrome. The *MYO7A* gene is responsible for USH1B, the most common subtype. Homozygosity mapping is a robust approach that is highly suited for genetically heterogeneous autosomal recessive disorders in populations where consanguinity is prevalent. For genotyping, DNA samples were processed on Gene Chip Human Mapping 250K Arrays (Affymetrix, Santa Clara, CA) following the instructions provided by the manufacturer. Homozygosity mapping was carried out using CNAG software that uses SNP genotypes generated by a DM algorithm for the detection of copy number changes and blocks of homozygosity. For affected individuals where the homozygosity blocks were identified on chromosome 11q13.5 harboring *MYO7A*, direct sequencing was performed for the entire coding and flanking intronic regions. Because of the high frequency of consanguinity in our patients this approach allowed us to quickly identify eleven *MYO7A* mutations in twenty families with Usher Syndrome. Though many different mutations in this gene have been identified worldwide, the purpose of this study was to document the *MYO7A* variants present in the Saudi population.

2735T

Mutation Analysis in an Autosomal Recessive Retinitis Pigmentosa Family. *S. Saini[1], J.R. Singh[2], V. Vanita[1].* 1) Department of Human Genetics, Guru Nanak Dev University, Amritsar, Punjab, India; 2) Central University of Punjab, Bathinda, India.

Purpose: To do mutation screening in an autosomal recessive retinitis pigmentosa (arRP) family of Indian origin in candidate genes; cyclic nucleotide gated channel alpha-1 (*CNGA1*), elongation of very long fatty acid 4 (*ELOVL4*), rhodopsin (*RHO*) and retinal degeneration slow (*RDS*)/*Peripherin*. Methods: A detailed family history and clinical data including visual acuity test, fundus examination and optical coherence tomography were recorded. Mutation screening in four candidate genes known for autosomal recessive RP (*CNGA1*, *ELOVL4*, *RHO* and*RDS/Peripherin*) was performed by bi-directional sequence analysis of the amplified products. Results: Mutation screening in the four candidate genes in an arRP family revealed a novel c.205G>T change in heterozygous form in *RDS/Peripherin* resulting in p.Val69Leu substitution in the proband and his unaffected mother. This substitution was absent in proband's unaffected father, his other unaffected siblings and 50 ethnically matched controls. Amino acid conservation analysis revealed valine to be highly conserved at codon 69 of *RDS/Peripherin*. Three already known polymorphisms (p.Gln304Glu, p.Arg310Lys and p.Asp338Lys) were also detected in *RDS/Peripherin* in the present family in affected as well as unaffected members. Further, mutation analysis in *RHO* indicated a novel polymorphism i.e. c.163A>T in heterozygous form leading to p.Asn55Tyr in the proband and his unaffected parents while his unaffected sister had both wild type alleles. Conclusions: We identified previously unreported substitution i.e. p.Val69Leu in *RDS/Peripherin* in an arRP family. Mutations in RDS/Peripherin have not yet been reported with arRP although it has been linked with autosomal recessive central areolar choroidal dystrophy type 2 (CACD2. There is a possibility that in present analyzed family the disease is either being carried in autosomal recessive pattern with pseudodominance or p.Val69Leu substitution in *RDS/Peripherin* in association with polymorphism (p.Asn55Tyr) in RHO, both in heterozygous state, may lead to RP in the proband. Another possibility is that another mutation in compound heterozygous state in some other gene segregating with the mutation p.Val69Leu in *RDS/Peripherin* may lead to RP in present family.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2736F

The role of APCS and RBP4 genes as modifiers of age-at-onset in Familial Amyloid Polyneuropathy (FAP ATTRV30M). *D. Santos[1], T. Coelho[2], J.L. Neto[1,3], J.P. Basto[1,3], J. Sequeiros[1,3], I. Alonso[1,3], C. Lemos[1,3], A. Sousa[1,3].* 1) UnIGENe, IBMC, Porto, Portugal; 2) Unidade Clínica de Paramiloidose, Centro Hospitalar do Porto (CHP), Porto, Portugal; 3) ICBAS, Instituto Ciências Biomédicas Abel Salazar, Univ. Porto, Portugal.

Familial amyloid polyneuropathy (FAP ATTRV30M) is an autosomal dominant inherited disease, due to a point mutation in the TTR gene (chr18q11.2-q12.1). Remarkable differences in mean age-at-onset (AO) have been described in different clusters, including within Portuguese population. Among Portuguese families, FAP shows a wide variation in AO (17–82 yrs) and asymptomatic carriers aged 95 can be found; this variation is also often observed between generations. A previous study in Portuguese patients (Soares et al., 2005) found a modifier effect in AO for APCS gene and RBP4 gene, when comparing classic and late-onset patients with controls. However variation between generations was not taken into account. Our aim was to investigate if these two candidate-genes have a modifier effect in AO variation from parent to offspring in FAPATTRV30M families. We collected a sample of 36 FAP families with at least 2 generations affected. We selected 5 tagging SNPs through the degree of linkage disequilibrium (LD) existing between SNPs and also the 5 SNPs previously described. These SNPs were analysed by SNaPshot and RFLP, respectively. Samples' genotyping is currently underway and results are being analyzed with the GeneMapper™ v.4.0 software. Preliminary results in 5 FAP families showed that although for RBP4 gene we found different genotype's frequencies in patients for rs7079946 and rs17484721 from HapMap, no striking differences were found between generations in the families analyzed for the two genes. In the total sample of Portuguese families ascertained, we expect to find or exclude the potential role of these candidate-genes as modifiers of FAP ATTRV30M AO. The study of genetic modifiers is crucial to understand the mechanisms involved in AO variability within and between families and may have an important impact in genetic counselling.

2737W

Linkage Analysis in Families with Autosomal Dominant Corneal Dystrophies. *D. Singh[1], V. Vanita[2].* 1) Dr. Daljit Singh Eye Hospital, Sheranwala Gate, Amritsar, Punjab, India; 2) Department of Human Genetics, Guru Nanak Dev University, Amritsar, Punjab, India.

Purpose: To perform linkage analysis in two autosomal dominant corneal dystrophy families of Indian origin so as to identify the disease linked gene. Material and Methods: We came across two autosomal dominant corneal dystrophy families at the Dr. Daljit Singh Eye Hospital, Amritsar. In one family in 4-generations 17 members were found to be affected with Reis Bucklers type and in the second family 15 members in 3-generations had granular type corneal dystrophy. Linkage analysis was carried out using highly polymorphic fluorescently labeled microsatellite markers for already known candidate gene regions for corneal dystrophies. Genotyping was carried out on a DNA sequencer ABI-377, data collected and analyzed by GENESCAN version 3.1.2, and genotyping was done using GENOTYPER 2.5.1 software. 2-point and multipoint linkage analysis were performed by using the program packages linkage v.5.2 and Genehunter v.2.1. Mutation screening in the potential candidate gene at the mapped interval was carried out by bi-directional sequence analysis of the amplified products. Results: In two-point and multipoint lod score analysis we obtained significant positive lod scores indicative of linkage with markers at chromosome 5q31 in these two autosomal dominant corneal dystrophy families. Mutation screening in all the exonic regions of the candidate gene i.e. transforming growth factor beta-induced gene TGFBI) at the mapped interval revealed segregation of Arg124Cys mutation in all the affected members of family with Reis Bucklers corneal dystrophy and in none of the unaffected members of this family. Second family with granular type corneal dystrophy showed the segregation of Arg555Trp change in all the tested affected members and in none of the unaffected individuals. Conclusions: Present findings depict the co-segregation of two mutations Arg124Cys and Arg555Trp in gene for TGFBI in two large families of Indian origin with Reis Bucklers and granular type corneal dystrophy respectively. These findings substantiate the role of TGFBI for corneal dystrophy. Further, present findings suggest a strong phenotype-genotype correlation of Reis Bucklers and granular type corneal dystrophy with these two identified mutations.

2738T

A novel chromosomal locus in a family with primary ciliary dyskinesia. *R. Parvari[1,2], M. Mazor[1], S. Alkrinawi[3], V.C. Sheffield[4], M. Aviram[3].* 1) Dept of Virology& Dev. Genetics, Ben Gurion Univ of the Negev, Beer Sheva, Israel; 2) National Institute of Biotechnology in the Negev, Ben Gurion University of the Negev, Beer Sheva, Israel; 3) Division of Pediatrics, Soroka University Medical Center, Beer-Sheva, Israel; 4) Department of Pediatrics - Division of Medical Genetics, Howard Hughes Medical Institute, University of Iowa, Iowa City, IA.

Motile cilia are essential for fertilization, respiratory clearance, cerebrospinal fluid circulation and establishing laterality. Cilia motility defects cause primary ciliary dyskinesia (PCD, MIM244400), a disorder affecting 1:15,000–30,000 births. In a consanguineous Bedouin family with 2 affected and 6 healthy siblings we have searched for the mutation causing the disease. Assuming disease by homozygosity of a mutation from a common ancestor, linkage to the DNAH5 and DNAI1 genes was excluded by the finding of heterozygosity of alleles at known polymorphic markers adjacent to these genes. We further searched for homozygous regions consistent with linkage by genotyping the patients and both parents with the Affymetrix (Santa Clara, CA) GeneChip® Human Mapping 250K Sty arrays. Five homozygous regions larger than 18cM were identified using the KinSNP program. To determine linkage to a single region, all family members were tested with both known polymorphic microsatellite markers and with additional markers developed for this purpose. Linkage was identified to a chromosomal locus on chromosome 18p11.1-q13.3. The multipoint Lod score calculated was 3.0, using the Pedtool server assuming recessive inheritance with 99% penetrance and an incidence of 0.01 or 0.001 for the disease allele in the population. The region contains 10 genes encoding ciliary proteins, a putative mutation causing PCD was not identified in any of these.

2739F

Polymorphism of VEGF +674 gene as potential risk factor for age related macular degeneration (AMD) in Indian patients. *D. Gupta[1], V. Gupta[2], S. Chawla[3], S. R. Phadke[1].*1) Department of Medical Genetics, Sanjay Gandhi Post Graduate Institute of Medical Sciences, Lucknow, Uttar Pradesh, India; 2) Department of Physiology, Chhatrapati Sahuji Maharaj Medical University Chowk, Lucknow Uttar Pradesh, India; 3) Prakash Netra Kendra, Gomti Nagar Lucknow, Uttar Pradesh, India.

Age related macular degeneration (AMD) is a degenerative disease of the retina that causes progressive impairment of central vision and is the leading cause of irreversible vision loss in elderly people. Vascular endothelial growth factor (VEGF)+674 gene (rs1413711) found on chromosome 6p12 has been determined to be strongly associated with AMD and a major player in the control of angiogenesis. We studied the VEGF +674 (rs1413711) polymorphism in VEGF gene as potential risk factor for AMD in Indian patients. Hundred patients and controls were genotyped for the C-to-T polymorphism in Intron1, 1695 bp downstream from the start of Exon1 of the VEGF gene by Amplification refractory mutation system polymerase chain reaction (ARMS PCR). Genotyping analysis of VEGF gene of 100 AMD patients shows that, 27% homozygous (CC); 19% mutant (TT); and 54% samples had heterozygous (CT) genotype. Whereas in Control samples 45% (CC); 14% (TT); and 41% (TC). The odds ratio of VEGF (rs1413711) gene for development of AMD in carriers of mutant allele (T) vs (C) wild allele is 2.59, homozygosity for the genotype (TT) increased likelihood for AMD 2.5 fold (p-value 0.036 with 95%; CI). +674CC genotype is a potential risk factor for AMD and that carriage of the T allele is potentially protective. While heterozygosity for genotype of VEGF +674 gene CT increased likelihood 2.0 fold which indicate that significant contribution of VEGF gene C>T polymorphisms. Identification of such would not only increase our understanding of the biology of the disease but might also allow a prediction of those `at risk' in the general population. KEY WORDS: Indian, Age related macular degeneration (AMD), Age related maculopathy (ARM), Vascular Endothelial Growth Factor (VEGF), Amplification refractory mutation system polymerase Chain Reaction (ARMS PCR).

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2740W

Whole genome analysis identifies mutations in inositol polyphosphate phosphatase-like 1 (INPPL1) as the cause of opsismodysplasia. *J.E. Below[1], D. Earl[2,3], K.M. Bofferding[3], M.J. McMillin[3], J. Smith[1], E. Turner[1], M. Stephan[3], L.I. Al-Gazali[4], J. Hertecant[4], D.H. Cohn[5], D.A. Krakow[6], D.A. Nickerson[1], M.J. Bamshad[3].* 1) Genome Sciences, Univ Washington, Seattle, WA; 2) Medical Genetics, Seattle Children's Hospital, Seattle, WA; 3) Department of Pediatrics, Univ Washington, Seattle, WA; 4) Department of Pediatrics, United Arab Emirates Univ, Al Ain, UAE; 5) Department of Molecular, Cell and Developmental Biology, Univ of California Los Angeles, Los Angeles, CA; 6) Department of Human Genetics, Univ of California Los Angeles, Los Angeles, CA.

Opsismodysplasia is a rare chondrodysplasia characterized by micromelia, platyspondyly, delayed skeletal ossification, metaphyseal cupping, and in some individuals severe phosphate wasting. Based on the occurrence of multiple affected siblings born to consanguineous parents, opsismodysplasia is hypothesized to segregate in an autosomal recessive pattern. We performed whole genome sequencing on a parent-child trio of a consanguineous family with two siblings with opsismodysplasia and phosphate wasting. Using linkage analysis we identified regions of homozygosity in the proband. The whole genome sequence data revealed homozygous mutations in several high priority candidate genes within these regions, only one of which, INPPL1, is involved in phosphate metabolism. Sanger sequencing of INPPL1 confirmed homozygosity for a missense mutation, p.Pro659Leu, in both siblings. In a validation cohort of an additional 12 unrelated families with opsismodysplasia, homozygosity or compound heterozygosity for INPPL1 missense, nonsense or splicing mutations was found in 5 families. Thus mutations in INPPL1 that are predicted to result in loss of enzymatic function were found in 6/13 or (~45%) of opsismodysplasia families. INPPL1 encodes an SH2-containing inositol polyphosphate-5 phosphatase. These membrane-bound signaling molecules regulate cell proliferation and survival, and are involved in cytoskeletal reorganization and vesicular trafficking through recruitment of effector proteins to the cell surface. Loss of function mutations in other members of the 5-phosphatase family cause Lowe oculocerebrorenal syndrome, which is, among other findings, characterized by vitamin D-resistant rickets. Comparison of INPPL1 mutation negative and mutation positive families provided insights into the relationship between opsismodysplasia with and without phosphate wasting. These discoveries enable direct genetic testing of individuals in whom the diagnosis of opsismodysplasia is considered They also will facilitate more accurate predictions about the natural history and outcome of individuals with opsismodysplasia, and provide a starting point for exploring the molecular and developmental mechanism underlying opsismodysplasia.

2741T

Novel Loci for Autosomal Dominant Nocturnal Frontal Lobe Epilepsy. *S.E. Heron[1], M. Bahlo[2,3], K.R. Smith[2,4], P. Hickey[2], C. Bromhead[2], C.P. Derry[5], J.S. Duncan[6], S. Howell[7], S.F. Berkovic[5], I.E. Scheffer[5,8], L.M. Dibbens[1].* 1) Epilepsy Research Program, School of Pharmacy and Medical Sciences, University of South Australia, Adelaide, South Australia, Australia; 2) Bioinformatics Division, The Walter and Eliza Hall Institute of Medical Research, Melbourne, Victoria, Australia; 3) Department of Mathematics and Statistics, The University of Melbourne, Melbourne, Victoria, Australia; 4) Department of Medical Biology, The University of Melbourne, Melbourne, Victoria, Australia; 5) Epilepsy Research Centre, Department of Medicine, University of Melbourne, Melbourne, Victoria, Australia; 6) Department of Clinical and Experimental Epilepsy, Institute of Neurology, University College London, Queens Square, London, United Kingdom; 7) Department of Neurology, Royal Hallamshire Hospital, Sheffield, United Kingdom; 8) Centre for Neuroscience and Melbourne Brain Centre, Florey Neuroscience Institute, University of Melbourne, Melbourne, Victoria, Australia.

Autosomal dominant nocturnal frontal lobe epilepsy (ADNFLE) is a seizure disorder characterized by violent motor seizures arising during sleep. Approximately 10% of cases are caused by mutations in the nicotinic acetylcholine receptor genes (nAchR) CHRNA4, CHRNB2 and, in rare cases, CHRNA2. The genetic basis of the majority of cases remains unknown. We ascertained two families with severe ADNFLE accompanied by psychiatric features and intellectual disability. No nAchR mutations were found in either family and linkage to the known loci was excluded, suggesting that both families had mutations at a novel locus. We genotyped the members of both families using Affymetrix or Illumina SNP arrays and performed genome-wide linkage analysis using MERLIN. Linkage for Family A was found to a single locus at chromosome 9q34.2. Linkage for Family B was found to two loci with identical LOD scores: a locus at chromosome 8q22.1–24.13 and a locus at chromosome 15q25.1–26.1. Both the chromosome 8 and chromosome 9 linkage intervals represent new loci for ADNFLE. The chromosome 15 locus has been previously described. The chromosome 8 and 15 loci contain 131 and 73 genes respectively. The chromosome 9 locus contains 99 genes. The identification of the causative genes at these loci will reveal new pathways in the pathogenesis of ADNFLE.

2742F

Identification of a new gene for X-linked Charcot-Marie-Tooth neuropathy using linkage and exome analysis. *M. Kennerson[1,2,3], E. Yiu[4,5], D. Chuang[6], S-C. Tso[6], C. Ly[1], A. Kidambi[1], G. Perez Siles[1], M. Ryan[4,5], G. Nicholson[1,2,3].* 1) Northcott Neuroscience, ANZAC Res Inst, Concord, Australia; 2) Molecular Medicine Laboratory, Concord Hospital, Concord, NSW Australia; 3) Sydney Medical School, University of Sydney, NSW Australia; 4) Children's Neuroscience Centre, Royal Children's Hospital, Victoria, Australia; 5) Murdoch Childrens Research Institute, Victoria, Australia; 6) UT Southwestern Medical Center, Dallas, Texas, USA.

Charcot Marie Tooth neuropathy (CMT) is a clinically and heterogeneous group of disorders characterized by chronic motor and sensory polyneuropathy. It is the most common hereditary neuromuscular disorder presenting in neurogenetics clinics and affects approximately 1 in 2,500 population. X-linked Charcot-Marie-Tooth (CMTX) disease accounts for 10–15% of all CMT. To date, there are five loci reported (CMTX1, CMTX2, CMTX3, CMTX4 Cowchock and CMTX5) and two genes Connexin 32 (Cx32) and phosphoribosyl pyrophosphate synthetase (PRPS1) have been identified for CMTX1 and CMTX5 respectively. We have identified a large family with probable X-linked dominant axonal motor and sensory polyneuropathy. The affected male proband presented at 9 years of age with foot deformity and gait abnormalities. Clinical findings included bilateral pes cavus, clawed toes, distal wasting and weakness and absent ankle jerks. There is no male to male inheritance and carrier females have very mild clinical features. We have screened the Cx32 gene in the proband (coding, UTR and promoter) and have excluded pathogenic mutations. Using traditional linkage analysis we mapped a new locus for X-linked dominant CMT to a 1.6-Mb interval on Xp11.1-p21.3. Whole exome sequencing of family members (2 affected males and one normal male) identified a non-synonymous exonic variant that maps within the established linkage region. The variant fully segregates with the affected phenotype and has been excluded in 1200 ethnically matched control chromosomes. The variant occurs in a highly conserved amino acid residue of the protein and is predicted to be pathogenic. Genetic validation and functional studies supporting pathogenicity of the variant will be presented. The combination of linkage analysis and exome sequencing continues to be a powerful approach for inherited peripheral neuropathy gene discovery.

2743W

A novel splice site Mutation in Erlin2 gene causes complicated form of Hereditary Spastic Paraplegia in a Saudi Family. *S. Majid[1], S. Bohlega[3], S. Hagos[1], H. Dossari[1], K. Ramzan[1], Z. Hassnan[2].* 1) Department of Genetics; 2) Department of Medical Genetics; 3) Department of Neurosciences,King Faisal Specialist Hospital and Research Center, Riyadh, Saudi Arabia.

Hereditary Spastic paraplegias (HSP) encompass clinically and genetically heterogeneous group of neurodegenerative disorders characterized by insidiously progressive weakness and spasticity of the lower extremities. We describe a consanguineous family segregating a complicated form of hereditary spastic paraplegia in autosomal recessive pattern. The 2 affected siblings have spasticity of the lower extremities, speech delay, and mild intellectual disability. Their upper extremities were mildly hypertonic. An intronic splice acceptor site mutation in the Erlin 2 gene was found to be responsible for causing this disorder. Erlin-2 protein is a mediator of Endoplamic Recticulum Degradation Pathway (ERAD) which helps to remove the aberrant proteins. An ERAD pathway targets certain proteins which are needed to maintain cellular homeostasis in tightly regulated level. Our results, in concurrence with previous studies of Erlin-2 suggest that alteration in Erlin-2 gene is a common cause of Complicated HSP thereby increasing the spectrum of mutations in SPG18 locus.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2744T

Clinical and genetic study of a new form of familial multiple keratoakanthoma in a Tunisian family: The palmoplantar familial multiple keratoacanthoma. *O. Mamaï[1], L. BoussofaraARA[2], L. Adala[1], I. Ben Charfeddine[1], S. Bouraoui[1], A. Amara[1], R. Nouira[2], M. Denguezli[2], B. Sriha[3], M. Gribaa[1], A. Saad[1].* 1) Laboratory of Human Cytogenetics, Molecular Geneti, Farhat Hached University Hospital, Sousse, Sousse, Tunisia; 2) Departement of Dermatology and Venerology, Farhat Hached University Hospital, Street Ibn El Jazzar, 4000 Sousse, Tunisia; 3) Department of Pathology and Cytology, Farhat Hached University Hospital, Street Ibn El Jazzar, 4000 Sousse, Tunisia.

Keratoacanthoma (KA) is a benign, self-healing epithelial neoplasm; it is characterized by rapid growth and spontaneous regression with residual scarring. It is considered benign growth but evidence is accumulating in favor of a low grade squamous cell carcinoma (SCC). There is tree form of familial multiple keratoacanthoma (FMKA). We investigated a five generation Tunisian family in which 13 members were affected by FMKA and seemed to be Ferguson Smith types transmitted as an autosomal dominant pattern with ulcerate lesions. This disease has been linked to 9q22.3–31 locus, and caused by specific mutations in TFGBR1gene. A detailed disease history with particular attention about the age of onset, distribution, and clinical course of their skin lesions were noted. The whole family was genotyped for this locus; karyotype and the fragility test were performed for all affected persons. After this study a genome wide scan genotyping by Human660W-Quad BeadChip Illumina SNP array was performed. We found particular palmoplantar localization for the lesion with conjuntival keratoacanthoma association and an early age of onset. Haplotype analysis, for 9q22.3–31 locus, did not found any common haplotype linked to FMKA in this family and no mutation was detected in the TGFB11 gene. All patients had a normal karyotype with absence of chromatin fragility. After SNP array analysis we found a delineated single candidate locus of 10 Mb on chromosome 17. We suggested a new form of FMKA named as: the Palmoplantar Familial Multiple Keratoacanthoma (PPFMKA) linked to this new locus that we project to sequence and identify the causative mutation.

2745F

FBLN5 mutation as the cause of HMSN I in a Czech family -usefulness of combination of SNP linkage and exome sequencing analysis. *D. Brožková, P. Laššuthová, J. Lisoňová, M. Krůtová, P. Seeman.* DNA laboratory, Department, University hospital Motol, Prague, Czech Republic.

Hereditary motor and sensory neuropathy is genetically very heterogeneous and can be caused by mutations in more than 40 genes. Sometimes it is difficult to find the causal mutation even in the families with more affected members. A Czech family with an autosomal dominant type of demyelinating HMSN was initially tested in 2001 for the most common causes of HMSN I: CMT1A/HNPP, GJB1, MPZ, PMP22 mutations. Later linkage analysis and additionally exome sequencing were performed. The disease onset is between 30–50 years. Two sisters 57 and 61 old show muscle atrophies below knees, unsecure walking and extremely reduced vibration sense on lower limbs. In three of their sons HMSN I was diagnosed mainly by slow nerve conduction velocity in the second decade, one other son is healthy. Genomewide SNP genotyping was performed on 8 persons with 10k Affymetrix chips. The linkage analysis revealed 20 regions with the LOD score 1.204. The regions with linkage comprised 2 HMSN associated genes - MFN2 and RAB7, but no mutation was detected by sequencing in them. Other HMSN genes and regions were excluded by the linkage, therefore we expected a new not yet known form of the HMSN. Exome sequencing of the 2 patients was performed on Illumina platform and over 18 000 variants were detected in each. The filtering criteria for candidate variants were as follows: common heterozygous variants, coding, not present in db SNP. The criteria fulfilled 357 variants, but after applying of linkage intervals only 5 of them remained. Two of the variants have lower in silico predictions to be a pathogenic. Sanger sequencing confirmed the segregation of the remained 3 variants in all affected patients, but only variant p.R373C in FBLN5 gene seemed most likely to be causal for HMSN I in family. Recently a paper about FBLN5 as a novel gene associated with demyelinating HMSN was published by another group reporting a family with AD HMSNI from Austria with the same p.R373C mutation. Since the Czech family originates from a region close to the Austria, we expect the common origin of this mutation in both families. We conclude that the mutation p.R373C in FBLN5 is causal for HMSN I in our family and we did not observed any apparent signs of macular degeneration or carpal tunel syndrome in any of the patients yet as reported in the Austrian family. Our family seems to by only the second family reported worldwide with FBLN5 mutation and classical HMSN I phenotype. Support- IGA MHCZ NT 11521/4.

2746W

Expansion of a repeat motif in the 5' untranslated region of EIF4A3 causes craniofacial and limb defects (Richieri-Costa-Pereira syndrome). *F. Favaro[1,2,3], RM. Zechi-Ceide[1], L. Alvizi[2,3], J. Souza[4], TM. Felix[5], SRF. Twigg[6], SJ. McGowan[6], AOM. Wilkie[6], A. Richieri-Costa[1], ML. Guion-Almeida[1], MR. Passos-Bueno[2, 3].* 1) Department of Clinical Genetics, Hospital of Rehabilitation of Craniofacial Anomalies (HRCA), University of São Paulo, SP, Brazil; 2) Department of Genetic, Institute of Biosciences of University of São Paulo, São Paulo, SP, Brazil; 3) Human Genome Research Center, Institute of Biosciences of University of São Paulo, São Paulo, SP, Brazil; 4) Assistance Center for Cleft Lip and Palate (CAIF), Curitiba-PR, Brazil; 5) Division of Medical Genetics, Hospital of Porto Alegre, Brazil; 6) Weatherall Institute of Molecular Medicine, University of Oxford, Oxford, England.

Richieri-Costa-Pereira syndrome (RCPS; OMIM 268305) is an autosomal recessive acrofacial dysostosis comprising Robin sequence, cleft mandible and limb defects. Despite previous efforts, no causative genetic mechanism for RCPS has been identified. We conducted the present study to identify the pathogenic mechanism of RCPS. We included DNA samples of 20 affected individuals ascertained in a geographic isolate and their non-affected relatives whenever possible. We conducted linkage analysis using SNP microarray 50K in 8 affected individuals and in 3 non-affected sibs. The best homozygous candidate region was at chromosome 17q25 (physical position: 68465456bp–80253028bp). Segregation analysis with markers mapped at this region in all the affected patients and non-affected relatives (n=70) supported linkage between markers rs2289534 and rs3829612, spanning 128.5kb. Neither Sanger sequencing (SS) nor Exome sequencing (ES) revealed any obvious pathogenic mutation in the coding regions of the 4 genes mapped in the 17q25 region or in the rest of the genome. Resequencing of exon 1 of one of the candidate genes revealed that the affected individuals have 16 repeats of a 20 nucleotide (nt) region at the 5'UTR of the EIF4A3 gene. We next verified that homozygosity of the 16 repeats were exclusively present among the affected individuals and not found in the parents or non-affected sibs or in 520 Brazilian controls. We also verified that the 20 nt repeat motif is transcribed in only one of the 2 EIF4A3 isoforms and through qRT PCR, we observed that the 16 repeat-isoform is associated with very low transcription levels (P=0,039) of the EIF4A3. We also found that 3 other Brazilian RCPS patients ascertained elsewhere were also homozygous for the 16 repeats at the 5' UTR of the EIF4A3. Homozygous mutations in EIF4A3 are very likely the cause of RCPS. Exome sequencing did not detect this alteration possibly because it is a CG rich region and have many repetitions. Of the 4 genes in this candidate region EIF4A3 is also the best functional gene. It is involved in splicing processes of many types of RNA, adding RCPS to the list of craniofacial syndromes caused by mutations in genes involved in RNA processing and translation. Reduction in EIF4A3 is possibly interfering in important molecular events during early development, as supported by knockdown studies in zebrafish (Alvizi et al 2012 submitted to ASHG). CEPID/FAPESP, CNPq.

2747T

A recurrent mutation of keratin 4 gene causing white sponge nevus in a Japanese family. *M. Kimura[1], T. Nagao[1,2], J. Machida[1,3], S. Yamaguchi[1], A. Shibata[1], G. Takeuchi[1,2], H. Miyachi[1], K. Shimozato[1], Y. Tokita[4].* 1) Maxillofacial Surgery, Aichi-Gakuin University School of Dentistry, Nagoya, Japan; 2) Department of Oral and Maxillofacial Surgery, Okazaki City Hospital, Okazaki, Japan; 3) Department of Oral and Maxillofacial Surgery, Toyota Memorial Hospital, Toyota, Japan; 4) Perinatology, Institute for Developmental Research, Aichi Human Service Center, Kasugai, Japan.

White sponge nevus (WSN;OMIM 193900) is a rare autosomal dominant disorder characterized by painless, white, and `spongy' plaques of the oral mucosa. Less frequently, nasal, esophageal, laryngeal, vaginal and anal mucosa may also be affected. Although it is a pathologically benign lesion, there is no effective therapy so far. Mutations in keratin 4 (*KRT4*) and 13 (*KRT13*) genes have been reported as responsible genes of WSN. The purpose of this study is to identify pathogenic mutations of *KRT4* or *KRT13* in a Japanese family with three generations affected by WSN, and to reveal its dysfunction. Proband in this family was a 38-year-old female and four individuals of the family were affected by WSN. Mutation search was performed using DNA from peripheral blood lymphocyte and functional analysis was carried out using HEK293 cells. A heterozygous 3bp deletion (N160Del) was found in the helix-initiation motif at the beginning of alpha-helical domain 1A of *KRT4*. In addition, this mutation was found to lead the disruption of the keratin filament network. From the results of the study, we conclude that the amino acid residue of asparagine in position 160 of *KRT4* is functionally important for keratin filament network formation.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2748F

Molecular analysis of benign familial neonatal-infantile convulsion in a Japanese family. *Y. Yamada[1], K. Yamada[1], D. Fukushi[1], K. Miura[2], N. Nomura[1], A. Yamano[1], T. Kumagai[3], N. Wakamatsu[1].* 1) Dept Gen, Inst Dev Res, Aichi Human Service Ctr, Kasugai Aichi, Japan; 2) Dept Pediatrics, Nagoya Univ Sch Med, Nagoya, Japan; 3) Ctrl Hosp, Aichi Human Service Ctr, Kasugai Aichi, Japan.

Benign familial neonatal-infantile convulsion (BNFC) is an autosomal dominant disorder in which afebrile seizures occur in clusters during the first year of life, without neurologic sequelae. As the gene responsible for BNFC, sodium channel and potassium channel genes have been reported. Focusing on a large family including 8 patients with benign spasm, we investigated to define the genetic mutation causing BNFC and establish a method for the simple detection of the mutation. Linkage analysis of the family pathogenesis narrowed down the gene locus to the region between D2S335 and D2S142 (Lodscore 2.1 maximum). Some sodium channel genes (*SCN1A, SCN2A, SCN3A, SCN7A*) were located in this chromosome 2q23-2 domain. Firstly, we analyzed the nucleotide sequences in all exons and adjacent splice sites of *SCN2A* likely to be the causative gene, but no responsible mutations were detected. In the analysis of *SCN1A* and*SCN3A*, we also could not identify any mutations. However, we identified a single nucleotide substitution (IVS18+7A>G) in a new candidate gene, *SCN7A*. As the identified mutation was located at the 5'-end region in intron 18 as the splice-donor site, it is suggested that the mutation caused abnormal splicing of *SCN7A* mRNA. Then, we analyzed the mRNA using the cultured B-lymphoblast from patients in the family, but have not been able to confirm the expression of abnormal mRNA. Next, we constrated a mini chimeric gene containing a part of patients' intron 18, transfected HEK293 cells with it, and analyzed the *SCN7A* mRNA. However, these studies have not proved that the substitution IVS18+7A>G causes splice error. Most recently, the IVS18+7A>G has been reported as a rare single nucleotide polymorphism (SNP, A:0.992, G:0.008). Therefore, we performed array CGH analysis considering the posibility of gene deletion, but chromosome abnormalities causing BFNC have not been identified. Exome analysis of all genes of the affected and non-affected individuals in the family are now in progress.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2749W

A Newly Identified Locus for Benign Adult Familial Myoclonic Epilepsy on Chromosome 3q26.32-3q28. *P. Yeetong[1,2,3], S. Ausavarat[2,3], R. Bhidayasiri[4], K. Piravej[5], N. Pasutharnchat[4], T. Desudchit[6], C. Chunharas[4], J. Loplumlert[4], C. Limotai[4], K. Suphapeetiporn[2,3], V. Shotelersuk[2,3].* 1) Inter-Department Program of Biomedical Sciences, Faculty of Graduate School, Chulalongkorn University; 2) Center of Excellence for Medical Genetics, Department of Pediatrics, Faculty of Medicine, Chulalongkorn University; 3) Excellence Center for Medical Genetics, King Chulalongkorn Memorial Hospital, Thai Red Cross; 4) Division of Neurology, Department of Medicine, Faculty of Medicine, Chulalongkorn University; 5) Department of Rehabilitation Medicine, Faculty of Medicine, Chulalongkorn University; 6) 6Division of Neurology, Department of Pediatrics, Faculty of Medicine, Chulalongkorn University, Bangkok, 10330, Thailand.

Benign Adult Familial Myoclonic Epilepsy (BAFME) is an autosomal dominant disorder characterized by adult-onset cortical tremor or action myoclonus predominantly in the upper limbs, and generalized seizures. We investigated a Thai BAFME family. Clinical and electrophysiological studies revealed that 13 were affected with BAFME. There were a total of 24 individuals studied. Genetic analysis by genome-wide linkage study (GWLS) was performed using 400 microsatellite markers and excluded linkage of the previous BAFME loci, 8q23.3-q24.1, and 2p11.1-q12.2. GWLS showed that the disease-associated region in our Thai family was linked to a newly identified locus on chromosome 3q26.32-3q28. Next-generation sequencing was used to sequence the whole critical region of 18 Mb. In addition, array comparative genomic hybridization was performed to identify copy number variations. Even with both advanced assays, the etiologic gene remains elusive. Discovery of the fourth BAFME chromosomal region will facilitate the identification of the responsible gene. This may provide further understanding into the molecular basis of epilepsy and better insight into the disease mechanism leading to more effective treatment of this disorder.

2750T

Calpain-5 causes autoimmune uveitis, retinal neovascularization and photoreceptor degeneration. *V.B. Mahajan[1], J.M. Skeie[1], A.G. Bassuk[2,3], J.H. Fingert[1], T.A. Braun[1,4], H.T. Daggett[1], J.C. Folk[1], V.C. Sheffield[1,2,5], E.M. Stone[1,5].* 1) Department of Ophthalmology and Visual Sciences, University of Iowa, Iowa City, IA; 2) Department of Pediatrics, University of Iowa, Iowa City, IA, United States; 3) Department of Neurology, University of Iowa, Iowa City, IA, United States; 4) Department of Biomedical Engineering, University of Iowa, Iowa City, IA, United States; 5) Howard Hughes Medical Institute, University of Iowa, Iowa City, IA, United States.

Background: Autosomal dominant neovascular inflammatory vitreoretinopathy (ADNIV) is an autoimmune condition that sequentially mimics uveitis, retinitis pigmentosa, and proliferative diabetic retinopathy as it relentlessly progresses from a relatively subtle onset to complete blindness over five decades. The purpose of this study was to identify the disease-causing gene and to understand how a single gene could cause this wide range of clinical findings.

Methods: Eye examinations were performed in three different kindreds. High-density genotyping narrowed the disease interval, and exome sequencing of two individuals from the largest kindred identified the disease-causing gene. Sanger DNA sequencing was used to confirm the disease-causing mutation and to identify the mutations in the other two families. Immunohistochemistry, structural modeling, and heterologous cell expression of the mutant alleles were performed to investigate the pathophysiology.

Results: Two different missense mutations were identified on chromosome 11q13 in exon 6 of the *CAPN5* gene, which encodes the calcium-activated cysteine protease calpain-5. Structural modeling revealed that both mutations lie within a putative regulatory domain near the active site. Calpain-5 is expressed in retinal photoreceptor cells. Both mutations cause mislocalization from the cell membrane to the cytosol in vitro.

Conclusions: This is the first report of a specific molecular cause autoimmune eye disease. *CAPN5* is only the second member of a large protease gene family to cause a human Mendelian disorder. Further investigation of these mutations is likely to provide new insight into the pathophysiologic mechanisms of common eye diseases ranging from autoimmune disorders to diabetic retinopathy.

2751F

Soluble ICAM-1 levels are associated to a protective prognosis in sickle cell retinopathy. *P.R.S. Cruz[1], T.R. Zaccarioto[2], F.N. Mitsuuchi[3], S.A.C. Pereira Filho[3], R.P.C. Lira[3], I. Meyer[4], M.A.C. Bezerra[5], K.Y. Fertrin[6], F.F. Costa[6], M.B. Melo[1].* 1) Laboratory of Human Molecular Genetics, Center of Molecular Biology and Genetic Engineering (CBMEG), University of Campinas - UNICAMP, Campinas, São Paulo, Brazil; 2) Department of Clinical Pathology, School of Medical Sciences, State University of Campinas-UNICAMP, Campinas, São Paulo, Brazil; 3) Department of Ophthalmology, Faculty of Medical Sciences, University of Campinas - UNICAMP, Campinas, São Paulo, Brazil; 4) Eyes Instute of Recife - IOR, Recife, Pernambuco, Brazil; 5) Center of Hematology and Hemotherapy of Pernambuco - HEMOPE, Recife, Pernambuco, Brazil; 6) -Hemotherapy Center, Department of Clinical Medicine, University of Campinas - UNICAMP, Campinas, São Paulo, Brazil.

Hemoglobinopathies are among the most common inherited diseases globally. The HbS and HbC variants of the beta globin gene lead to sickle cell anemia (HbSS) and SC hemoglobinopathy (HbSC). Sickled red blood cells are more likely to adhere to the vascular endothelium, resulting in vascular obstruction and ischemia. Sickle cell retinopathy is an ocular condition derived from vaso-occlusion in the microvasculature of the retina leading to hypoxia, neovascularization and eventually to blindness. The reasons for HbSC patients being more likely to exhibit ocular manifestations than patients with HbSS are not clear. A possible relation to the rate of sickling, hematocrit and blood viscosity is postulated. These facts combined to endothelium activation and altered expression of adhesion molecules such as P-selectin, E-selectin, intercellular adhesion molecule-1 (ICAM-1) and vascular cell adhesion molecule-1 (VCAM-1) offer a more feasible view of the vaso-occlusion process in retinopathy development. There is also evidence of angiogenesis regulation molecules participation, namely angiopoetin-1 (Ang-1), angiopoetin-2 (Ang-2), pigment epithelium-derived factor (PEDF) and vascular endothelial growth factor (VEGF). In order to evaluate the involvement of these factors, the plasmatic Ang-1, Ang-2, sVCAM-1 and sICAM-1 levels were quantified by ELISA in 37 Brazilian sickle cell retinopathy patients (13 HbSS and 24 HbSC) which were compared to 34 Brazilian sickle cell patients without retinopathy (19 HbSS and 15 HbSC), along with healthy sex and age-matched individuals (n=10). Moreover, two polymorphisms (-358G/A and -790T/C) in the PEDF gene promoter region were analyzed in 92 patients (32 HbSC e 60 HbSS) by direct sequencing. It was observed that soluble ICAM-1 is elevated in the HbSC non-retinopathy group compared to the affected HbSC group ($p=0.012$). Furthermore Ang-1 and Ang-2 levels were altered in SCD compared to healthy controls ($p=0.001$ and $p<0.001$, respectively). Ang-2 levels were also heightened in HbSS in comparison to HbSC individuals ($p<0.001$). Ang-2 levels were negatively correlated to hematocrit in both HbSC (Pearson's correlation coefficient, $r = -0.719$, $p< 0.001$) and HbSS patients ($r = -0.438$, $p = 0.046$). However, sVCAM-1, sE-selectin, sP-selectin, VEGF and PEDF analysis exhibited no statistical significant differences among groups. These results underline the importance of carefully analyzing sickle cell disease different genotypes separately.

2752W

X-linked CHARGE-like Abruzzo-Erickson syndrome and classic cleft palate with ankyloglossia result from TBX22 splicing mutations. *P. Stanier[1], E. Pauws[1], E. Peskett[1], C. Boissin[1], K. Mengrelis[1], E. Carta[1], M.A. Abruzzo[2], M. Lees[1,3], G.E. Moore[1], R.P. Erickson[4].* 1) UCL Institute of Child Health, 30 Guilford Street, London, WC1N 1EH, UK; 2) Dept. of Biology, California State University, Chico, CA, 95929-0515, USA; 3) Great Ormond Street Hospital, London, WC1N 3JH, UK; 4) Dept. of Pediatrics, University of Arizona Health Science Center, Tucson, AZ, 85724-5073, USA.

X-linked cleft palate (CPX) is caused by mutations in the gene encoding the TBX22 transcription factor and is known to exhibit phenotypic variability, usually involving either a complete, partial or submucous cleft palate, with or without ankyloglossia. This study hypothesized a possible involvement of TBX22 in a family with X-linked, CHARGE-like Abruzzo-Erickson syndrome, of unknown etiology. The phenotype extends to additional features including sensorineural deafness and coloboma, which are suggested by the Tbx22 developmental expression pattern but not previously associated in CPX patients. A novel TBX22 splice acceptor mutation (c.593-5T>A) was identified that tracked with the phenotype in this family. A novel splice donor variant (c.767+5G>A) and a known canonical splice donor mutation (c.767+1G>A) affecting the same exon were identified in patients with classic CPX phenotypes and were comparatively analyzed using both in silico and in vitro splicing studies. All three variants were predicted to abolish normal mRNA splicing while the in vitro assay indicated that use of alternative splice sites was a likely outcome. Collectively, the data demonstrated the functional effect of several novel splice site variants but most importantly confirms that TBX22 is the gene underlying Abruzzo-Erickson syndrome, expanding the phenotypic spectrum of CPX.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

† Trainee Award Finalist

2753T

Enlarged vestibular aqueduct syndrome - characterization of three new mutations. *E. Prera, S. Arndt, A. Aschendorff, R. Birkenhäger.* Otorhinolaryngology H&N Surg, Univ Med Ctr Freiburg, Freiburg, Germany.

Introduction: Enlarged vestibular aqueduct (EVA) syndrome is a congenital autosomal recessive anomaly, characterised for a progressive bilateral sensorineural hearing impairment, with temporal bone abnormalities (enlarged vestibular aqueduct). In combination with a euthyroid goiter it is called a Pendred syndrome. 40% of affected individuals have a mutation in the SLC26A4 gene. The gene product Pendrin is expressed in the inner ear, thyroid gland and kidney. This gene product is an iodide/chloride transporter. Here we describe the identification of three new mutations, which in combination with an already known mutation lead to an enlarged vestibular aqueduct syndrome. Patients and Methods: In this analysis four members of two different families were included. Complementary to the audiologic and radiologic diagnosis complete molecular-genetic analyses were performed with all family members. Sensorineural hearing impairment occurred on one of the involved members at the age of 22 years and on the second family on one member at the age of 12 years. All intron transitions and exons from the SLC26A4 gene were sequenced directly. Results: Two loss of function mutations (c.1353_1354insGCT;IVS7-2T>C) and two missense-mutations (c.412G>T, p.V138F; c.1238A>G, p.E413R) were identified. Conclusion: The discovery of these new mutations can derive to a better understanding of the phenotype variability of EVA and Pendred syndrome, and leading to a better understanding of phenotype-genotype correlation.

2754F

A Novel Molecular Basis for Carpenter Syndrome. *S. Movva, U. Kotecha, R. Puri, I.C. Verma.* Center of Medical Genetics, Sir Ganga Ram Hospital, New Delhi, India.

Carpenter syndrome (OMIM:201000) is a rare autosomal recessive disorder characterized by craniosynostosis, obesity, polydactyly and soft-tissue syndactyly. The gene for Carpenter syndrome (RAB23), was identified in 2007, following which there have been three reports describing different types of mutations in the gene, although the spectrum of identified mutations in has been limited. In the present report we describe a novel homozygous mutation in an Asian Indian patient. A non-consanguineously married woman from Punjab was referred during her third conception. Her first child had died within 24 hrs of life, and had an abnormal shape of the skull with ambiguous genitalia. The second born is a healthy girl. In the current pregnancy antenatal ultrasound study at 31 weeks of gestation showed polyamnios, clover leaf skull, polydactyly in both lower and one upper limb. In view of the previous history, possibility of an autosomal recessive disorder was entertained. Examination of the neonate after delivery, along with radiographs, CT scan of the brain with bone window 3D reconstruction was recommended. A male baby was born at 32 weeks and succumbed within 2 hrs of life. As the family was from another state and the newborn could not be examined by us, the attending pediatrician took photographs with blood sample for karyotyping and DNA studies. Three years later the family came for preconception counseling. The neonatal photographs were reviewed and a diagnosis of Carpenter syndrome was strongly suspected. RAB23 gene was sequenced using the stored DNA of the baby. We identified a novel mutation which converted the stop codon into a coding one, resulting in formation of a lengthened protein. The Stop238E causes an open reading frame of the coding sequence in the exon 7 of the gene up to nucleotide no. 918 followed by a subsequent stop codon sequence at nucleotide nos.919–921. This results predictably in an additional 69 amino acid residue sequence leading to a longer mutated Rab23 protein containing 307 amino acids in contrast to the wild type protein containing 237 amino acid residues. The present study identified a novel molecular basis for the Carpenter syndrome due to a stop codon mutation in the RAB23 gene adding to the spectrum of reported mutations. The case also highlights the importance of clinical history, taking photographs and storing the DNA, all of which helped to establish the diagnosis.

2755W

De novo activating mutations in Phosphatidylserine Synthase 1 cause a rare mental retardation-multiple congenital anomalies syndrome. *S.B. Sousa[1,2], D. Jenkins[3], E. Chanudet[4], A. Barnicoat[5], K. Chrzanowska[6], J. Sá[2], J. Saraiva[2], R. Scott[5], D. Wattanasirichaigoon[7], G. Anderson[8], G. Tasseva[9], J. Vance[9], P. Beales[3], G.E. Moore[1].* 1) Clinical and Molecular Genetics Unit, UCL Institute of Child Health, London, United Kingdom; 2) Serviço de Genética Médica, Hospital Pediátrico de Coimbra, Coimbra, Portugal; 3) Molecular Medicine Unit, UCL Institute of Child Health, London, UK; 4) GOSgene, UCL Institute of Child Health, London, UK; 5) Clinical Genetics Department, Great Ormond Street Hospital, London; 6) Department of Medical Genetics, The Children's Memorial Health Institute, Warsaw, Poland; 7) Department of Pediatrics, Faculty of Medicine, Ramathibodi Hospital, Mahidol University, Bangkok, Thailand; 8) Histopathology Department, Great Ormond Street Hospital for Children, London; 9) Group on the Molecular and Cell Biology of Lipids, Department of Medicine, University of Alberta,Edmonton, Canada.

Phosphatidylserine is a component of mammalian cell membranes, usually representing 2–10% of total phospholipids. Phosphatidylserine is known to play important roles in biological processes such as apoptosis and cell signalling. We collected clinical information and DNA/tissue samples from four unrelated patients with a rare syndrome characterized by mental retardation and multiple congenital anomalies, including facial dysmorphology and skeletal abnormalities. Affected individuals of both sexes have been identified on several ethnic backgrounds, but no familial cases are known and parental consanguinity has not been reported. We therefore hypothesized that the disease is caused by a dominant de novo mutation in affected individuals. We performed whole exome sequencing in the four unrelated affected patients, and identified novel heterozygous missense mutations in the Phosphatidylserine Synthase 1 gene (PTDSS1) in all affected individuals. These mutations were validated by Sanger sequencing, were not present in the available control databases, affected conserved residues and were confirmed de novo in 2/2 cases where parental samples were available, further suggesting pathogenicity. The structure of the enzyme Phosphatidylserine Synthase 1 is not known in detail, but the identified missense mutations are both located in transmembrane domains. The reported PTDSS1 null-mutant mice were normal, suggesting a gain-of-function mechanism. This was supported by measuring PTDSS1 enzymatic activity in primary fibroblasts, which was strongly increased in patient samples compared to controls. Furthermore, expression of the corresponding missense mutant versions of human PTDSS1 in zebrafish embryos caused severe defects of the craniofacial skeleton that were not observed in embryos expressing the wild-type form. Ongoing studies aim to clarify the physiopathology of these mutations. In summary, we report the first human mutations affecting the phosphatidylserine pathway, which emphasize its importance and give insights into its role in neural and skeletal development.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2756T

Exome sequencing of a consanguineous Palestinian kindred with syndromic cleft palate. *H. Shahin[1], U. Sharaha[1], M. Lee[2], S. Saed[4], J. van Aalst[3], MC. King[2], T. Walsh[2].* 1) Life Sciences Department, Bethlehem University, Bethlehem, Palestinian Territory; 2) Department of Medical Genetics, University of Washington, Seattle, WA, USA; 3) Plastic & Reconstructive Surgery, University of North Carolina, Chapel Hill, NC, USA; 4) Palestinian Cleft Society, Palestine.

Cleft lip and cleft palate (CLP) are among the most common birth defects worldwide and confer lifelong morbidity on affected individuals. The trait is highly heterogeneous both phenotypically and genetically. We are studying a 4-generation consanguineous Palestinian kindred including seven affected relatives. The participants in the present project are three siblings (two male, one female), ages 23, 21, and 12 years, and their unaffected first-cousin parents. All three siblings have complete clefting of the soft palate, incomplete bilateral clefting of the hard palate, severe velopharyngeal insufficiency, and moderate to severe bilateral hearing loss. One male has retarded growth. Vision and cognitive development are normal in all siblings. The sibship also includes four unaffected brothers, two unaffected sisters, and a fourth affected child who is deceased. Exome sequencing of genomic DNA of the three affected siblings, followed by filtering for rare potentially damaging variants, yielded exactly one exome variant homozygous in all three affected siblings: COL2A1 c.2563G>A (p.G855S). COL2A1 is a collagen of cartilage. Glycine is conserved at this site in all sequenced vertebrates, and p.G855S is predicted to be damaging by all bioinformatics tools. Both parents were heterozygous and unaffected and 3 of 320 controls from the same region of the West Bank were also heterozygous and unaffected. Severe heterozygous mutations in COL2A1 are responsible for multiple human disorders, including achrondrogenesis, chrondrodysplasia, early onset osteoarthritis, Stickler syndrome type I, and spondyloepiphyseal dysplasia congenita (SEDC). Many disease-associated mutations in COL2A1 are missenses at glycine residues in the collagen triple helix repeats. The glycine at residue 855 lies between the first (aa 513–571) and second (aa 969–1016) collagen triple helix repeats. We speculate that although p.G855 and neighboring residues are highly conserved throughout vertebrates, p.G855S may be less severe than mutations in the triple helix repeats themselves, so that homozygosity for p.G855S is compatible with life. We are presently sequencing COL2A1 in other Palestinian families with cleft palate to determine if COL2A1-associated phenotypes may be more frequently recessive than previously appreciated.

2757F

Mutations in the DOCK6 gene in Adams-Oliver syndrome. *W. Wuyts[1], A. Laureys[1], E. Van Hul[1], M. De Smedt[2], K. Storm[1], N. Van der Aa[1].* 1) Dept Med Gen, Univ Antwerp, Edegem, Belgium; 2) Dept Med Genet, Leuven University Hospital, Leuven, Belgium.

Introduction Adams-Oliver syndrome (AOS)is a rare genetic disorder characterized by aplasia cutis congenita (ACC) and terminal transverse limb defects. Mutations in the ARHGAP31 gene have been described as a cause of autosomal dominant AOS while mutations in the DOCK6 gene were shown to be responsible for an autosomal recessive form of AOS. To investigate the contribution of both genes to the AOS phenotype we performed mutation analysis of ARHGAP31 and DOCK6 genes in a cohort of AOS patients. **Materials and Methods** A cohort of 20 patients with clinical characteristics of AOS were included in the study, including familial and sporadic cases. Mutation analysis of the ARHGAP31 gene was performed by sequence analysis of the terminal exon 12. Mutation analysis of the DOCK6 gene was performed by sequence analysis of all coding exons. **Results and discussion** For 12 patients potential recessive inherritance was suspected based on observed inherritance pattern or known or suspected consanguinity of the parents. Mutation analysis of all DOCK6 coding exons revealed the presence of homozygous inactivating mutations in 3 patients, including a nonsene mutation, a frameshift mutation and a multi-exon deletion. Mutation analysis of the ARHGAP31 terminal exon (exon 12) did not reveal a mutation in any of the patients analyzed. Our results confirm the involvement of the DOCK6 gene in Adams-Oliver syndrome. However, as the majority of patients in our cohort did not show a DOCK6 or ARHGAP31 mutation, further genetic heterogeneity is supected for AOS.

2758W

Novel mutations of the FGD1 gene in 16 patients with Aarskog-Scott syndrome. *T. Kaname[1], K. Yanagi[1], N. Okamoto[2], K. Kurosawa[3], Y. Izumikawa[1], Y. Fukushima[4], Y. Makita[5], M. Tsukahara[6], A. Altincik[7], S. Mizuno[8], K. Naritomi[1].* 1) Dept Med Gen, Univ Ryukyus, Nishihara, Japan; 2) Dept Dev Med, Osaka Med Centr Res Inst Maternal and Child Health, Izumi, Japan; 3) Div Med Genet, Clin Res Inst, Kanagawa Child Med Centr, Kanagawa, Japan; 4) Dept Med Genet, Shinshu Univ, Matsumoto, Japan; 5) Edu Centr, Asahikawa Med Univ, Asahikawa, Japan; 6) Faclt Health Sci, Yamaguchi Univ, Ube, Japan; 7) Faclt Med, Dokuz Eylul Univ, Turkey; 8) Dept Pediatr, Cent Hosp, Aichi Hum Serv Centr, Aichi, Japan.

Faciogenital dysplasia 1 (FGD1) gene was identified as a responsible gene for Aarskog-Scott syndrome (AAS) or facio-digito-genital dysplasia (MIM#305400), which is an X-linked disorder characterized by genital and skeletal abnormalities. To date, more than 25 mutations have been reported. However, neither phenotype/genotype correlations nor phenotypic features on patients with FGD1 mutation have been cleared. We identified nine mutations of the FGD1 gene in 16 Japanese or Turkish patients with clinical AAS. Of nine mutations, five are missense mutations, two are splice site mutations, and two are deletion mutations. In addition, we found two synonymous variations in the patients by PCR-HRM analysis. All mutations were not found in 200 controls, suggesting that the mutations would be pathogenic. Although the evidence for phenotype/genotype correlations between the type of mutations and clinical features was not found, many clinical features met the primary diagnostic criteria for AAS in the patients with mutations in the FGD1 gene.

2759T

Exome Sequencing Identifies an *IFIH1* Mutation Causing Singleton-Merten Syndrome by Dysregulation of the Innate Immune Response. *F. Rutsch[1], M. MacDougall[2], C. Lu[2], Y. Nitschke[1], I. Buers[1], O. Mamaeva[2], D.K. Crossman[2], J. Dong[2], C. Müller[1], H.G. Kehl[3], J. Kleinheinz[4], P. Barth[5], K. Barczyk[6], D. Bazin[7], J. Altmüller[8], H. Thiele[8], P. Nürnberg[8], W. Höhne[9], A.S. Feigenbaum[10], R. Hennekam[11].* 1) Department of General Pediatrics, University Children's Hospital, Münster, Germany; 2) Institute of Oral Health Research, University of Alabama Medical School, Birmingham, AL, USA; 3) Department of Pediatric Cardiology, University Children's Hospital, Münster, Germany; 4) Department of Cranio-Maxillofacial Surgery, Münster University Hospital, Münster, Germany; 5) Gerhard-Domagk Institute of Pathology, Münster University Hospital, Münster, Germany; 6) Institute of Immunology, Münster University Hospital, Münster, Germany; 7) Laboratoire de Physiques des Solides, Paris Sud University, Orsay, France; 8) Cologne Center for Genomics, University of Cologne, Cologne, Germany; 9) Department of Biochemistry, Charité University Medicine, Berlin, Germany; 10) Department of Clinical and Metabolic Genetics, Hospital for Sick Children, Toronto, Canada; 11) Department of Pediatrics, Academic Medical Center, Amsterdam University, Amsterdam, The Netherlands.

Singleton-Merten Syndrome (SMS; MIM 182250) is a rare autosomal-dominant disorder characterized by calcification of the ascending aorta and cardiac valves leading to premature death with psoriasis, dentin dysplasia, delayed tooth eruption, early loss of permanent teeth through root resorption and widened medullary cavities of the phalanges with focal osteoporosis. To identify the underlying disease gene, whole exome sequencing was performed on four SMS patients from three families. We identified the same heterozygous missense mutation segregating with the disease in all four patients in *IFIH1*, encoding for melanoma differentiation associated protein 5 (MDA5), not present in the Single Nucleotide Polymorphism or the 1000 Genomes databases. MDA5 regulates inflammatory signals belonging to the RIG-I-like receptor family of sensors for cytoplasmic double-stranded RNA and initiates antiviral responses by producing type I and type II interferons. Genome wide association studies have linked MDA5 SNPs with type I diabetes and psoriasis. The identified p.R822Q MDA5 mutation is at a highly conserved amino acid in the helicase domain. Structural analysis of this region revealed a positive surface potential for RNA orientation during unwinding that is changed, altering the potential affinity and/or activity of the helicase. Expression analysis (qRT-PCR and immunohistochemistry) showed MDA5 expression in heart tissue, skin, alveolar bone and tooth ameloblasts (enamel producing), odontoblasts (dentin producing) and the periodontal apparatus; all disease targeted tissues. Osteogenic array analysis of dental mineralizing SMS versus control cells showed ~55% of the gene tested were dysregulated. Ingenuity pathway analysis of Affymetrix GeneChip expression profiles revealed that inflammatory response/immunity networks and interferon and type I diabetes mellitus signalling pathways were the most significantly dysregulated. Transient and stable transfection of HEK 293 cells with wildtype, SMS-MDA5 and corrected SMS-MDA5 (corSMS-MDA5) vectors showed enhanced (dominant) cytokine production of INF-β and TNF-α associated with the SMS MDA5 mutation that was recovered to baseline levels with the corSMS-MDA5 construct. These data suggest that the deficient dental structure formation, abnormal root resorption, localized osteoporosis, cardiac calcification and psoriatic skin changes seen in SMS result from the p.R822Q MDA5 mutation by modulation of the innate immune response.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2760F

Mutations in PVRL1 and PVRL4 genes define nectinopathies as developmental field disorders of ectodermal derivatives and cutaneous syndactyly with or without cleft lip/palate. *E. Agolini[1,2], P. Fortugno[3,9], L.B. Ousager[4], A. Bygum[5], L. Graul-Neumann[6], K. Stieler[7], A. Kuechler[8], G. Zambruno[9], F. Brancati[10,1].* 1) Mendel Laboratory, IRCCS Casa Sollievo della Sofferenza, Rome, Italy; 2) Department of Molecular Medicine, Sapienza University, Rome, Italy; 3) Dermatology Unit, IRCCS Ospedale Pediatrico Bambino Gesù, Rome, Italy; 4) Department of Clinical Genetics, Odense University Hospital, Odense, Denmark; 5) Department of Dermatology and Allergy Centre, Odense University Hospital, Odense, Denmark; 6) Institute of Medical and Human Genetics, Charité Universitätsmedizin, Berlin, Germany; 7) Department of Dermatology and Allergy, Clinical Research Center for Hair and Skin Science (CRC), Charité Universitätsmedizin, Berlin, Germany; 8) Institut für Humangenetik, Universitätsklinikum, Essen, Germany; 9) Molecular and Cell Biology Laboratory, IRCCS Istituto Dermopatico dell'Immacolata, Rome, Italy; 10) Department of Biopathology and Diagnostic Imaging, Tor Vergata University, Rome, Italy.

The presence of hair and teeth anomalies, alopecia, and partial syndactyly of the fingers and/or toes defines the ectodermal dysplasia-syndactyly syndrome (EDSS1; OMIM #613573), a rare autosomal recessive disorder so far reported in no more than five families worldwide. Its disease-causative gene is PVRL4, encoding nectin-4, an immunoglobulin superfamily cell adhesion molecule implicated in the formation of cadherin-based adherens junctions. Nectin-4 function is closely related to its ability to form heterodimers with nectin-1, encoded by the paralog PVRL1, whose mutations are responsible for the autosomal recessive cleft lip/palate-ectodermal dysplasia syndrome (CLPED; OMIM #225000) also showing variable syndactyly. Here we describe two novel patients with EDSS in which we have identified two novel homozygous mutations in PVRL4. A third patient with abnormalities in hair, teeth and cutaneous syndactyly without cleft lip/palate, initially diagnosed with EDSS, was found mutated in PVRL1. The first family is a consanguineous Lebanese family with two members affected by EDSS, in which sequence analysis revealed a novel PVRL4 homozygous nonsense mutation. In the second family, the EDSS proband is the single child of healthy unrelated Danish parents. He has a large biallelic exonic deletion involving one PVRL4 exon, confirmed in his heterozygous parents by quantitative real-time PCR. Interestingly, a large 1.8 Mb homozygous region involving the PVRL4 locus at 1q23 was identified, arguing for a possible founder effect in the Danish population. In the third patient from Germany, no mutations in the PVRL4 gene were identified, while a novel nonsense mutation truncated the coding sequence of PVRL1. Notably this is the third patient mutated in PVRL1 and the first without oral defects. Our results define "nectinopathies" as clinically overlapping ED syndromes caused by defective nectins-1 and 4, two molecules essential for correct ectodermal derivatives and limb development.

2761W

ELN Mutations in Autosomal Dominant Cutis Laxa. *E. Lawrence[1], M. McGowan[1], C. Su[1], M. Mac Neal[1], K. Levine[1], A. Zaenglein[2], M. Bodzioch[3], A. Kiss[4], Z. Urban[1].* 1) Department of Human Genetics University of Pittsburgh; 2) Dept. of Dermatology Penn State Milton S. Hersey Medical Center; 3) Jagiellonian University Medical College; 4) Dept. of Clinical Genetics Universidade Federale de Ciencias de Porto Alegre.

Cutis laxa (CL) is an inherited skin disease with remarkable locus heterogeneity. Mutations in the elastin gene (ELN) cause autosomal dominant cutis laxa (ADCL). We have sequenced the last 5 exons of ELN in 63 consecutive probands with CL. Mutations were identified in 4 probands. Two had a previously published c.2296delA mutation in exon 32 and one had an earlier reported mutation in exon 30 (c.2177delC). Finally, one proband carried a novel mutation in exon 34 (c.2351delG). Examination of the origin of the mutations was possible in 3 families, 1 showing inherited, and 2 de novo mutations. In addition to CL, 3/5 affected individuals had mild-moderate obstructive pulmonary disease and one had a mild aortic root dilatation. Immunoblotting showed elevated canonical, but unaffected non-canonical transforming growth factor-beta signaling in fibroblasts from patients with ELN mutations. We conclude that ELN-positive CL explains approximately 6% of cutis laxa cases and elevated transforming growth factor-beta signaling is a disease mechanism in ADCL shared with other connective tissue diseases.

2762T

***RIPK4* mutations explain a portion of *IRF6*-negative popliteal pterygium cases.** *E.J. Leslie[1], S. Goudy[2], M. Cunningham[3], J.C. Murray[1].* 1) Department of Pediatrics, University of Iowa, Iowa City, IA; 2) Department of Otolaryngology, Vanderbilt University, Nashville, TN; 3) Department of Pediatrics, University of Washington, Seattle, WA.

Van der Woude syndrome (VWS) and popliteal pterygium syndrome (PPS) are two allelic, autosomal dominant cleft lip/palate disorders caused by heterozygous mutations in *IRF6*. However, mutations have not been detected in approximately 30% of VWS and 3% of PPS patients. There is substantial clinical overlap between PPS and the recessive pterygia syndrome Bartsocas-Papas syndrome (BPS), which has recently been determined to be caused by mutations in *RIPK4*. *Irf6*$^{-/-}$ and *Ripk4*$^{-/-}$ mice have nearly identical phenotypes and both genes are transcriptional targets of p63. Given these data, we hypothesized that mutations in *RIPK4* could explain a portion of PPS or VWS patients that have no detectable mutation in *IRF6*. We sequenced *RIPK4*in 27 VWS and 3 PPS patients without mutations in *IRF6*. We identified novel homozygous mutations in two PPS patients. The first, A448P, was present in a patient with bilateral cleft lip and palate, oral synechia, bilateral popliteal pterygia, and syndactyly. This patient, from consanguineous parents, also had a dorsal-ventral patterning defect of the hands and feet. The second mutation, R618H, was found in a patient with bilateral cleft lip and palate, lip pits, oral synechia, and popliteal pterygium from healthy, non-consanguineous parents. The mutation was carried by the mother, but not by the father. No deletion was detected in the patient or father by array CGH. SNP genotyping demonstrated maternal isodisomy of chromosome 21, where *RIPK4* is located. Our results show RIPK4 mutations may explain a significant portion of PPS cases that are *IRF6* mutation negative, demonstrate phenotypic overlap for mutations in the *RIPK4* gene, and have important implications for genetic counseling given the differing patterns of inheritance for PPS caused by *IRF6* as opposed to *RIPK4*. Finally, the results suggest a common biological pathway uniting *IRF6* and *RIPK4*. It remains possible that non-coding mutations in *RIPK4* or other members of this pathway could explain the remaining causes of VWS.

2763F

Sacsin-related ataxia caused by the novel missense mutation Arg272His in a patient from Southern-Italy. *F. Cavalcanti[1], A. Nicoletti[2], G. Annesi[1], P. Tarantino[1], M. Gagliardi[1,3], G. Mostile[2], V. Dibilio[2], A. Quattrone[4,5], A. Gambardella[1], M. Zappia[2].* 1) Inst. of Neurological Sciences, National Research Council, Mangone (CS), CS, Italy; 2) Department of Neurosciences, University of Catania, Catania, Italy; 3) University of Magna Graecia, Catanzaro, Italy; 4) Institute of Neurology, Department of Medical Sciences, University of Magna Graecia-Catanzaro, Italy; 5) Neuroimaging Research Unit, National Research Council, Catanzaro, Italy.

Autosomal recessive spastic ataxia of Charlevoix-Saguenay (ARSACS) is an early-onset cerebellar ataxia with spasticity, amyotrophy, nystagmus, dysarthria, and peripheral neuropathy. SACS is the only gene known to be associated with the ARSACS phenotype. The gene was initially reported to be encoded by a single gigantic exon. Recently eight additional exons were identified upstream of the giant exon. The SACS protein product, sacsin, is believed to integrate the ubiquitin-proteasome system and Hsp70 chaperone machimery. Recent studies indicates a role for sacsin in regulation of mitochondrial dynamics. To date, more than 70 different mutations, predominantly located in the giant exon, have been identified worldwide; of these only seven, exclusively located in the giant exon, were identified in Italian patients. The authors identified a novel homozygous variation c.815G>A, which results in the missense mutation p.Arg272His (R272H) in a patient from Southern-Italy. This mutation was present in heterozygosis in both unaffected parents as in one unaffected sibling and in several unaffected relatives. The phenotype of our patient closely resembled the classic phenotype. R272H missense mutation falls in the seventh coding exon of the SACS gene, hence not in the carboxyterminal giant coding exon which contains most reported sacsin mutations. R272H represents the first described Italian mutation located upstream of the giant exon of the SACS gene. The residue is completely conserved in evolution. The missense mutation R272H was not detected among 200 control chromosomes. Although it is also in a well-conserved region of the protein, no functional domain is defined for this region.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2764W

SCA36 molecular analysis in patients with spinocerebellar ataxia. *D. Di Bella, E. Sarto, C. Mariotti, C. Gellera, S. Magri, F. Taroni.* Genet Neurodeg & Metab Dis, IRCCS Istituto Neurologico C Besta, Milan, Italy.

Autosomal dominant spinocerebellar ataxias (SCA) are a heterogeneous group of neurological disorders characterized by cerebellar dysfunction mostly due to Purkinje cell degeneration. Recently, a novel form of spinocerebellar ataxia (SCA36) with motor neuron involvement was described in Japanese and Galician patients. It is caused by a GGCCTG repeat expansion in intron 1 of NOP56 gene. This gene encodes a component of the ribonucleoprotein complex and plays a role in transcription and splicing. We screened a large cohort of Italian unrelated patients with familiar (n=102) or sporadic (n=94) spinocerebellar ataxia for GGCCTG repeat expansion in NOP56 intron 1. All patients were negative for the common SCA1 and SCA2 mutations. The NOP56 repeat was analysed by fluorescent triplet repeat-primed PCR (TP-PCR) analysis, using three primers including one fluorescent-dye-conjugated forward primer, a first reverse primer consisting of 4 repeat units and a 5' anchor tail, and a second reverse anchor primer. NOP56 intron1 repeat expansion was detected in 4 probands from 4 different unrelated Italian families with dominant ataxia. No expansions were identified in sporadic patients. Normal alleles in the SCA population ranged from 6 to 14 repeats, with the 9-repeat allele being the most frequent. Clinically, mutated patients presented with late-onset (range 40–65 yrs), slow-progressing spinocerebellar ataxia characterized by pyramidal signs and eye movement abnormalities. Interestingly, all probands originated from a relatively small area in central Italy, suggesting a common ancestor and a founder effect for this mutation. In conclusion, SCA36 accounts for approximately 4% of Italian autosomal dominant SCA families negative for the common SCA mutations, a frequency lower than that (6.9%) observed in Galician patients but similar to that (3.6%) found in the Japanese population. Clinically, SCA36 is characterized by slow progressive gait ataxia with upper motor neuron involvement. Neuroimaging revealed prominent cerebellar atrophy affecting the vermis, with minor involvement of cerebellar hemispheres and brainstem [Supported in part by Telethon-Italia grant GGP09301].

2765T

Germline CAG repeat instability causes extreme CAG repeat expansion with infantile onset of spinocerebellar ataxia type 2 in offspring. *J. Ek[1], T. Vinther-Jensen[2,3], M. Dunoe[1], F. Skovby[1], L.E. Hjermind[2,3], J.E. Nielsen[2,3], T.T. Nielsen[2,3].* 1) Molecular Genetic Laboratory, Copenhagen University Hospital, Copenhagen, Denmark; 2) Memory Disorders Research Group, Neurogenetics Clinic, Department of Neurology, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark; 3) Department of Cellular and Molecular Medicine, Section of Neurogenetics, The Panum Institute, University of Copenhagen, Denmark.

The spinocerebellar ataxias (SCA) are a genetically and clinically heterogeneous group of diseases, characterized by dominant inheritance, progressive cerebellar ataxia and diverse extracerebellar symptoms. Mutations in 26 different genes are known to cause autosomal dominant inherited SCA, and a subgroup of the ataxias is caused by unstable CAG repeat expansions in their respective genes leading to pathogenic expansions of polyglutamine stretches in the encoded proteins. They belong to the CAG repeat disorders, which share a tendency of anticipation, implying that the CAG repeat is unstable and tends to expand further in successive generations resulting in earlier onset and more severe symptoms. In general, unstable CAG repeats have an uninterrupted CAG repeat whereas stable CAG repeats are either short or interrupted by CAA codons, which - like CAG codons - code for glutamine. Here we report a family in which infantile-onset SCA2 was diagnosed in a girl several years post-mortem, after presentation of the disease in her father. The girl presented with an atypical clinical phenotype with delayed motor milestones and epilepsy. She had myoclonic jerks and athetoid movements, and the only features indicative of a cerebellar disorder were uncoordinated eye movements and subtle findings on an MRI scan. Molecular genetic analysis revealed that the girl had a CAG repeat length in *ATXN2* of at least 124 repeats. Her father had a 45 CAG repeat allele. Surprisingly, the expanded allele of the father was an interrupted CAG repeat sequence. Furthermore, analyses of single spermatozoa from the father showed a high frequency of paternal germ line instability in the SCA2 locus. In conclusion, this study shows a high degree of germline instability in the SCA2 locus in spite of a CAA interruption in the CAG repeat. Our observations question the general assumption that instability is primarily limited to pure CAG repeats. Finally, this study highlights the importance of considering rare genetic disorders during diagnosis of infants and children even without known neurodegenerative disorders in the family.

2766F

Genetic Aspects of Dravet Syndrome in Southern Italy. *M. Gagliardi[1,2], P. Tarantino[1], F. Cavalcanti[1], M. Sesta[5], P. Conti[5], D. Galeone[5], G. Tortorella[3], A. Labate[4], M. Caracciolo[1], T. Mirante[1], A. Bagala[1], A. Quattrone[4], A. Gambardella[1], G. Annesi[1].* 1) Institute of Neurological Sciences, National Research Council, Mangone, Cosenza, Italy; 2) University Magna Graecia, Catanzaro, Italy; 3) Departement of Neurosciences, Psychiatry and Anaesthesiology University of Messina, Messina, Italy; 4) Institute of Neurology, University Magna Graecia Catanzaro and Neuroimaging Research Unit, National Research Council, Germaneto, Italy; 5) Departement of Neurosciences, Psychiatry and Anaesthesiolog Department of Neurology, Pediatric Hospital, Bari, Italy.

Dravet syndrome (DS), or severe myoclonic epilepsy of infancy (SMEI) is a genetically determined encephalopathy mainly caused by de novo mutations in the SCN1A gene. A small percentage of female patients with a DS-like phenotype might carry in the gene encoding protocadherin-19 (PCDH19) who were negative for mutations in the SCN1A gene. The aim of this study was to investigate the frequency of SCN1A mutations in 70 SMEI patients with epileptic encephalophathy of infancy (EEI) from southern Italy and to investigate the frequency of PCDH19 mutations in female patients who resulted negative for SCN1A mutations. All patients had seizures onset before 12 months of age, mild to severe mental retardation with poor language development, with or without ataxia. Genomic DNA from the patients was analysed by direct sequencing of the SCN1A and PCDH19 genes on an ABI 3130XL Avant automated sequencer. The genomic anomalies of the SCN1A gene were screened using MLPA and confirmed by real-time PCR. We identified 13 different heterozygous mutations of SCN1A gene in 13/70 patients with SMEI (18,57%) (Lys1246fsX1268, Phe807Leu, Arg1636X, p.1502del, Glu1021X, IV242A-G, Arg1886Gly, Thr1289Ile, IVS7+4delA, 3840insT, Ser1505X, del. ex 1–25, Tyr1460X). Thirty-one/70 female with genetic test negative for SCN1A mutations were analyzed for the PCDH19 gene. This analysis revealed three different heterozygous novel mutations (9,67%) (Ile508ProfsX59, Arg550Pro, Ser856Cys). We found 13 patients carrying different SCN1A mutations, thus confirming the high genetic heterogeneity related to DS. The proportion of patients with SCN1A mutations in our population is 18,57%. Previous studies a higly variable rate of SCN1A mutations, ranging from 30% to 80%. These discrepancies may be due to the sizes of the series, the use of different clinical criteria, or the inclusion of broader epileptic phenotypes. Also the results of this study indicate that PCDH19 mutations are a relatively frequent cause of EEI in Southern Italy. This frequency (9,67%) is comparable to that reported in previous studies.

2767W

Vacuolar-type H+-ATPase V1A subunit is a molecular partner of Wolfram syndrome 1 (WFS1) protein, which regulates its stability and expression. *S. Gharanei[1], M. Zatyka[1], D. Astuti[1], J. Fenton[1], E. Rabai[2], A. Sik[3], Z. Nagy[2], T. Barrett[1].* 1) School of Clinical and Experimental Medicine, University of Birmingham, Birmingham, United Kingdom; 2) Department of Pharmacology, School of Clinical and Experimental Medicine, University of Birmingham, Birmingham, United Kingdom; 3) Department of Neurophysiology, School of Clinical and Experimental Medicine, University of Birmingham, Birmingham, United Kingdom.

Wolfram syndrome is an autosomal recessive disorder characterised by neurodegeneration and diabetes mellitus. The gene responsible for the syndrome (WFS1) encodes an endoplasmic reticulum (ER) resident transmembrane protein, that also localises to secretory granules in pancreatic beta cells. Although its precise functions are unknown, WFS1 protein deficiency affects the unfolded protein response, intracellular ion homeostasis, cell cycle progression, and granular acidification. In this study, immunofluorescent and electron-microscopy analyses confirmed that WFS1 also localizes to secretory granules in human neuroblastoma cells. We demonstrated a novel interaction between WFS1 and the V1A subunit of the vacuolar-type H+-ATPase (proton pump) by co-immunoprecipitation in HEK293 cells, and with endogenous protein in human neuroblastoma cells. We mapped the interaction to the WFS1-N terminal, but not the C-terminal domain. V1A subunit expression was reduced in WFS1 stably and transiently depleted human neuroblastoma cells and depleted NT2 cells. This reduced expression was not restored by adenoviral over-expression of BiP to correct the ER stress. Protein stability assays demonstrated that the V1A subunit was degraded more rapidly in WFS1 depleted neuroblastoma cells compared with wild type. We conclude that WFS1 has a specific interaction with the V1A subunit of H+-ATPase; this interaction may be important both for pump assembly in the ER; and for granular acidification.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2768T

Altered expression of carbonic anhydrase-related protein XI in neuronal cells expressing mutant ataxin-3. *M. Hsieh[1], W.H. Chang[1], C.F. Hsu[1], C.L. Kuo[2], I. Nishimori[3], T. Minakuchi[3], N. Nukina[4].* 1) Department of Life Science, Tunghai University, Taichung 407, Taiwan, Republic of China; 2) Department of Biomedical Science, Chung-Shan Medical University, Taichung 402, Taiwan, Republic of China; 3) Department of Gastroenterology and Hepatology, Kochi Medical School, Nankoku, Kochi 783-8505, Japan; 4) Laboratory for Structural Neuropathology, RIKEN Brain Science Institute, RIKEN, Wako-shi, Japan.

Spinocerebellar ataxia type 3/Machado-Joseph disease (SCA3/MJD) is a late-onset neurodegenerative disorder caused by the expansion of a polyglutamine tract within the gene product, ataxin-3. Microarray analysis revealed a dramatic differential expression of carbonic anhydrase-related protein XI (CA-RPXI/CA11) in the presence or absence of mutant ataxin-3. Therefore, we examined the expression and distribution of all three CA-RPs (CA8, 10 and 11) in human neuronal cells stably expressing mutant ataxin-3. The protein expression of CA8 and CA11 is significantly increased in human neuroblastoma cells harboring mutant ataxin-3, compared to cells containing normal ataxin-3. Semi-quantitative RT-PCR demonstrated that all three CA-RPs exhibited significantly higher transcript levels in neuronal cells expressing mutant ataxin-3. Interestingly, CA11 is distributed not only in the cytoplasm but also translocated within the nuclei of the stably transfected mutant cells, compared to the sole cytoplasmic distribution in cells containing normal ataxin-3. In addition, results from transient transfection assays in SK-N-SH and Neuro2a cells also confirmed the nuclear localization of CA11 in the presence of truncated ataxin-3. Most importantly, immunohistochemical staining of the MJD transgenic mice and postmortem MJD human brain also revealed that CA11 highly expresses in both cytoplasm and nuclei of the brain cells. Recruitment of CA11 into nuclear inclusions containing mutant ataxin-3 revealed a possible correlation between CA11 and disease progression. Although the exact function of CA-RPs is still undefined in the central nervous system, our findings suggest that CA-RPs, especially CA11, may play specific roles in the pathogenesis of Machado-Joseph disease.

2769F

Alterations in RNA Splicing Underlie Development of Ataxia in SCA1 Mice. *M.A.C. Ingram[1], P. Bergmann[2], H. Kang[3], H.Y. Zoghbi[3,4], H.T. Orr[1].* 1) Institute for Translational Neuroscience, University of Minnesota, Minneapolis, MN; 2) Foresight Logic, Inc. St. Paul, MN; 3) Department of Molecular and Human Genetics, and; 4) Department of Pediatrics, and the Howard Hughes Medical Institute, Baylor College of Medicine, Houston, TX.

SCA1, a late onset neurodegenerative disease, is caused by a CAG polyglutamine expansion in the Ataxin-1 (ATXN1) protein. Evidence indicates that ATXN1 normally interacts with regulators of transcription and RNA splicing. Previous work showed that domains other than the polyQ tract, notably a serine residue at site 776, mediate mutant ATXN1 toxicity. A wild type ATXN1[30Q] with phosphomimetic Asp at 776 triggers disease initiation and development of ataxia as seen in mice expressing mutant ATXN1[82Q]. Towards identification of cellular pathways critical for development of SCA1 we performed RNA-seq to examine gene expression and RNA splicing patterns in ATXN1[82Q], ATXN1[30Q]-D776, and wild type FVB cerebella at 12 weeks of age. Gene expression analysis indicated 2500–4000 of significant gene changes (q-value <0.05) occur in ATXN1[82Q] and ATXN1[30Q]-D776 in reference to FVB, 2240 significant gene changes when compared to each other. However, while 5000–6000 significant splicing changes are identified in comparison to FVB, only one splicing change is identified between ATXN1[82Q] and ATXN1[30Q]-D776. Additionally, previous biochemical evidence showed ATXN1[30Q]-D776 increases interaction with the splicing factor RBM17 to the level of ATXN1[82Q]. We propose, ATXN1[82Q] and ATXN1[30Q]-D776 interact with RBM17 to the same extent, resulting in a splicing pattern that leads to the development of ataxia in SCA1.

2770W

Clinical spectrum of intragenic CAMTA1 rearrangements: From non-progressive congenital ataxia to intellectual disability. *E. Lopez[1], J. Thevenon[1,2], B. Keren[3,4], D. Heron[5], C. Mignot[5], C. Altuzara[6], M. Béri-Dexheimer[7], C. Bonnet[7], L. Burglen[5,8], D. Minot[2], J. Vigneron[9], P. Charles[5], A. Brice[3,4], L. Gallagher[10], J. Amiel[11], C. Mach[3], C. Depienne[3,4], D. Doummar[5], L. Duplomb[1], V. Carmignac[1], P. Callier[1,12], A.L Mosca-Boidron[1,12], V. Roze[1,13], B. Aral[1,14], F. Razavi[15], P. Jonveaux[7], L. Faivre[1,2], C. Thauvin-Robinet[1,2].* 1) Equipe Génétique et Anomalies du Développement, Université de Bourgogne, Faculté, Dijon, France; 2) Centre de Génétique et Centre de Référence (Anomalies du Développement et Syndromes Malformatifs), Hôpital d'Enfants, CHU Dijon, 21033 Dijon, France; 3) INSERM, U975 and CNRS 7225, CRICM, Hôpital Pitié-Salpêtrière, 75 013 Paris, France; 4) AP-HP, Département de Génétique et de Cytogénétique, Centre de Génétique Moléculaire et Chromosomique, Hôpital Pitié-Salpêtrière, 75 013 Paris, France; 5) Département de Génétique et Centre de Référence "Déficiences intellectuelles de causes rares", Groupe Hospitalier Pitié-Salpêtrière et CRICM, INSERM UMRS975, La Pitié Salpêtrière, 75 013 Paris, France; 6) Service de Neuropédiatrie, CHU de Besançon, 25000 Besançon, France; 7) Laboratoire de Génétique et EA 4368, CHU de Nancy, 54035 Nancy, France; 8) Centre de référence des malformations et maladies congénitales du cervelet et Service de Génétique Médicale, APHP, Hôpital Trousseau, et INSERM U676, Hôpital Robert Debré, 75019 Paris, France; 9) U.F. de Génétique Clinique, Maternité Régionale, 54035 Nancy, France; 10) Department of Psychiatry, Trinity College Dublin, Dublin 2, Irelan; 11) Département de Génétique, Hôpital Necker Enfant Malade, AP-HP et Université René Descartes, 76004 Paris, France; 12) Laboratoire de Cytogénétique, Plateau Technique de Biologie, CHU de Dijon, 21033 Dijon, France; 13) Laboratoire de Cytogénétique, CHU de Besançon, 25000 Besançon, France; 14) Laboratoire de Génétique Moléculaire, Plateau Technique de Biologie, CHU de Dijon, 21033 Dijon, France; 15) Service d'Histologie-Embryologie-Cytogénétique, Hôpital Necker-Enfants Malades, 149 rue de Sèvres, 75015 Paris, France.

Non-progressive congenital ataxias (NPCA) with or without intellectual disability (ID) are clinically and genetically heterogeneous conditions. As a consequence, the identification of the genes responsible for these phenotypes remained limited. Using high resolution microarrays, we identified intragenic copy number variations in the CG-1 domain of the Calmodulin-binding Transcription Activator 1 (CAMTA1) gene, segregating with autosomal dominant ID with NPCA in two unrelated families, and a de novo deletion located in the same domain in a child presenting with NPCA. In the ID patients, the deletion led to a frameshift, producing a truncated protein, while the deletion was in frame in the patient with isolated childhood ataxia. Brain MRI of the patients revealed a pattern of progressive atrophy of cerebellum medium lobes and superior vermis, parietal lobes and hippocampi. Although DNA sequencing of the CG-1 domain in 197 patients with sporadic or familial non-syndromic intellectual deficiency, extended to full DNA sequencing in 50 patients with ID and 47 additional patients with childhood ataxia, identified no pathogenic mutation, there is considerable evidence that CAMTA1 rearrangements are being disease causing. Indeed, CAMTA1 is a brain specific calcium responsive transcription factor expressed in the brain and cerebellum during development and later implicated in memory processes; intragenic rearrangements are concentrated in a highly conserved functional domain with transcription regulation ability and nuclear trafficking functions; CAMTA1 transcriptional studies have shown upregulation of genes already implicated in intellectual disability and autism spectrum disorder. Altogether, we show that CAMTA1 loss-of-function is responsible for NPCA with or without ID.

2771T

Analysis of Human Mitochondrial Aminoacyl-tRNA Synthetase complexes in patients with MARS2 mutations. *I. Thiffault, F. Sasarman, T. Nishimura, E.A. Shoubridge.* Lab Neurogenetics, 676, McGill Montreal Neurological Hosp, Montreal, PQ, Canada.

In the past few years mutations in at least 10 of the human mitochondrial aminoacyl-tRNA synthetases (ARSs) have been reported. Mutations in these ARSs are associated with a strikingly broad range of gene-specific clinical phenotypes, the molecular basis for which remains unknown. Many of the cytoplasmic ARSs are organized into a supramolecular complex that is coordinated by two proteins not directly involved in translation; but little is known about the organization of the mitochondrial ARSs. Here, we show that at least some of the mitochondrial ARSs exist in high molecular weight complexes of ~250 kDa and 1 MDa in human fibroblast cell lines. We previously reported that mutations in MARS2 (methionyl-tRNA synthetase 2) result in decreased mitochondrial translation in ARSAL subjects (SPAX3 OMIM ID: 611390). We show here that complexes containing MARS2 do not co-sediment with mitochondrial ribosomes, and that MARS2 mutations result in an altered association of MARS2 with the mitochondrial ARSs complexes. This study suggests that like some of the cytoplasmic ARSs, mitochondrial ARSs occur in high-molecular-weight complexes, and that these associations might be important for the regulation of mitochondrial translation.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2772F

Mutations in *GPSM2* Cause the Brain Malformations and Hearing Loss of Chudley-McCullough Syndrome. *D. Doherty[1], A.E. Chudley[2], G. Coghlan[3], G.E. Ishak[4], A.M. Innes[5], E.G. Lemire[6], R.C. Rogers[7], A.A. Mhanni[2], I.G. Phelps[1], S.J.M. Jones[8], S.H. Zhan[8], A.P. Fejes[8], H. Shahin[9], M. Kanaan[9], H. Akay[10], M. Tekin[11,12], B. Triggs-Raine[2], T. Zelinski[2,3] FORGE Canada Consortium.* 1) Department of Pediatrics, University of Washington, Seattle Children's Hospital, Seattle, Washington, USA; 2) Department of Pediatrics and Child Health and Department of Biochemistry and Medical Genetics, University of Manitoba, Winnipeg, Manitoba, Canada; 3) Rh Laboratory, Department of Pediatrics and Child Health, University of Manitoba, Winnipeg, Manitoba, Canada; 4) Department of Radiology, University of Washington, Seattle Children's Hospital, Seattle, Washington, USA; 5) Department of Medical Genetics, University of Calgary, Calgary, Alberta, Canada; 6) Department of Pediatrics, University of Saskatchewan, Saskatoon, Saskatchewan, Canada; 7) Greenwood Genetic Center, Greenville, South Carolina, USA; 8) Canada's Michael Smith Genome Sciences Centre, BC Cancer Agency, Vancouver, British Columbia, Canada; 9) Department of Life Sciences, Bethlehem University, Bethlehem, Palestine; 10) Memorial Hospital, Diyarbakir, Turkey; 11) Dr. John T. Macdonald Foundation, Department of Human Genetics and John P. Hussman Institute for Human Genomics, Miller School of Medicine, University of Miami, Miami, FL USA; 12) Division of Pediatric Genetics, Ankara University School of Medicine, Ankara, Turkey.

Autosomal recessive inheritance, severe to profound sensorineural hearing loss and partial agenesis of the corpus callosum are hallmarks of the clinically well-established Chudley-McCullough syndrome (CMS). Although not always reported in the literature, frontal polymicrogyria and gray matter heterotopia are uniformly present, while cerebellar dysplasia, ventriculomegaly and arachnoid cysts are nearly invariant. Despite these striking brain malformations, individuals with CMS do not in general present with significant neurodevelopmental abnormalities, except for hearing loss. Using a combined homozygosity mapping and whole exome sequencing approach in affected individuals from eight families (including the first-reported individuals), we identified four molecular variations (two single base deletions, a nonsense mutation and a canonical splice site mutation) in the G protein signaling modulator 2 gene, *GPSM2*, that underlie CMS. Mutations in *GPSM2* have been previously identified in people with profound, congenital, non-syndromic hearing loss. Subsequent brain imaging of these individuals revealed frontal polymicrogyria, abnormal corpus callosum and gray matter heterotopia, consistent with a CMS diagnosis; however, no ventriculomegaly was observed. The gene product, GPSM2, is required for orienting the mitotic spindle during cell division in multiple tissues, suggesting that the sensorineural hearing loss and characteristic brain malformations of CMS are due to defects in asymmetric cell divisions during development. Ongoing work includes *GPSM2* sequencing and detailed characterization of the neuroimaging findings in additional patients with CMS, as well as protein quantification and localization in patient fibroblasts.

2773W

Thrombospondin genes are required for development of normal cochlear function and maintenance. *M. Mustapha, D. Mendus, V. Sundaresan.* Otolaryngology-HNS, Stanford, Stanford, CA.

Recent studies reveal important roles for cell-adhesion molecules such as thrombospondins (TSPs) in promoting synapse formation during brain development and in repairing synaptic connection after stroke. For our efforts to identify genes involved in cochlear synapse formation and functional maturation, we hypothesize that cell adhesion proteins that are also upregulated during cochlear synaptogenesis and/or adulthood are important for these processes. In this study, we examine whether TSPs are involved in synapse formation and maintenance in the cochlea using recombinant mouse models of these genes. Our preliminary data using microarray and qPCR analysis on TSP genes 1–4 from whole cochlea revealed significant increase in TSP2–4 expression during a critical window of postnatal cochlear development with TSP1 expression picking up later in adulthood. Our auditory brain stem response (ABR) test on TSP1 and TSP2 mutants reveal elevated threshold in TSP2 mutants but not in TSP1 mutants at an early age. These data suggest a role for TSP2 in synapse functional maturation that is important for the onset of hearing. However, one-year old TSP1 mutant mice show an increase in ABR threshold at higher frequencies and decreased amplitude of response compared to wild type controls, suggesting a role for TSP1 in synapse maintenance. Our physiology data is in agreement with the roles predicated for these genes based on our expression studies. Counting of afferent and efferent synapse numbers in TSP2 and TSP1/2 mutants versus wild type mice does not show significant differences between these groups. Our next step would be to examine whether the synapses in the TSP mutant mice are functional using electrophysiology. To test the role for these genes in the recovery of the cochlea from noise injury, we exposed TSP2 and TSP1/2 mutant mice to a noise insult and measured ABR threshold after a recovery period. Data from these experiments showed that these mice were unable to recover from the noise injury compared to wild type controls indicating a deficit in repair mechanisms in these mice. Our results, taken together, indicate that TSPs 1 and 2 are both involved in different aspects of cochlear functional maturation, repair and maintenance and these genes are now candidates for screening in human patients with congenital and age-related hearing impairment.

2774T

Lack of the VPS35 Asp620Asn mutation in southern Italian patients with familial Parkinson's disease. *G. Annesi[1], M. Gagliardi[1, 2], P. Tarantino[1], F. Cavalcanti[1], T. Mirante[1], M. Caracciolo[1], A. Bagalà[1], A. Gambardella[1], A. Quattrone[3].* 1) Inst of Neurological Sciences, National Research Council, Mangone, CS, Italy; 2) University Magna Graecia,(CZ), Italy; 3) Institute of Neurology, University Magna Graecia, Catanzaro, Italy Neuroimaging Research Unit, Institute of Neurological Sciences, National Research Council, Germaneto (CZ).

Parkinson's disease (PD) is a common neurodegenerative disorder, affecting 2% of those over the age of 75 years. Although generally considered a sporadic disease, Mendelians forms of the disease are described (SNCA and LRRK2 for causing autosomal dominant PD and 3 genes causing autosomal recessive, juvenile PINK1, Parkin, DJ1). Recently using an exome sequencing based approach, 2 indipendent groups have identified a missense mutation in vacuolar protein sorting 35 homolog (VPS35 c.1858G>A; p.Asp620Asn) as the probable cause of late onset PD in a number of kindreds. To estimate the frequency of the Asp620Asn mutation in VPS35 in familial PD, we screened for this variant in a southern Italy PD cohort. Our population included 114 patients with familial PD, having at least 1 relative among their first degree, second degree, in third degree family members with a formal diagnosis of PD (major PD genes had been analyzed and positive cases (9 LRRK2) were not excluded. Genomic DNA was extracted from peripheral blood by standard method. The genotype was determined by genomic DNA amplification of a 229 bps fragment and sequencing was done on an ABI PRISM 3130 XL-AVANT Genetic Analyzer. This variant was not detected in any of the 150 analyzed familial PD cases, thus indicating that this mutation is rare among familial PD cases in the southern Italy. It would seem reasonable to conclude, therefore, that the recently published VPS35 mutation, is not a common cause of familial PD, in our population.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2775F

Ubiquilin 2 mutations in Italian patients with amyotrophic lateral sclerosis and frontotemporal dementia. *C. Gellera[1], C. Tiloca[2,3], R. Del Bo[4,5], L. Corrado[6], V. Pensato[1], j. Agostini[7], C. Cereda[8], A. Ratti[2,5], B. Castellotti[1], S. Corti[4,5], A. Bagarotti[6], A. Cagnin[7], P. Milani[8,9], C. Gabelli[10], G. Riboldi[4,5], L. Mazzini[11], G. Sorarù[7], S. D'Alfonso[6], F. Taroni[1], G.P. Comi[4,5], N. Ticozzi[2], V. Silani[2,5].* 1) Unit of Genetics of Neurodegenerative and Metabolic Diseases, IRCCS - Neurologico Carlo Besta, Milano, Italy; 2) Department of Neurology and Laboratory of Neuroscience, IRCCS Istituto Auxologico Italiano, Milan, Italy; 3) Doctoral School in Molecular Medicine, Department of Sciences and Biomedical Technologies, Università degli Studi di Milano, Italy; 4) IRCCS Foundation Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy; 5) Department of Neurological Sciences, "Dino Ferrari" Center, Università degli Studi di Milano, Milan, Italy; 6) Department of Health Sciences, Interdisciplinary Research Center of Autoimmune Diseases, `A. Avogadro' University, Novara, Italy; 7) Department of Neuroscience, University of Padua, Padua, Italy; 8) Laboratory of Experimental Neurobiology, IRCCS National Neurological Institute `C. Mondino,' Pavia, Italy; 9) Department of Neurological Sciences, University of Pavia, Pavia, Italy; 10) Research Consortium `Luigi Amaducci' CRIC, Arcugnano (Vicenza), Italy; 11) Department of Neurology, A. Avogadro University and Maggiore della Carita` Hospital, Novara, Italy.

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disease mainly involving cortical and spinal motor neurons. Recently, molecular studies identified different mutations in ubiquilin-2 (UBQLN2) gene as causative of a familial form of X-linked ALS, 90% penetrant in females. Aim of our study was to analyze the UBQLN2 gene in a large cohort of familial and sporadic ALS patients, with or without frontotemporal dementia (FTD), and in patients with FTD only. We analyzed the UBQLN2 gene in 819 sporadic ALS cases (SALS), 226 familial ALS cases (FALS), 53 ALS-FTD patients, and 63 patients with a clinical record of FTD. Molecular analysis of the entire coding sequence was carried out in all FALS and ALS-FTD patients, while SALS and FTD patients were analyzed specifically for the genomic region coding for the PXX repeat tract. Healthy controls were 845 anonymous blood donors and were screened for the PXX repeat region only. We found five different variants in the UBQLN2 gene in five unrelated ALS patients. Three of them, including two novel ones, involved a proline residue in the PXX repeat region, and were found in three FALS cases. The other two were novel variants, identified in one FALS and one SALS patient. None of these variants were present in controls, while one control carried a new heterozygous variant. Our data support the role of UBQLN2 gene in the pathogenesis of FALS, being conversely a rare genetic cause in SALS even when complicated by FTD.

2776W

Genetics of Primary Microcephaly in Indian families. *A. Kumar[1], V. Bhat[1], G. Mohan[1], S.C. Girimaji[2].* 1) Molecular Reproduction, Development & Genetics, Indian Institute of Science, Bangalore, Karnataka, India; 2) Department of Child and Adolescent Psychiatry, National Institute of Mental Health and Neuro Sciences, Bangalore, Karnataka, India.

The development of the human brain, specially the cerebral cortex, is an intricate and complex process that involves generation of neuronal progenitors in periventricular zones, neuronal cell proliferation through symmetric and asymmetric cell divisions, and finally the migration of post-mitotic neurons to their final destinations. An understanding of the molecular mechanisms that guide these intricate and complex processes is facilitated by discovery of genes for Mendelian forms of cortical (brain) malformations, such as primary microcephaly (MCPH), which is characterized by a smaller than normal brain and mental retardation. Occasionally, MCPH patients also have other brain abnormalities such as pachygyria, polymicrogyria, simplified gyral pattern etc. We have been working on the genetics of MCPH for some time, which led to the discovery of a novel seventh gene for MCPH, STIL, in 2009. The other six MCPH genes are MCPH1, WDR62, CDK5RAP2, CEP152, ASPM and CENPJ. In May 2012, the eighth gene, CEP135, has been identified. We have continued to ascertain MCPH families and collected so far a total of 56 families. Interestingly, Sanger sequencing has shown mutations only in WDR62, ASPM and STIL genes in our families. The ASPM gene is found to be the most commonly mutated gene followed by STIL in Indian population. There are still many families in our dataset which do not harbor mutations in any of the eight known genes, suggesting the discovery of additional novel genes in future. We are currently using homozygosity mapping and whole-exome sequencing to identify the novel gene(s). A comprehensive analysis of our genetic analysis of Indian MCPH families will be presented and discussed.

2777T

ZIC1 AND ZIC4 DELETION AT 3q24 IS NOT CONSTANTLY ASSOCIATED TO DANDY-WALKER MALFORMATION (DWM). *S. Loddo[1], V. Parisi[1], L. Bernardini[1], A. Ferraris[1], A. Capalbo[1], L. Travaglini[2], G. Zanni[2], M.L. Di Sabato[3], S. Tumini[4], A. Novelli[1], V. Sabolic[5], E. Bertini[2], B. Dallapiccola[2], E.M. Valente[1].* 1) Cytogenetics Lab, CSS-Mendel Inst, Rome, Italy; 2) Bambino Gesù Children Hospital Rome-Italy; 3) Sapienza University of Rome Department of Child Neurology and Psychiatry Rome-Italy; 4) University of Chieti Department of Pediatrics Chieti-Italy; 5) St. Cyril and Methodius University Skopje-Macedonia.

The Dandy-Walker malformation (DWM) is one of the most frequent form of congenital cerebellar hypoplasia and is characterized by upward rotation of the cerebellar vermis and cystic dilatation of the fourth ventricle. It 's a heterogeneous condition, which may be associated with different cytogenetic abnormalities and malformation syndromes. The first genes implicated in its pathogenesis, ZIC1 and ZIC4, are to date deleted in a limited number of patients. We have recently identified by SNP-array 3 new cases of interstitial deletion 3q. Two patients, recruited within the project focused on birth defects of the cerebellum and brainstem (CBCD Project) with a diagnosis of classical DWM, show, respectively, the del3q23q26.1 and the del3q22.3q25.32 (~ 20Mb). In contrast, in the third patient a del3q23q24 (about 4.9 Mb) was detected, but MRI ruled out with certainty the DWM. The minimal deleted region shared by the three patients included the genes ZIC1 and ZIC4, while the two patients with DWM shared two additional regions: one proximal (0.5MB) and one distal (10 Mb) to candidate genes. The clinical phenotype of the patients was variable, with psychomotor retardation, facial dysmorphism and anomalies of extremities. This study confirms the usefulness of the search for chromosomal microrearrangements, especially in cases of syndromic DWM. The absence of cerebellar defects in the third patient suggests that hemizygosity of genes ZIC1 and ZIC4 is not sufficient for the onset of DWM and that there are other factors that may modulate phenotypic expression.

2778F

A frameshift mutation in the gene encoding folliculin-interacting protein 2 (FNIP2) associated with hypomyelination and tremors in the Weimaraner. *P. Patel[1], T. Pemberton[2], S. Choi[1], J. Mayer[3], F-Y. Li[1], N. Safra[4], D. Bannasch[4], K. Sullivan[5], B. Breuhaus[6], I. Duncan[3].* 1) Inst Gen Med, Univ Southern California, Los Angeles, CA; 2) Department of Biology, Stanford University, Stanford, CA 94305; 3) Department of Medical Sciences, School of Veterinary Medicine, University of Wisconsin - Madison, Madison, WI 53711; 4) Department of Population Health and Reproduction, School of Veterinary Medicine, University of California Davis, Davis, CA 95616; 5) Department of Medicine, University of Washington, Seattle, WA 981956; 6) College of Veterinary Science, North Carolina State University, Raleigh, NC 27606.

Proper myelination requires intricate interactions between glial cells and neurons. An important approach to unraveling the factors involved in these interactions and their respective roles in the process is the study of naturally occurring mutants in whom the process has gone awry. An autosomal recessive disorder characterized by hypomyelination and tremors was first described in 1987 in the Weimaraner breed. It is a widespread disorder with affected animals reported from throughout the US and also from Europe. At 10–12 days of age, a tremor involving the trunk and all four limbs is noted that worsens on ambulation. The degree of tremor may vary between siblings in a single litter; likewise the recovery is seen more quickly in less affected dogs, usually by 3–4 months while some severely affected dogs may retain a mild tremor for life. On necropsy, a unique zone of non- or hypomyelination is seen around the periphery of the spinal cord, especially of the ventral and lateral columns. In the brain, there is a more subtle but generalized decrease in myelin compared to controls. The cause of this myelin defect was heretofore unknown but it had been speculated to involve the early differentiation of oligodendrocytes, their migration throughout developing white matter or their early death. Towards positional cloning of the genetic defect leading to this disorder, we genotyped 48 animals : 35 from 3 unrelated pedigrees - 9 affected animals, 18 known carriers, and 8 phenotypically normal ("unaffected") animals who may or may not be carriers - and 13 unrelated animals ("singletons") - five affected animals, one known carrier, and seven phenotypically normal animals. Genome-wide association mapping localized the gene to a 5 Mb interval on chromosome 15. Homozygosity mapping further narrowed the interval to a 3.57 Mb region. Seventeen genes were examined for mutations and a deletion of a single A residue was identified in the gene encoding folliculin-interacting protein 2 (FNIP2). All 14 affected dogs were homozygous for the mutation, 8 obligate and 15 predicted carriers were heterozygous for the mutation, and 4 unaffected animals from the two pedigrees as well as 7/8 unrelated "control" dogs did not bear the mutation. One unrelated "control" dog was found to be heterozygous for the mutation. Further work is underway to determine the expression pattern of the gene and it relationship to the disease phenotype.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2779W

Familial cortical myoclonus with a mutation in NOL3. *J. F. Russell[1], J. L. Steckley[2], G. Coppola[3], A. F. G. Hahn[2], M. A. Howard[1], Z. Kornberg[1], A. Huang[3], S. M. Mirsattari[2], B. Merriman[3], E. Klein[3], M. Choi[4], H.-Y. Lee[1], A. Kirk[5], C. Nelson-Williams[4], G. Gibson[6], S. C. Baraban[1], R. P. Lifton[4], D. H. Geschwind[3], Y.-H. Fu[1], L. J. Ptacek[1].* 1) UCSF, San Francisco, CA; 2) London Health Sciences Center and University of Western Ontario, London, Ontario, Canada; 3) UCLA, Los Angeles, CA; 4) Yale, New Haven, CT; 5) Royal University Hospital, Saskatoon, SK, Canada; 6) University of British Columbia, Vancouver, BC, Canada.

Objective: Myoclonus is characterized by sudden, brief involuntary movements and its presence is debilitating. We identified a family suffering from adult-onset, cortical myoclonus without associated seizures. We performed clinical, electrophysiological, and genetic studies to define this phenotype. **Methods**: A large, four-generation family with history of myoclonus underwent careful questioning, examination, and electrophysiological testing. Thirty-five family members donated blood samples for genetic analysis, which included SNP mapping, microsatellite linkage, targeted massively parallel sequencing, and Sanger sequencing. In silico and in vitro experiments were performed to investigate functional significance of the mutation. **Results**: We identified 11 members of a Canadian Mennonite family suffering from adult-onset, slowly progressive, disabling, multifocal myoclonus. Somatosensory evoked potentials indicated a cortical origin of the myoclonus. There were no associated seizures. Some severely affected individuals developed signs of progressive cerebellar ataxia of variable severity late in the course of their illness. The phenotype was inherited in an autosomal dominant fashion. We demonstrated linkage to chromosome 16q21-22.1. We then sequenced all coding sequence in the critical region, identifying only a single co-segregating, novel, nonsynonymous mutation, which resides in the gene *NOL3*. Furthermore, this mutation was found to alter post-translational modification of NOL3 protein in vitro. **Interpretation**: We propose that Familial Cortical Myoclonus (FCM) is a novel movement disorder caused by mutation in *NOL3*. Further investigation of the role of *NOL3* in neuronal physiology may shed light on neuronal membrane hyperexcitability and pathophysiology of myoclonus and related disorders.

2780T

Mutations in PRRT2 result in familial infantile convulsions with marked variability in clinical expression and SUDEP. *P. Tarantino[1], G. Palamara[2], A. Labate[2,3], M. Gagliardi[1], L. Mumoli[2], E. Ferlazzo[2], F. Cavalcanti[1], U. Aguglia[2], G. Incorpora[4], A. Gambardella[1,2], G. Annesi[1].* 1) Institute of Neurological Sciences, National Research Council, Mangone, cosenza, Italy; 2) Institute of Neurology University Magna Græcia, Catanzaro, Italy; 3) Neuroimaging Research Unit, National Research Council, Germaneto, Catanzaro, Italy; 4) Unit of Paediatrics, University Hospital "Arezzo-Trifiletti", Ragusa, Italy.

Benign familial infantile seizures (BFIS) is an autosomal dominant disorder characterized by afebrile seizures that begin at age 3–12 months with a favorable outcome. A recent discovery has identified heterozygous mutations in PRRT2 on chromosome 16p, which encodes proline-rich transmembrane protein 2, in most families affected by BFIS. PRRT2 is also the major causative gene for familial paroxysmal kinesigenic dystonia (PKD), a rare disorder characterized by episodic attacks of choreoathetosis or dystonia. These results have corroborated the existence of familial infantile convulsions with paroxysmal choreoathetosis (ICCA, OMIM 602066) that shares overlapping clinical features with BFIS and PKD. In this study, we performed mutation screening of PRRT2 in our collected families with BFIS, ICCA or PKD phenotypes.The whole genomic region of PRRT2 was sequenced in seven Italian families, of which six with BFIS or ICCA phenotype, and one family with PKD phenotype. The previously reported mutation, c.649dupC (p.R217Pfs*7), was found in two families with BFIS phenotype, and in one family with ICCA. In an additional BFIS family, a missense mutation, c.718C/T (R240X), was identified. All these mutations co-segregated with the disease in these families and were not observed in 300 control subjects of matched ancestry. In the ICCA family, two affected members displayed a more complex phenotype with episodic ataxia, mental retardation and migraine attacks. In one family that also carried the c.649dupC mutation, one affected member died at age of 13 years of SUDEP. This study confirms the major role of PPRT2 mutations in families with BFIS phenotypes. Our findings also enlarge the clinical spectrum related to PRRT2 mutations and underscore the complexity of the phenotypic consequences of mutations in this gene.

2781F

Molecular analysis of *NIPA1* gene mutations in patients with spastic paraplegia. *F. Taroni[1], S. Caldarazzo[1], E. Sarto[1], M. Plumari[1], V. Martinelli[2], D. Pareyson[3], S. Baratta[1], D. Di Bella[1].* 1) Unit of Genetics of Neurodegenerative and Metabolic Diseases, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milan, Italy; 2) Institute of Experimental Neurology, Fondazione Centro San Raffaele del Monte Tabor, Milan, Italy; 3) Unit of Neurology 8, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milan, Italy.

Hereditary spastic paraplegias (HSP) are a clinically and genetically heterogeneous group of neurodegenerative disorders. Pure and complicated forms of the disease have been described. About half of HSP cases result from autosomal dominant mutations in *spastin* (SPG4), *atlastin-1* (SPG3A), or *REEP1* (SPG31) genes. Mutations in the *NIPA1* gene have been reported to be a rare cause of autosomal dominant (AD) pure (10 families) or complicated (2 families) HSP phenotypes (SPG6). Recent data indicate that the HSP-associated proteins NIPA1 and atlastin-1 are members of a common biochemical pathway that supports axonal maintenance. In 9 of the 12 SPG6 families reported to date, two recurrent mutations at a mutational hot-spot in exon 3 have been reported (c.316G>A or c.316G>C), both resulting in the same amino acid change (p.G106R). We screened a large cohort of Italian patients with spastic paraplegia for *NIPA1*mutations. *NIPA1* was analysed by high-resolution melting analysis (HRM) and/or direct sequencing in 240 unrelated HSP index cases, including 140 AD and 100 sporadic cases, negative for SPG4 mutations. The previously reported missense mutation c.316G>A/p.G106R was detected in a 2-generation family affected with a pure form of AD-HSP. Notably, the mutation was also found in 1 out of 100 sporadic patients. In both cases, the disease is characerised by adult-onset pure spastic paraplegia. In conclusion, in our population of patients with HSP, we detected*NIPA1* mutations in only 2 patients, both found to carry one of the two recurrent mutations in exon 3. Our data strongly support the hypothesis that *NIPA1* nucleotide c.316G is a hot-spot for recurrent mutations in SPG6. Furthermore, they indicate that *NIPA1* mutations are a rare cause of HSP, with a frequency <1%. As in the majority of SPG6 families reported to date, the clinical phenotype is characterised by pure slowly-progressive spastic paraplegia with adult onset.

2782W

Incomplete nonsense-mediated decay facilitates detection of a multi-exonic deletion mutation in SCGE. *J. Xiao, K. Marshall, M. LeDoux.* Department of Neurology, University of Tennessee Health Science Center, Memphis, Tennessee, 38163, USA.

Mutations in the maternally imprinted gene encoding ε-sarcoglycan (SGCE) represent the major cause of the myoclonus-dystonia syndrome (DYT11), an autosomal dominant disorder of reduced penetrance. Virtually all affected individuals have myoclonus, which is concentrated in the upper extremities, neck and trunk with infrequent involvement of the legs. Over half of patients have dystonia, usually affecting the neck or arms. Disease onset is usually in childhood or early adolescence but may appear in adults. Many affected individuals reported a dramatic reduction in myoclonus after consumption of alcohol. Psychiatric disorders including depression and anxiety can be prominent co-morbidities. Given that SGCE is maternally imprinted, penetrance is significantly higher when the mutant allele is inherited from the father. DYT11 may be caused by a variety of mutations in SGCE (nonsense, missense, insertions and deletions). Of the more than 70 SGCE mutations reported in the literature, 18 were large deletions disrupting at least one exon. Therefore, testing for exonic deletions should be considered in individuals with a classic phenotype in whom Sanger sequencing is unrevealing. However, standard methodologies for detection of exonic deletion mutations are expensive, labor intensive and can be produce false negatives. Herein, we report the use of cDNA derived from leukocyte RNA to identify a deletion mutation (Exons 4 and 5) of SGCE in a family with DYT11. Residual RNA from incomplete nonsense-mediated decay permitted reverse transcription to cDNA. Breakpoints of the 8,939 bp heterozygous deletion were then defined with long-range PCR and Sanger sequencing. Use of cDNA generated by reverse transcription of leukocyte RNA can reduce the costs associated with diagnostic genetic testing and can facilitate detection of deletion mutations.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2783T

Isoform-selective MeCP2-e1 deficient mice recapitulate the neurologic defecits of Rett syndrome. *D.H. Yasui[1], K.W. Dunaway[1,2], M.S. Golub[3], N.C. Schanen[4], J.M. LaSalle[5].* 1) Med Microbiology & Immunology, Univ California Davis, Med Sch, Davis, CA; 2) Genome Center, UC Davis, Davis, CA 95616; 3) Department of Environmental Toxicology, UC Davis, Davis, CA; 4) Nemours Biomedical Research Institute, Al Du Pont Hospital, Wilmington, DE; 5) UC Davis, MIND Institute, Sacramento, CA 95817.

Rett syndrome (RTT) is an X-linked dominant neurodevelopmental disorder characterized by the decline of language, social and motor skills between 6–18 months of age caused by mutations in MECP2 encoding methyl CpG binding protein 2 (MeCP2). Two protein isoforms, MeCP2-e1 and MeCP2-e2 are produced by alternative translation start sites encoded in alternatively spliced exon 1 and exon 2, respectively. As MECP2 mutations in RTT were primarily found in exons 3 and 4, previous mouse models of Rett syndrome focused on deletion of exons 3 and/or 4 that recapitulated some neurologic phenotypes of RTT including decreased survival. In contrast, Mecp2 exon 2 knockout mice exhibited normal neurologic function. Recent screening of RTT patients have revealed mutations in exon 1 but not exon 2, suggesting that MeCP2-e1 may be the critical RTT deficient isoform. To determine the clinical relevance of the less well characterized MeCP2-e1, we generated mice with isoform-selective ablation of MeCP2-e1 production through a single point mutation in the exon 1 translational start site. Western blotting and immuno-fluorescent (IF) analyses confirm the absence of MeCP2-e1 and reveal elevated levels of MeCP2-e2 despite an overall 50% reduction in total MeCP2 in Mecp2-e1 mutant brain. Mecp2-e1 mutant mice were born healthy and at the expected frequency suggesting that MeCP2-e1 is not essential for embryonic development. However, adult Mecp2-e1 mutant mice exhibited RTT like features including altered anxiety in the elevated plus maze, reduced sociability in the three chamber paradigm and gait abnormalities in tread-scan analysis in addition to reduced lifespan. Distinct sub-nuclear distribution of MeCP2-e1 versus MeCP2-e2 was observed by IF in brain, as MeCP2-e1 primarily localized to the nucleoplasm while MeCP2-e2 localized to heterochromatin domains in neurons. Consistent with immuno-localization studies, mass spectroscopic analysis of factors co-purifying with MeCP2-e1 revealed association with SFPQ splicing and Matrin3 nuclear matrix proteins in neuronal cells. These associations were confirmed by subsequent reciprocal co-immunoprecipitation and Western blot analysis along with MeCP2-e1 co-localization with SFPQ and Matrin3 in brain by IF. Collectively, these studies suggest unique biologic functions for Mecp2-e1 and MeCP2-e2 and demonstrate the relevance of MeCP-e1 to the RTT phenotype.

2784F

ARHGEF26/SGEF controls arteriosclerosis, retinal fovea, immunity and neurodevelopment. *P. Bitoun[1], E. Pipiras[2,3], B. Benzacken[2,3], A. Delahaye[2,3].* 1) Gen Med, CHU Paris-Nord, Hopital Jean Verdier, BONDY, France; 2) Histo-Embryo-Cytogenetics, CHU Paris-Nord,Hopital Jean Verdier, BONDY, France; 3) Inserm U676, Hopital Robert Debre, Paris.

Purpose: To identify the cause of a novel congenital syndrome with macular dystrophy, absence of fovea identified in a consanguineous family 6 yo girl with congenital nystagmus and low vision after informed consent. Methods: The family had complete ophthalmologic, neurologic examination, ERG, VEP, OCT, angiograms, brain MRI, genetic tests, array CGH and lipid profiles. Parents had vascular Doppler US. Antibody In situ hybridisation was performed in fetal Mouse eye and human fetal brain. Results: Proband had congenital nystagmus and prenatal corpus callosum agenesis. Vision was 20/200,no photophobia, night blindness or dyschromatopsia. Fundus showed macular atrophy with normal vessels w/o pigmentary deposits and small optic papilla. Angiography confirmed symetric macular atrophy with normal auto-fluorescence. Flash ERG, OP and dark adaptation were normal, absent pattern VEP response confirmed macular bundle defect. OCT showed macular atrophy with absent foveal pit, thin retina(102 µm OD and 92 OS) and interrupted photoreceptor layer. MRI also showed hippocampi hypoplasia. ABCA4 mutation testing was negative as well as screening of 18 known ARRP genes. Parents have normal visual acuity, ERG, VEP and OCT and retinal angiographs and brain MRI. Array CGH showed a 114kb homozygous 3q23 deletion upstream of 6th intron of the SGEF gene in proband and 2 sibs. Her sister was found to have a similar absence of foveal development on OCT as well as low vision and macular dystrophy, her brother had bilateral subclinical decreased foveal function on multifocal ERG with normal vision. Both girls have repeated infections: EBV hepatitis,dental abcesses... Asymptomatic parents harbor the same heterozygous deletion.SGEF antibody in situ hybridisation localised to E13.5 fetal mouse retina and hybridised to cortical region in 13.5WG human fetal brain. SGEF controls leucocyte trans-endothelial migration,1st phase of atherosclerosis. Mother with hypercholesterolemia of 212mg/dl and hyper-triglyceridemia had normal cervical arteries doppler ultrasound imaging without any sign of atherosclerosis at age 42 years. Conclusions: Present data suggests ARHGEF26/SGEF gene controls arteriosclerosis, macular cone development, immune function and axon midline crossing with apparently recessive inheritance. This deletion possibly has a protective effect against the development of arteriosclerosis blocking the 1st phase of atherosclerotic plaque formation.

2785W

Homozygosity mapping and exome sequencing in a Saudi-Arabian family with unspecified juvenile neurodegeneration. *E. Frackelton[1], P.M.A. Sleiman[1,2,3], L. Tien[1], S. Panossian[1], K. Xu[1], B. Tweddale[1], N. Abdel-Magid[1], C. Kim[1], Y. Housawi[4], H. Hakonarson[1,2,3].* 1) Ctr Applied Genomics, Children's Hosp Philadelphia, Philadelphia, PA; 2) Division of Human Genetics, Children's Hosp Philadelphia, Philadelphia, PA; 3) Dept Pediatrics, Perelman School of Medicine, Univ of Pennsylvania, Philadelphia PA; 4) King Fahad Specialist Hospital-Dammam, Saudi Arabia.

A Saudi Arabian family of two healthy parents and three children all of whom were affected with an unspecified neurodegenerative condition was referred to us for sequencing after screening negative for a panel of potential candidate genes. All five family members were initially genotyped on a high density array (Illumina HH610) and homozygosity mapping carried out in plink. The three affected children shared 7 regions of homozygosity on chromosomes 1, 5, 6, 7, 11, 14 and 22. Two of the three affected sibs also had their exomes sequenced on an Illumina HiSeq 2000 to an mean depth of ~ 100x after exome capture and enrichment with the Agilent SureSlect 50Mb kit. Variants annotation and filtering was performed in ANNOVAR. The filtration gates included retaining variants that mapped to the regions of homozygosity, those that were shared between the affected sibs and were homozygous, and excluding variants that were represented in the publically available datasets. We identified a homozygous stop gain mutation Y790X in exon 17 of the PLA2G6 gene which was subsequently confirmed by sanger sequencing and shown to segregate with the phenotype in the family. PLA2G6 is a member of the phospholipase A2 (PLA2s) family which catalyze the hydrolysis of the sn-2 acyl-ester bonds in phospholipids, leading to the release of arachidonic acid and other fatty acids. Mutations of the PLA2G6 gene have previously been associated with Infantile Neuroaxonal dystrophy, Karak syndrome, Neurodegeneration with brain iron accumulation 2B, and with adult onset Parkinsonism (PD14). The Y790X mutation has been reported in classic infantile neuroaxonal dystrophy (INAD) Morgan et al. (2006). INAD is a rare, rapidly progressing disorder that usually begins between ages six months and three years with developmental regression, hypotonia, progressive psychomotor delay, and progressive spastic tetraparesis. Diagnosis of INAD is based on the identification of dystrophic axons on electron microscopic examination of nerve ultrastructure in a tissue biopsy. As axonal spheroids accumulate with age and may not be evident in all tissues, individuals with INAD may require multiple biopsies over time before axonal spheroids are identified. This study demonstrated the utility of exome sequencing in rapidly identifying the cause of conditions that can be difficult to classify clinically and thus sequence the appropriate candidate genes.

2786T

Neuropathology in POL-III related disorders: Myelin loss in 4H syndrome. *A. Vanderver[1], J. Schmidt[1], G. Carosso[1], D. Tonduti[2], G. Bernard[3], J. Lai[4], C. Rossi[5], M. Quezado[4], K. Wong[6], R. Schiffman[7].* 1) Center for Genetic Medicine, Children's National Medical Center, Washington DC; 2) Mondino National Institute of Neurology Foundation, University of Pavia, Pavia, Italy; 3) Montreal Children's Hospital, Montreal, Canada; 4) Department of Pathology, National Institutes of Health, Bethesda, MD; 5) Department of Pathology, Children's National Medical Center, Washington DC; 6) Department of Pathology, USAF, San Antonio, TX; 7) Institute of Metabolic Disease, Baylor Research Institute, Dallas, TX.

INTRODUCTION: Mutations in POLR3A and POLR3B, encoding the largest (catalytic core) subunits of human RNA Polymerase III (Pol III), were identified as causative of 4H Syndrome (hypomyelination with hypogonadotropic hypogonadism and hypodontia) and related hypomyelinating phenotypes, now referred to as Pol III related leukodystrophy. While mechanisms of pathogenesis remain to be elucidated, here we provide a first neuropathological description of brain tissue. METHODS: Brain tissue from a POLIII mutation-proven patient and age-matched controls were obtained according to an IRB approved protocol. Paraffin-embedded tissue was examined using standard immunohistochemistry approaches. Neuroimaging and clinical history on the 4H patient was obtained from the Myelin Disorders Bioregistry Project. RESULTS: Cerebral MRI findings show diffuse white matter signal abnormality consistent with a diagnosis of hypomyelination. Histologic analysis reveals variable myelin rarefaction and loss, reduction in axons, and reduction of oligodendroglia underlying most gyri, although myelin and oligodendroglia are better preserved surrounding small blood vessels. Astrogliosis is seen mostly in perivascular white matter, and rare CD68+ macrophages are seen juxtaposed to oligodendroglia. CONCLUSIONS: The present 4H patient's brain pathology reveals marked loss of oligodendroglia with severe loss of myelin. Astrogliosis, microgliosis, and macrophagic phagocytosis are mild compared to that found in other myelin compromising disorders. The observed CD68+ macrophagic reaction involving viable oligodendrocytes, as well as the observed reduction in IFN- β, are suggestive of an immunologic process. This affirms a primary myelin disorder in Pol III related leukodystrophy.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2787F

Identification of Homozygous WFS1 Mutations (p.Asp211Asn, p.Gln486*) Causing Severe Wolfram Syndrome and First Report of Male Fertility. *A. Haghighi[1], A. Haghighi[2], A. Setoodeh[3], N. Saleh-Gohari[4], D. Astuti[5], T.G. Barrett[5].* 1) The Hospital for Sick Children, University of Toronto, Toronto, Canada; 2) Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK; 3) Growth & Development Research Centre, University of Tehran/ Medical Sciences, Tehran, Iran; 4) Genetic Department, Kerman University of Medical Sciences, Kerman, Iran; 5) Diabetes Department, Institute of Child Health, Diabetes Department, University of Birmingham, Birmingham, UK.

Wolfram syndrome (WS) is a neurodegenerative genetic condition characterized by childhood-onset diabetes mellitus and optic atrophy. Other features of this syndrome include diabetes insipidus, sensorineural hearing loss, cerebellar ataxia, psychiatric illness, neuropathic bladder and hypergonadotrophic hypogonadism. There is no treatment available for WS. The mortality is high and only about 40% of patients survive beyond the age of 35 years. WFS1 is the main gene identified as the cause of Wolfram syndrome and encodes Wolframin, a transmembrane protein that is involved in the regulation the unfolded protein response. There is as yet no clear genotype-phenotype correlation in WS. We undertook a clinical and molecular study of severe WS in two consanguineous families from Iran. The clinical features of these two families were compared with each other and with the previously published cases of WS in order to correlate genotypes with phenotypic features. Direct sequencing of all the exons and intron-exon boundaries of WFS1 was undertaken in the index cases on an ABI 3730 DNA sequencer. We screened for mutations in first degree relatives of both families. The clinical diagnosis of Wolfram syndrome was confirmed by the presence of diabetes mellitus and optic atrophy in the index cases, who in addition had severe neurodegenerative complications. Despite the neurodegenerative complications, two affected males in the first family had fathered unaffected children. We identified two homozygous mutations: family 1: c.631G>A (p.Asp211Asn) in exon 5, and family 2: c.1456C>T (p.Gln486*) in exon 8. Heterozygous carriers were unaffected. Genotpying of a panel of markers in both father/child pairs confirmed a high probablility of paternity. This is the first report of male Wolfram patients who have successfully fathered children. Surprisingly, our patients also had, almost all the complications associated with Wolfram syndrome. Male hypogonadism has been a well reported feature of Wolfram syndrome, in contrast to affected females, for whom there are several reports of successful pregnancies. Data from mouse knockout models show reduced spermatogenesis and reduced fertility in males compared with females. Our findings have implications for genetic counselling for future families affected by these mutations, and raise the hope that some affected males may be able to father children.

2788W

Whole Exome Sequencing in Children with Unclassified, Sporadic, Early-Onset Epileptic Encephalopathies. *K.R. Veeramah[1], L. Johnstone[1], D. Wolf[1], R. Sprissler[1], T.M. Karafet[1], D. Talwar[2,3,4], R.P. Erickson[2,5], M.F. Hammer[1].* 1) Arizona Research Laboratories, University of Arizona, Tucson, AZ; 2) Department of Pediatrics, Arizona Health Science Center, Tucson, AZ; 3) Department of Neurology, Arizona Health Science Center, Tucson, AZ; 4) Center for Neurosciences, Tucson, AZ; 5) Department of Molecular and Cellular Biology, University of Arizona, Tucson, AZ.

Early-onset Epileptic Encephalopathies (EEEs) are brain disorders characterized by seizures that usually begin within the first year of life, followed by severe cognitive and behavioral impairment and deterioration. Many children are refractory to treatment and early death is not unusual. Some children with EEEs are classified within particular syndromes described by the International League Against Epilepsy (ILAE) with an emphasis on clinical utility, but because EEEs are highly phenotypically heterogeneous, many children remain unclassified. While a handful of genes such as *SCN1A* and *ARX* have been associated with EEEs, the underlying genetic etiologies in most children are unknown—a result that makes sense given the large number of genes that potentially could disrupt brain development. In fact, the large `genic target' has been presented as an evolutionary explanation for the relatively high incidence of severe neurological disorders in the general population. Thus, it is reasonable to hypothesize that many sporadic cases of EEE result from a spontaneous new (*de novo*) mutation in a diverse set of genes. Therefore whole genome and exome approaches are promising avenues for identifying pathogenic EEE mutations and associated genes. We recently identified a causal mutation in a young girl with EEE in a gene not previously associated with any seizure disorder, *SCN8A*, using a whole genome sequencing family approach. Building on this work we perform high-coverage (50–80X) whole exome sequencing in 15 children with undiagnosed sporadic EEEs as well as their unaffected parents. We then identify potentially pathogenic *de novo* (as well as recessive) mutations using our custom bioinformatics pipeline. In addition, we develop a systematic classification system that attempts to capitalize on detailed clinical and laboratory parameters scored in patients with EEEs over their lifespan. This system is based on various clustering, machine learning, and data-dimensionality-reducing methods that can quantify the phenotypic similarities and differences among EEE cases. Application of this classification system enables us to quantify the amount of phenotypic variation and subclass structure among our cohort. We then quantitatively explore the hypothesis that children with more similar phenotypic characteristics share causal EEE genes that are found in more similar developmental or metabolic pathways.

2789T

Determining the genetic architecture of Joubert syndrome. *D. O'Day[1,3], B.J. O'Roak[2], I.G. Phelps[1,3], J.C. Dempsey[1,3], I. Glass[1,3], H.C. Mefford[1,3], J. Shendure[2], D. Doherty[1,3].* 1) Department of Pediatrics, University of Washington, Seattle, WA; 2) Department of Genome Sciences, University of Washington, Seattle, WA; 3) Seattle Children's Hospital, Seattle, WA.

Joubert Syndrome (JS) is a recessive ciliopathy characterized by cognitive impairment, hypotonia, abnormal respiratory pattern, abnormal eye movements and a distinctive hindbrain malformation, “the molar tooth sign.” Subsets of patients also develop retinal dystrophy, cystic kidney disease and liver fibrosis. JS is genetically heterogeneous, with causal mutations identified in 18 genes; however, no cohort of patients has been comprehensively sequenced. To determine the prevalence of each genetic cause, we used an efficient and inexpensive molecular inversion probe capture technique and next generation sequencing to sequence 16 of the known JS genes in 308 subjects representing 258 families. For captured basepairs, the sensitivity and positive predictive value were 99% and 98% respectively, compared to previously generated bidirectional Sanger sequencing data. Using the same data, we were also able to identify the common deletion of *NPHP1* as well as other small copy number variants. We identified causal mutations in 55% of the families, with 5 genes (*CC2D2A*, *TMEM67*, *CEP290*, *AHI1*and *TMEM216*) accounting for 35% of patients. The majority of mutations in *CEP290* were truncating, while other genes harbored mainly missense mutations, potentially indicating that each gene requires a different severity of loss of function to generate the JS phenotype. Missense mutations highlight domains of the proteins important for function. Previous genotype-phenotype correlations were borne out in the more comprehensive data; mutations in *CEP290* and *AHI1* were associated with retinal disease, while mutations in *TMEM67* were associated with liver fibrosis. Families in which the known genetic causes have been excluded provide a resource for identifying the remaining undiscovered causes of JS.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2790F

Genetics of Joubert Syndrome in the French Candian population. *M. Srour[1], J. Schwartzentruber[2], F.F. Hamdan[1], L.H. Ospina[3], L. Patry[1], D. Labuda[4], C. Massicotte[4], S. Dobrzeniecka[1], J.M. Capo-Chichi[1], S. Papillon-Cavanagh[4], M.E. Samuels[4], K.M. Boycott[5], M.I. Shevell[6], R. Laframboise[7], V. Désilets[4], D. Amron[1], E. Andermann[8], G. Sebire[9], B. Maranda[10], G.A. Rouleau[11], J. Majewski[8], J.L. Michaud[1] FORGE Canada Consortium.* 1) Centre of Excellence in Neurosciences of Université de Montréal and Sainte-Justine Hospital Research Center, Montréal, Canada; 2) McGill University and Genome Quebec Innovation Centre, Montréal, Canada; 3) Department of Ophthalmology, Sainte-Justine Hospital Research Center, Montréal, Canada; 4) Sainte-Justine Hospital Research Center, Montréal, Canada; 5) Children's Hospital of Eastern Ontario Research Institute, Ottawa, Canada; 6) Division of Pediatric Neurology, Montreal Children's Hospital-McGill University Health Center, Montreal, Canada; 7) Department of Medical Genetics, Centre Hospitalier Universitaire Laval, Québec, Canada; 8) Department of Human Genetics, McGill University, Montréal, Canada; 9) Division of Neurology, Centre Hospitalier Universitaire de Sherbrooke, Sherbrooke, Canada; 10) Division of genetics, Centre Hospitalier Universitaire de Sherbrooke, Sherbrooke, Canada; 11) Centre of Excellence in Neurosciences of Université de Montréal, CHUM Research Center and Department of Medicine.

Joubert syndrome (JBTS) is an autosomal recessive disorder characterized by a mid-hindbrain malformation, developmental delay, ocular-motor apraxia and breathing abnormalities. To date, causal mutations have been identified in at least 16 genes, all of which are involved in non-motile cilia function. Though JBTS was first described in French Canadian (FC) siblings over 40 years ago, the causal gene has not yet been identified in this family, nor in most FC cases. This study aims to study the genetic basis of JBTS in the FC population. We have ascertained 28 FC individuals with JBTS from 22 families. We noted a regional clustering of 16 cases in the Lower St-Lawrence region, suggesting the possibility of a regional founder effect. SNP genotyping excluded the presence of a common homozygous mutation. We performed exome sequencing in 15 subjects. 9 individuals from 7 families (including the original JBTS family) carried rare compound heterozygous mutations in *C5ORF42*. Two missense variants (p.R1336W, p.A1564T) and a splicing mutation (c.7400+1G>A) were found in multiple subjects, whereas three other mutations, all truncating (p.P2136fs, p.R1602X, p.R2493X), were identified in single individuals. Genotyping 477 FC controls did not identify any of the 6 mutations. Exome sequencing also identified compound heterozygous mutations (p.D1556V, p.E1126K) in *CC2D2A*, a gene already known to be responsible for JBTS in 2 unrelated individuals. We performed whole-exome sequencing in 10 other FC individuals with JBTS from our cohort originating from other regions of Quebec. We identified 1 individual with mutations in *C5ORF42* (c.7400+1G>A and p.K2753fs), and 5 other individuals with mutations in *CC2D2A* (p.D1556V, p.E1126K, p.Y1568H, p.N1520S, p.1150_1151del). Screening for the recurrent p.D1556V and p.E1126K mutations in the remainder of the cohort identified 2 additional individuals with mutations in *CC2D2A*. Mutations in *C5ORF42* and CC2D2A are responsible for a large fraction of JBTS cases in the FC population, and there is presence of a complex founder effect. In the as yet unexplained individuals with JBTS, exome sequencing identified interesting variants in several candidate genes which we are in the process of validating.

2791W

Mechanisms underlying non-recurrent microdeletions causing neurofibromatosis type-1 (NF1). *J. Vogt[1], K. Bengesser[1], K. Claes[2], K. Wimmer[3], L. Messiaen[4], L. Kluwe[5], V.-F. Mautner[5], H. Kehrer-Sawatzki[1].* 1) Institute of Human Genetics, University of Ulm, Ulm, Germany; 2) Center for Medical Genetics, Gent University Hospital, 9000 Gent, Belgium; 3) Division of Human Genetics, Medical University Innsbruck, Innsbruck, Austria; 4) Medical Genomics Laboratory, Department of Genetics, University of Alabama at Birmingham, Birmingham, USA; 5) Department of Neurology, University Hospital Hamburg Eppendorf, Hamburg, Germany.

NF1 microdeletions encompassing the *NF1* gene region at 17q11.2 are present in 5–10% of patients with NF1. Whereas the mechanisms underlying recurrent *NF1* microdeletions have been investigated in greater detail, those underlying non-recurrent (atypical) *NF1* microdeletions are not well delineated. *NF1* microdeletions with non-recurrent breakpoints are heterogeneous in terms of their size, breakpoint position and number of deleted genes. Furthermore, extended sequence homology is not observed in the respective breakpoint regions. In this study, we investigated 12 atypical *NF1* deletions using high resolution custom made array CGH. We could assign the breakpoints to regions of 1.2–6 kb. In seven of these 12 atypical *NF1* deletions, we identified the breakpoints at the level of a few basepairs. Five of these seven deletions were mediated by non-homologous end joining (NHEJ) as concluded by the absence of or only minor (1–2bp) homology at the breakpoints. Two of these seven *NF1* deletions exhibited microhomologies of 24 and 33 bp at the breakpoint sites indicative of microhomology-mediated end joining (MMEJ) as underlying mechanism. We conclude that NHEJ or MMEJ are the prevailing mechanisms underlying non-recurrent *NF1* deletions that lack any complexity at the deletion breakpoint sites. However, four of the 12 *NF1* deletions investigated by us seem to represent complex rearrangements most likely caused by replication associated erroneous template switches. For one of these complex rearrangements we already have a breakpoint-spanning PCR fragment, which harbours multiple breakpoints, indicative of several rearrangements at the boundaries of the*NF1* microdeletion. This is most likely the result of replication-based errors. In order to characterize the complex rearrangements at the breakpoint sites of all four complex atypical *NF1*microdeletions, we want to perform a targeted capture and next generation sequencing (TCNGS) approach. Chimeric sequence reads will indicate breakpoint-spanning sequences and thus reveal the structure and orientation of sequences at the breakpoints. In summary, our study indicates for the first time, that also non-recurrent *NF1* microdeletions are mediated by a variety of mutagenic processes including mechanisms of double strand repair such as NHEJ and replication based mechanisms such as Fork stalling and template switches (FoSTeS).

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2792T

The DYRK1A gene is a rare cause of Angelman-like syndrome with severe microcephaly and epilepsy. *JB. Courcet[1], E. Lopez[1], J. Thevenon[1,2], N. Gigot[1,2,3], L. Jego[1], C. Ragon[1,4], A. Masurel-Paulet[2], P. Callier[1,4], N. Marle[1,4], AL. Mosca-Boidron[1,4], F. Huet[1,2], C. Philippe[5], L. Lambert[6], P. Malzac[7], L. Faivre[1,2], A. Moncla[8], C. Thauvin-Robinet[1,2].* 1) Equipe émergente GAD (Génétique des Anomalies du développement), IFR Santé STIC, Université de Bourgogne, Dijon, France; 2) Centre de Génétique et Centre de Référence Anomalies de Développement et Syndromes Malformatifs de l'interrégion Grand-Est, Hôpital d'Enfants, CHU, Dijon, France; 3) Laboratoire de Génétique Moléculaire, Plateau Technique de Biologie, CHU Dijon, France; 4) Laboratoire de Cytogénétique, Plateau Technique de Biologie, CHU de Dijon, Dijon, France; 5) Laboratoire de Génétique et EA 4368, CHU de Nancy, Nancy, France; 6) Service de médecine infantile III et de génétique clinique, Laboratoire de génétique, Unité de génétique du service néonatale, Maternité régionale de Nancy, Nancy, France; 7) Département de Génétique Médicale, Laboratoire de génétique moléculaire, CHU de Marseille, Hôpital de la Timone, Marseille, France; 8) Département de Génétique Médicale, Unité de Génétique Clinique, CHU de Marseille, Hôpital de la Timone, Marseille, France.

Angelman syndrome (AS) belongs to the group of syndromic mental retardation (MR). Methylation anomalies at the SNRPN locus and UBE3A mutations are found in 85% of cases. Patients presenting with features compatible with AS (severe MR, microcephaly, epilepsy, absent speech and ataxia) but no AS genetic causes define the group of « Angelman-like » syndrome (AS-like). Recently, we identified a 69 kb deletion in 21q22 including the 5' part of the DYRK1A gene in a 3 year-old girl, presenting with intra uterine growth retardation (IUGR), facial dysmorphism, congenital microcephaly, seizures, ataxic gait, absent speech, and developmental delay. Four patients with intragenic DYRK1A rearrangements or overlapping 21q22 microdeletions including only DYRK1A have previously been reported with a similar phenotype. Thus, we hypothesized that DYRK1A mutations could be responsible for AS-like. The DYRK1A gene was studied in a cohort of 105 patients previously screened by methylation analysis on SNRPN locus +/− UBE3A sequencing and the following criteria: MR and at least two symptoms from AS spectrum (microcephaly < -2.5 SD, ataxic gait, seizures and/or speech delay). We identified a de novo frameshift DYRK1A mutation (c.290_291delCT; p.Ser97Valfs*) leading to DYRK1A haploinsufficiency in a 14 year-old patient with IUGR, persistent growth retardation, feeding difficulties during childhood, delayed language, MR, primary and secondary severe microcephaly (OFC= -6 SD) and epileptic seizures. Therefore, we aimed to delineate from this study and previously reported patients the criteria for sequencing the DYRK1A gene. When microcephaly, epilepsy and language delay/absent speech are considered as mandatory features, a DYRK1A mutation is found in 1/43 patients (2.3%). DYRK1A is believed to have an important role in controlling brain growth through neuronal proliferation and neurogenesis has been shown to be affected in MR disorders with primary microcephaly. DYRK1A protein is expressed in a gene dosage dependent manner so that, DYRK1A haploinsufficiency induce a reduced brain size in animal models. Numerous evidences support also the hypothesis that DYRK1A is implicated in neurodevelopmental alterations underlying the cognitive deficits of Down syndrome. Altogether, we show that DYRK1A is a rare cause of AS-like and the DYRK1A analysis could be discussed when patients present with MR, severe microcephaly, epilepsy and language delay/absent speech.

2793F

Exome Sequencing Reveals Cause of Hypomyelinating Leukodystrophy. *D.L. Dinwiddie[1], N.A. Miller[1], C.J. Saunders[1], S.E. Soden[1], G.M. Hobson[2], S.F. Kingsmore[1].* 1) Center for Pediatric Genomic Medicine, Children's Mercy Hospital, Kansas City, MO; 2) Alfred I. duPont Hospital for Children, Nemours Biomedical Research, Wilmington, DE.

Here we present a case in which exome sequencing rapidly produced a molecular diagnosis for a patient with hypomyelinating leukodystrophy for whom years of costly conventional testing had not identified a molecular diagnosis. The patient had severe developmental delay, ataxia, nystagmus, external opthalmoplegia, hypometric saccades, and high myopia. Magnetic resonance imaging of the brain demonstrated severe failure of myelination of the corona radiate and centrum semi ovale. Through exome sequencing, two novel heterozygous missense variants were identified in the POLR3A gene (c.1160C>G; p.Ala387Gly and c.3781G>A; p.Glu1261Lys). We propose the use of exome sequencing for the diagnosis of genetically heterogeneous conditions such as hypomyelinating leukodystrophy as a cost-effective primary diagnostic tool.

2794W

Intractable epilepsy and tau protein pathology in a child with a defect in the synaptic vesicle recycling pathway. *D.A. Dyment[1], A. Smith[2], P. Humphreys[3], J. Woulfe[4], The. Forge Canada Consortium[5], D.E. Bulman[2], J. Schwartzentruber[6], J. Majewski[6], J. Michaud[7], K.M. Boycott[1].* 1) Department of Genetics, Children's Hospital of Eastern Ontario, Ottawa, Canada; 2) Ottawa Hospital Research Institute; 3) Division of Neurology, Children's Hospital of Eastern Ontario, Ottawa, Canada; 4) Department of Pathology and Laboratory Medicine, The Ottawa General Hospital, Ottawa, Canada; 5) Forge Steering Committee is listed in Acknowledgements; 6) McGill University and Genome Quebec Innovation Centre; 7) Department of Pathology and Laboratory Medicine, Childrens Hospital of Eastern Ontario.

Individuals with Alzheimer disease have a well-recognized increased incidence of epilepsy though the exact mechanism is unclear. Here, we present a 6-year-old child of a consanguineous marriage with an intractable seizure disorder. The child's severe seizure disorder began at 9 days of life and was often continuous and refractory to treatment. An exhaustive neurological, metabolic and genetic workup was negative. At the time of his passing, he required G-tube feeding, he was cortically blind, had little motor skills and had never developed the ability to communicate verbally. Pathology of his brain showed generalized atrophy and severe loss of white matter, with moderate loss of neurons in the substantia nigra and tau protein positive neurofibrillary degeneration in most of the remaining neurons. Exome sequencing identified nine novel variants, seven of which were within regions of homozygosity as defined by a SNP microarray. Only one of these variants was a truncating mutation and expressed in the central nervous system. The gene identified was Synaptojanin1 and the observed C>T transition changes a p.R97X. A western blot of brain tissue of the patient showed no detectable synaptojanin protein present. Mouse knock-out models of synaptojanin1 from the literature result in mice that are ataxic, have convulsions and eventually die early in life. The gene product functions at the axon terminal to uncoat clathrin from the recycled synaptic vesicle. The child's homozygous defect at synaptojanin1 is likely responsible for his refractory seizure disorder; moreover, the homozygous mutation suggests that the unusual pathological finding of tau bodies and neurofibrillary tangles in a pediatric patient are related to an inherent defect at the synapse.

2795T

Novel *CDKL5* splicing variant in a boy with early-onset seizures. *R. Polli[1], G. Ho[2], E. Leonardi[1], E. Bettella[1], E. Fontana[3], B. Dalla Bernardina[3], J. Christodoulou[2], A. Murgia[1].* 1) Deptartment of Pediatrics, University of Padua, Padua, Italy; 2) NSW Centre for Rett Syndrome Research, The Children's Hospital at Westmead, Sydney, Australia; 3) Child Neuropsychiatry, University Hospital of Verona, Verona, Italy.

Genetic variants of the X-linked cyclin-dependent kinase-like 5 gene (*CDKL5*; OMIM 300203) cause a condition characterized by severe developmental delay, with a seizure onset in most before the age of three months. The CDKL5 disorder, previously considered a Rett syndrome variant, is now believed to be an independent clinical entity associated with early-onset epileptic encephalopathy. Mutations of the *CDKL5* gene have been mainly reported in females, while a few male subjects have been described with concomitant chromosomal alterations or carrying missense variants. In this study we report the case of a 2 year old boy with a very severe developmental disorder with early-onset drug-resistant seizures, hypotonia, and microcephaly. Folowing a metabolic work-up which showed no abnormalities, high resolution karyotype, and *MECP2* and *CDKL5* genetic testing was performed by dHPLC analysis and direct sequencing, followed by quantitative real-time PCR. A pathogenic mutation was found in the donor splice site of *CDKL5* intron 10, c.825+1G>T (NM_003159.2). The variant, not present in Rett syndrome specific databases, and not previously reported in literature, was found to be de novo. To establish the effect on splicing and to try and define the pathogenic relevance of this alteration, a study was conducted using a minigene system with a modified pBudCE 4.1 vector (Life Technologies, Carlsbad, CA) and transfection in COS-7 cells. The minigene experiments revealed that the mutation only partially affects the splicing of exon 10, allowing a proportion of the wild type transcript to be synthesized. This “leaky” mutation explains how an otherwise potentially severe pathogenic *CDKL5* mutation can be compatible with survival in a chromosomally normal male subject beyond the age of one year.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2796F

The molecular role of GTF2IRD1, a protein implicated in the neurodevelopmental features of Williams-Beuren syndrome. *P. Carmona-Mora[1], J. Widagdo[1], K.M. Taylor[1], R. Tsz-Wai Pang[1], P.W. Gunning[2], E.C. Hardeman[1], S.J. Palmer[1].* 1) School of Medical Sciences, Neuromuscular and Regenerative Medicine Unit, University of New South Wales, Sydney, Australia; 2) School of Medical Sciences, Oncology Research Unit, University of New South Wales, Sydney, Australia.

Williams-Beuren syndrome (WBS) is an autosomal dominant disorder caused by a hemizygous microdeletion within chromosome 7q11.23. Many of the physical abnormalities can be attributed to haploinsufficiency of elastin but the remaining unattributed features include a characteristic set of craniofacial abnormalities and a very distinctive neurocognitive profile involving visuospatial problems, reduced social anxiety and high incidence of psychopathologies including ADHD and phobias. Analysis of patients with atypical deletion breakpoints have shown that two related genes, *GTF2I* and *GTF2IRD1* encoding transcriptional regulator proteins, are responsible for the craniofacial and neurological defects but the molecular and cellular mechanism underpinning their role in the disease remains unclear. Mice with targeted null mutations of *Gtf2ird1* recapitulate some of the defects of WBS including a facial soft tissue abnormality and behavioural defects. We aim to reveal the normal function of the GTF2IRD1 protein using these mice as a model system, in combination with biochemical analyses to study molecular interactions. We have used a yeast two-hybrid screening system as an unbiased way to identify the protein binding partners of GTF2IRD1 and the pathways it integrates with. Secondly, we have conducted microarray studies using corpus striatum tissue from *Gtf2ird1* knockout mice in order to identify the transcriptional changes arising in this part of the brain, which shows high-level *Gtf2ird1* expression and is relevant to the potential cause of ADHD. Our yeast two-hybrid work has identified a panel of novel interacting partners including DNA binding proteins, chromatin modifying enzymes and proteins involved in post-translational modifications such as SUMO ligases and ubiquitination components. Our analysis of transcript profile in the striatum of knockout mice has revealed significant changes in genes involved in neuronal development and a cluster of immediate-early response genes that are known to correlate with heightened levels of motor activity. Together, these data suggest GTF2IRD1 acts as a nuclear adaptor protein that bridges connections between specific transcription factors and components of the chromatin modification machinery in order to regulate gene expression. Secondly, GTF2IRD1 levels and activity appear to be tightly controlled by post-translational mechanisms, which may be a characteristic feature of dosage sensitive genes causative of WBS.

2797W

Contribution of *POLR3A* and *POLR3B* recessive mutations to hypomyelinating leucodystrophy. *H. Daoud[1], M. Tétreault[2], K. Choquet[2], W. Gibson[3], M. Synofzik[4], B. Brais[2], A. Vanderver[5], G. Bernard[6].* 1) Centre of Excellence in Neuroscience of Université de Montréal, CHUM Research Center and the Department of Medicine, Montreal, H2L 2W5, Quebec, Canada; 2) Neurogenetics of Motion Laboratory, Montreal Neurological Institute, McGill University, Montreal, Quebec H3A 2B4, Canada; 3) Child and Family Research Institute, Department of Medical Genetics, University of British Columbia, Canada; 4) Centre for Neurology and Hertie-Institute for Clinical Brain Research, University of Tübingen, Tübingen, Germany; 5) Department of Neurology, Children's National Medical Center, Washington D.C., 20010, USA; 6) Departments of pediatrics, Neurology and Neurosurgery, Division of Pediatric Neurology, Montreal Children's Hospital, McGill University Heath Center, Montreal, Quebec, H3H 1P3, Canada.

Aims and Background: Leukodystrophies are a heterogeneous group of inherited neurodegenerative disorders characterized by abnormal central nervous system white matter. We recently showed that mutations in *POLR3A* and *POLR3B* genes, encoding the two largest subunits of RNA polymerase III (Pol III), are a major cause of three clinically overlapping hypomyelinating leukodystrophy phenotypes. Our aim was to investigate the presence and frequency of *POLR3A* and *POLR3B* mutations in patients with genetically unexplained hypomyelinating leukodystrophies. **Methods:** The entire coding region and the flanking exon/intron boundaries of *POLR3A* and *POLR3B* genes were amplified and sequenced in 12 patients with hypomyelinating leucodystrophies. **Results:** We identified six patients with recessive mutations in *POLR3A* and *POLR3B* genes. Three patients carry five novel mutations in *POLR3A*, among who two were compound heterozygotes and one homozygote. The other three patients have compound heterozygous mutations in *POLR3B*, including a common missense shared by these three patients and two novel mutations. **Conclusions:** To date, our group has described 27 mutations in *POLR3A* and *POLR3B*. Altogether, our results further support that *POLR3A*and *POLR3B* mutations are a major cause of hypomyelinating leukodystrophy and suggest that *POLR3A* mutations are more frequent.

2798T

Identification of disease causing mutations in the serine/threonine kinase VRK1 gene in two siblings with primary microcephaly, neuropathy and spinal muscular atrophy. *C. Gonzaga-Jauregui[1], W. Wiszniewski[1], J. Reid[2], D. Muzny[2], R.A. Gibbs[1,2], J.R. Lupski[1,3,4], Baylor-Hopkins Center for Mendelian Genomics.* 1) Molec & Human Gen, Baylor College Med, Houston, TX; 2) Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX; 3) Department of Pediatrics, Baylor College of Medicine, Houston, TX; 4) Texas Children's Hospital, Houston, TX.

Targeted genomic capture and next-generation sequencing technologies have enabled the high-throughput sequencing of exomes to discover pathogenic variants in patients with rare or uncharacterized genetic diseases and the identification of new genes involved in disease. We have identified through genomic next-generation sequencing the genetic cause of the disease in two affected sisters that present with primary microcephaly, motor delay and infantile polyneuropathy for which the molecular diagnosis was not known and extensive clinical testing had been performed previously without concluding a specific etiologic diagnosis. We used targeted exome sequencing at 139 average depth of coverage in the youngest proband combined with whole genome sequencing using massively parallel sequencing by oligo ligation/detection technology, SOLiD, at ~30x average depth of coverage for the older sister in order to identify the most probable disease causing variants in these affected sisters. We identified more than 20,000 coding variants in each of the affecteds, with half of them being potential deleterious nonsynonymous or frameshifting variants; 3,000 of these were shared between both probands. Given the recurrent phenotype in this family, we hypothesized a recessive mode of inheritance and therefore examined the variant calls in this family for homozygous or compound heterozygous alleles amongst the affected sisters. We identified novel compound heterozygous missense deleterious mutations in the vaccine related kinase 1 gene, VRK1. VRK1 is a serine/threonine protein kinase involved in cell cycle progression and cell division; it has been observed to directly interact and phosphorylate p53, ATF2 and c-Jun, among other transcription factors important for cell proliferation. Mutations in VRK1 were recently reported to be responsible for Pontocerebellar hypoplasia type 1 (PCH1) in one consanguinous family. PCH1 is characterized by progressive microcephaly with central and peripheral nervous system involvement and motor dysfunction that resembles spinal muscular atrophy type 1, which can be misdiagnosed as axonal neuropathy. We demonstrate the utility of next-generation genomic sequencing approaches to answer a 10 year quest for an answer in a family with an undiagnosed rare phenotype and for which the findings can have tremendous clinical impact for the family including the enabling of prenatal diagnosis.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2799F

Molecular screening of patients with CMT2 disease. *G. Bergamin[1], C. Briani[2], E. Pegoraro[2], A. Martinuzzi[3], E. Battiston[1], F. Boaretto[1], M.L. Mostacciuolo[1].* 1) Departement of Biology, University of Padava, Padova, Padova, Italy; 2) Departement of Neurosciences, University of Padova; 3) E. Medea Scientific Institute, Conegliano Research Centre, Conegliano, Italy.

Charcot-Marie-Tooth disease type 2 (CMT2) comprises a group of clinically heterogeneous inherited axonal neuropathies. CMT2 is also genetically heterogeneous, but among the 13 genes associated to this phenotype, only 3 are found more frequently mutated in patients: MFN2, NEFL, and MPZ. In addition, mutations in CX32 can be found in CMT2 patients, especially in females with intermediate CMT phenotype. The aim of this study was to investigate the mutational spectrum of CMT2 disease in a cohort of 44 Italian patients, both familiar (23 cases) and isolated (21 cases). Direct sequencing was performed to detect mutations in the coding and splicing site-flanking regions of the MFN2, NEFL, MPZ and CX32 genes on genomic DNA from 44 unrelated CMT2 patients. In addition, rearrangements in PMP22 gene, that represent the most frequent cause of CMT1, were evaluated. In 14 of 44 index cases (about 32% of the cohort) we identified 15 mutations including 12 MFN2 (27,2%), 2 MPZ (4.5%) mutations and 1 PMP22 duplication (2,3%). In the MFN2 gene we found 9 missense (2 never reported before), 2 novel single amino acid deletions and 1 splice-site mutation. It is noteworthy that 1 patient affected by a severe axonal CMT resulted a compound heterozygote for 2 MFN2 mutations. The percentage of MFN2 mutations in CMT2 patients we found is in agreement with previous reports (16–33%), thus confirming that mutations in this gene are the primary cause of autosomal dominant CMT2. However we found that severe CMT2 phenotypes can be associated to recessive MFN2 mutations. The number of families with homozygous or compound heterozygous mutations in MFN2 described is increasing, so it is possible that recessive MFN2 mutations could be more frequent than expected. For this reason we propose that sequencing of MFN2 should be considered in the presence of severe CMT2 phenotype and also in the case of recessively inherited CMT2 disease. Interestingly we identified also a duplication in the PMP22 gene, thus suggesting that this gene, which is responsible for the majority of demyelinating CMT cases, can also cause axonal disease. For this reason we suggest to consider the analysis of PMP22 duplication in CMT2 patients, especially when the analysis of the other frequently-mutated genes is negative. In conclusion our data expanded the knowledge about the mutation frequencies of CMT2 associated genes and could contribute to the development of a diagnostic flow chart for CMT2 diseases.

2800W

Novel autosomal recessive syndrome characterized by manganese deficiency, cerebellar atrophy, skeletal dysplasia, and cognitive impairment caused by mutation in *SLC39A8*. *CL. Beaulieu[1], M. Galvez-Peralta[2], D. Redl[3], L. Huang[1], EG. Puffenberger[4], SC. Hampson[1], AE. Chudley[5], JN. Scott[6], RA. Hegele[7], DR. McLeod[3], AE. MacKenzie[1], DE. Bulman[1], JS. Parboosingh[3], DW. Nebert[2], AM. Innes[3], KM. Boycott[1] FORGE Canada Consortium.* 1) Children's Hospital of Eastern Ontario Research Institute, University of Ottawa, Ottawa, ON, Canada; 2) Department of Environmental Health and the Center for Environmental Genetics, University of Cincinnati Medical Center, Cincinnati, OH, USA; 3) Department of Medical Genetics and Alberta Children's Hospital Research Institute for Child and Maternal Health, University of Calgary, Calgary, AB, Canada; 4) Clinic for Special Children, Stratsburg, and Franklin and Marshall College, Lancaster, PA, USA; 5) Section of Genetics and Metabolism, Children's Hospital and the Department of Pediatrics and Child Health, University of Manitoba, Winnipeg, MB, Canada; 6) Department of Radiology, Foothills Hospital, Calgary, AB, Canada; 7) Robarts Research Institute and University of Western Ontario, London, ON, Canada.

Manganese (Mn) is an essential trace element acting as a cofactor for enzymes including mitochondrial superoxide dismutase. Mn deficiency in animal models is known to cause impaired growth, skeletal abnormalities, and ataxia. Recently, mutations in *SLC30A10* have been found to cause hypermanganesemia resulting in dystonia, parkinsonism, polycythemia, and liver disease. We have identified a homozygous mutation in *SLC39A8* (*ZIP8*) associated with manganese deficiency and cerebellar atrophy, hypotonia, strabismus, developmental delay, short stature, mild skeletal dysplasia, and connective tissue abnormalities. This novel autosomal recessive disorder was identified in three patients from three Hutterite families. An identity-by-descent mapping approach identified an associated 11.7 Mb region on chromosome 4q22-25 and a mutation, p.Gly38Arg, was identified in a conserved residue of *SLC39A8*. As a member of the solute carrier gene (SLC) superfamily, SLC39A8 transports Mn, Zn, and Cd across the plasma membrane with very high-affinity binding constants. Evaluation of trace element levels in affected patients revealed a severe biochemical defect with very low blood Mn levels (20 nM, compared to normal range of 273–278 nM) and high Mn levels excreted in the urine; the Zn deficiency was milder, likely due to redundancy (presence of other Zn transporters). To our knowledge, this is the first gene associated with a human Mn deficiency syndrome and provides insight into the role of Mn homeostasis in development and health.

2801T

GNAL is a new causative gene for Primary Dystonia. *T. Fuchs[1], R. Saunders-Pullman[2,3], I. Masuho[4], M. San Luciano[5], D. Raymond[2], S. White[1], E. Ainehsazan[1], N. Sharma[6], M.E. Ehrlich[7,8], K. Martemyanov[4], S.B. Bressman[2,3], L.J. Ozelius[1,8].* 1) Gen & Genomic Sci, Mount Sinai Sch Med, New York, NY; 2) Department of Neurology, Beth Israel Medical Center, New York, NY; 3) Department of Neurology, Albert Einstein College of Medicine, Bronx, NY; 4) Department of Neuroscience, The Scripps Research Institute, Jupiter, FL; 5) Department of Neurology, University of California, San Francisco, CA; 6) Department of Neurology, Massachusetts General Hospital, Boston, MA; 7) Department of Pediatrics, Mount Sinai School of Medicine, New York, NY; 8) Department of Neurology, Mount Sinai School of Medicine, New York, NY.

Dystonia is a movement disorder characterized by involuntary movements and abnormal postures. Primary torsion dystonia (PTD), defined clinically by the absence of neurological signs other than dystonia (or tremor) represents a complex, clinically heterogeneous group of movement disorders. Loci for 9 forms of PTD (DYT1, 2, 4, 6, 7, 13, 17, 21, and CIZ1) have been identified, most of which are inherited in an autosomal dominant (AD) manner with reduced penetrance. Using exome sequencing in two PTD families we identified a novel causative gene, GNAL, with a nonsense mutation resulting in premature stop codon in one family and a missense mutation in the other. To confirm GNAL as a dystonia causative gene and to assess the phenotypic spectrum associated with GNAL mutations, we performed Sanger sequencing of the entire coding region in more than 250 PTD subjects of mainly European origin with various PTD phenotypes. This screening revealed seven additional novel mutations in this gene. The phenotype associated with these mutations was predominated by cervical dystonia, with spread to cranial muscles in about half. Age-onset was between 10–54 years with a mainly segemental distribution. GNAL encodes the stimulatory alpha subunit, Gαolf, which links G-protein coupled receptors to downstream effector molecules and function as heterotrimers composed of alpha, beta and gamma subunits. It is expressed in the brain, with prominent enrichment in the striatum where it presumably couples both the dopamine type 1 receptors (D1R) and adenosine A2A receptors (A2AR) to the activation of adenylyl cyclase type 5 (AC5). Abundant evidence supports dysfunction of the basal ganglia and dopamine signaling in dystonia. In order to obtain mechanistic insight into the potential impact of missense mutations we studied their effect on D1R mediated Gαolf activation using bioluminescence resonance energy transfer (BRET)-based reporter system where Gαolf function could be assessed by its ability to influence the interaction between Gβγ subunits during receptor initiated cycle of nucleotide binding and hydrolysis. Several of the mutations tested in the assay reveal deficiencies in the efficiency of heterotrimer formation or coupling to D1R suggesting loss of function in the D1R-Gαolf signaling pathway is likely causative for PTD. GNAL mutations directly point to the dopamine signal transduction system as one of the origins of dystonia pathophysiology.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2802F

Analysis of Aicardi syndrome genome using array CGH and EXOME sequencing. *W. Liu[1], M. Bainbridge[2], A. Balasa[3], X. Wang[1], D. Murdock[2], R. Gibbs[2], I. Van den Veyver[1,4,5].* 1) Dept OB/GYN, Baylor Col Med, Houston, TX; 2) Human Genome Sequencing Center, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA; 3) USDA/ARS Children's Nutrition Research Center, Department of Pediatrics, Baylor College of Medicine, Houston, TX 77030, USA; 4) Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA; 5) The Jan and Dan Duncan Neurological Research Institute, Texas Children's Hospital, Houston, TX 77030, USA.

Background: Aicardi syndrome is a severe sporadic neurodevelopmental disorder that mostly affects females. It is thought to be caused by heterozygous defects in an essential X-linked gene, but a defect in an autosomal gene with sex-limited expression cannot be excluded. Despite multiple efforts, the cause of Aicardi remains unknown. To date, functional candidate-gene sequencing or genome-wide array genomic hybridization (aGH) on 244K arrays by us on 38 patients, as well as similar studies by others, have not revealed a causative copy-number variant (CNV) or candidate gene. Methods: Additional screening for CNVs in DNA from 5 Aicardi patients using a higher resolution 1M array and screening for coding-gene variants in 14 Aicardi families using whole-exome sequencing (WES) using two different capture designs was performed. Results: (1) CNV analysis: We found 182 CNVs (109 losses and 66 gains) of which 7 were loss or gain in different patients. Nine of the 66 gains were not previously reported. Only one of these contained a known gene, NXRN1, but could not be validated with quantitative PCR. No new CNVs were detected on the X chromosome. (2) WES: The first “trio” analysis on 3 Aicardi families using one capture design yielded 5 to 6 de novo sequence variants in each trio. The second WES experiment on 12 trios (1 of which was also included in the first 3) identified 5 SNPs on X chromosome, one of which in IL1RAPL1 was validated to be de novo. There were no other changes in IL1RAPL1, and the same SNP was not found in 10 other Aicardi families and 4 controls. Neither WES experiment yielded de novo variants in genes shared by all trios. Conclusion: this data, combined with those of published results, suggests that it is unlikely that Aicardi syndrome is caused by constitutional CNVs detectable on high-resolution arrays, or caused by constitutional point mutations in protein-coding genes detectable by WES.

2803W

Molecular analysis for Gaucher, Tay-Sach's and Sandhoff disease in Indian patients. *J. J. Sheth[1], M. A. Mistri[1], C. M. Ankleshwaria[1], P. Tamhankar[2], A. Bavdekar[3], CA. Datar[4], M. Kamate[5], S. Gupta[6], F. J. Sheth[1].* 1) Biochemical and Molecular Genetics, Institute of Human Genetics, Ahmedabad, Gujarat, India; 2) Genetic Research Center, NIRRH, Jehangir Merwanji Street, Parel, Mumbai 400012, Maharashtra, India; 3) K.E.M. Hospital, 489, Rasta peth, Sardar Mudliar road, Pune- 411011, India. Tele: 020-66037300; 4) Sahyadri Medical Genetics and Tissue Engineering Facility, 170/01 Barve Memorial Complex, J.M. Road, Pune- 411005. Tele: 020-25510663; 5) K.L.E.S. Dr. Prabhakar Kore Hospital & Medical Research Centre, Nehru Nagar, Belgaum, Karnataka,India.Tele :0091-0831-2473777, Fax : 0091-0831-2470732; 6) Department of Biochemistry, M.S. University, Baroda, India.

Mutations leading to Gaucher (GD), Tay-Sachs (TSD) and Sandhoff (SD) disease in India are yet unknown. Hence, we aimed to determine mutations spectrum for these disorders in Indian patients. In the present study, twenty-six, eighteen and six unrelated individuals who were clinically and biochemically confirmed to have GD, TSD and SD respectively were included. All subjects referred from various regions of India were investigated for mutation identification by ARMS, RFLP and DNA sequencing method. Out of 26 subjects with GD, L444P mutation was found in eleven (42.30%) patients, L444P/Rec NciI in three (11.53%), Rec NciI in three (11.53%) followed by R463C and R398Q in one each (7.69%) patient respectively and one patient was heterozygous for R463C, three were heterozygous for Rec NciI. Novel mutation I466S / Rec Nci I was observed in one case (3.84%), while remaining patients study is under process. In 15 of 18 subjects with TSD, six novel missense mutations were detected. The most common novel mutation c.1385A>T (p.E462V) was found in six (40%) unrelated families from Gujarat. In five (33.33%) patients novel mutations viz. c.340G>A (p.E114K), c.964G>A (p.D322N), c.964G>T (p.D322Y), c.1178C>G (p.R393P) c.1432 G>A (p.G478R) were found respectively and in one (6.66%) Iraqi patient previously reported mutation c.508C>T (p.R170W) was observed. Another known mutation in Ashkenazi Jewish, c.1274_1277 (p.Tyr427IlefsX5) was observed in homozygous and heterozygous status in one each (13.32%). In remaining three patients mutation is not found and study is under process. Of 6 subjects with SD, four novel mutations C.534_541delAGTTTATC (p.V179RfsX10), c.1563_1573delTATGGATGACG (p.M522LfsX2), c.1591_1592insC (p.R531KfsX22) and c.1417+1 G>A in one each (66.66%) respectively, one patient(16.66%) was heterozygous for c.850 C>T (p.R248X) nonsense mutation and one(16.66%) was heterozygous for c.1597C>T (p.R533C) missense mutation. These both mutations are known mutations reported in Italian and Iranian respectively. This study clearly demonstrates that L444P is the most common mutation in Indian GD patients while TSD and SD patients are more likely to have novel mutations. Nonetheless p.E462V mutation in HEXA gene seems to be more common in TSD patients from Gujarat.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2804T

Progress and challenges in elucidating the molecular genetic basis of cluster headache. *L. Southgate[1], S. Scollen[2], W. He[3], A. Moss[4], M.A. Simpson[1], B. Zhang[3], L. Xi[3], T. Schlitt[1], M. Weale[1], C.L. Hyde[3], J.C. Stephens[3], C. Sjöstrand[5], M.B. Russell[6], M. Leone[7], S.L. John[3], R.C. Trembath[1,8].* 1) Medical and Molecular Genetics, King's College London, London, United Kingdom; 2) Pfizer, Neusentis, PTx Precision Medicine, Granta Park, Cambridge, United Kingdom; 3) Pfizer, PTx Precision Medicine, Eastern Point Road, Groton, CT; 4) Pain & Sensory Disorders, Pfizer, Neusentis, Granta Park, Cambridge, United Kingdom; 5) Department of Neurology, Karolinska University Hospital Huddinge, Karolinska Institute, Stockholm, Sweden; 6) Department of Neurology, Research Centre, Akershus University Hospital, Oslo, Norway; 7) Carlo Besta National Neurological Institute, Milan, Italy; 8) Barts and The London School of Medicine & Dentistry, Queen Mary, University of London, London, United Kingdom.

Cluster headache (CH) is a trigeminal autonomic cephalalgia characterised by recurrent, relatively short-lived attacks of excruciating pain, often with regular periodicity. A genetic predisposition to CH has long been debated. Numerous epidemiological studies have reported an increased risk in relatives of CH sufferers and familial clustering supports a model of autosomal dominant inheritance with reduced penetrance. Although a few candidate gene studies have been performed, detecting moderate associations with the *HCRTR2* and *ADH4*genes, these have been limited by small sample size. We have established a large cohort of CH families, in which we have previously reported a genome-wide linkage scan. These data isolated a number of putative linkage loci; however, a single causative gene is yet to be identified, largely due to substantial genetic heterogeneity. To further delineate the genetic architecture underlying this condition, we have used an exome sequencing strategy in a subset of Northern European families. Exome capture was performed for ten probands using the Agilent SureSelect Target Enrichment System and paired-end sequence reads were generated on an Illumina Genome Analyzer IIx. After filtering of annotated variants to exclude synonymous polymorphisms and known or common variation, a total of 1711 novel heterozygous variants were detected. Exome data from related affected subjects were examined to limit the analysis to variants segregating with the CH phenotype. These variants were subsequently Sanger sequenced in all family members and those which did not segregate with affected subjects were excluded. Following segregation analysis, we have reduced our candidate list to a total of 45 genes with novel variants (range: 1–13 genes per pedigree). These genes are now being screened in our extended CH cohort to identify additional variants, providing further insight into the role of each gene in CH pathogenesis. Whilst the power of exome sequencing in rare Mendelian disease gene identification is now well-established, approaches to identify pathogenic variation with complexities such as locus heterogeneity and incomplete penetrance remain challenging. We have used a combination of exome and Sanger sequencing to detect novel coding variation segregating across CH families. The significance of such variants in CH pathogenesis remains to be elucidated; however these results now point to a potential genetic predisposition to this debilitating disorder.

2805F

Japanese Chin dogs with GM2 gangliosidosis are homozygous for a HEXA missense mutation and are a potential large animal model for Tay-Sachs disease. *D. Gilliam[1], A.C. Freeman[2], D.N. Sanders[3,6], R. Zeng[1], D.A. Wenger[4], G.S. Johnson[1], M.L. Katz[3], M. Vandenberg[2], S. Holmes[2], M. Kent[2], R. Rech[2], E. Howerth[2], S.R. Platt[2], D.P. O'Brien[5].* 1) Department of Veterinary Pathobiology, University of Missouri College of Veterinary Medicine, Columbia, MO; 2) College of Veterinary Medicine, University of Georgia, Athens, GA; 3) Mason Eye Institute, University of Missouri School of Medicine, Columbia, MO; 4) Department of Neurology, Thomas Jefferson University Jefferson Medical College, Philadelphia, PA; 5) Department of Veterinary Medicine and Surgery, University of Missouri College of Veterinary Medicine, Columbia, MO; 6) Plymouth State University, Plymouth, NH.

We evaluated two half-sibling Japanese Chin dogs which appeared to have the same GM2 gangliosidosis described in Japanese Chins over 25 years ago by Cummings et al. (Acta Neuropathol, 1985;67:247) and Ishikawa et al. (J Neurochem, 1987;48:860). Similar to the earlier disease descriptions, the neurodegeneration in the current dogs became apparent when the dogs were 11 and 16 months old and progressed rapidly over a period of 5 to 7 months until the dogs were euthanized because they showed cerebellar ataxia, blindness, severe tremors, mental dullness and disorientation. Also, consistent with the earlier reports, tissues from the affected half siblings had markedly increased concentrations of GM2 ganglioside and normal or increased β-hexosaminidase activities when measured with fluorogenic substrates, either 4-methylumbelliferyl-β-N-acetylglucosamine (MUG) or 4-methylumbelliferyl-β-N-acetylglucosamine-6-sulfate (MUGS). These biochemical findings resemble those from human patients with the AB variant of GM2-gangliosidosis, caused by mutations in GM2A. Surprisingly, the coding regions and intron-exon junctions from all four GM2A exons from the affected dogs harbored only common polymorphisms. GM2-gangliosidosis can also be caused by mutations in HEXA and HEXB, the genes that encode the respective α and β subunits of hexosaminidase A. No sequence variants were found in HEXB from the affected dogs; however, both dogs were homozygous for a HEXA:c967G>A missense mutation that predicts a p.E323K amino acid substitution. This mutation is the likely cause of the GM2-gangliosidosis in the Japanese Chins because the glutamyl moiety at codon 323 serves as the general acid/base in the catalytic site in the α subunit of hexosaminidase A. In further support of our contention that the HEXA:c967G>A missense mutation is causal, none of the 128 healthy Japanese Chins with DNA in our collection were HEXA:c967A homozygotes; instead, 123 were HEXA:c967G homozygotes and 5 were A/G heterozygotes. We believe that the β-hexosaminidase activity detected with MUGS in the affected dog tissues, but not present in tissues from comparable human patients, resulted from species-specific differences in the substrate specificity of the active site on the β subunit in hexosaminidases A and B. In conclusion, we have identified a HEXA:c967G>A missense mutation as the likely cause of the GM2 gangliosidosis in Japanese Chins, making these dogs a potential large animal model for Tay-Sachs disease.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2806W

Identification of Novel X-linked Intellectual Disability Genes by Human X Chromosome Exome Sequencing. *T. Wang[1,2], T. Niranjan[1,3], C. Skinner[4], M. May[4], L. Holloway[4], C. Schwartz[4].* 1) McKusick-Nathans Institute of Genetic Medicine, The Johns Hopkins University School of Medicine, Baltimore, MD; 2) Department of Pediatrics, The Johns Hopkins University School of Medicine, Baltimore, MD; 3) Predoctoral Training Program in Human Genetics, The Johns Hopkins University School of Medicine, Baltimore, MD; 4) Greenwood Genetic Center, Greenwood, SC.

X-linked Intellectual Disability (XLID) is a group of genetically heterogeneous disorders caused by mutations in genes on the X-chromosome. More than 90 XLID genes have been identified, which account for ~2/3 of known XLID disorders. To systematically identify genes responsible for the remaining XLID disorders, we performed next-generation sequencing of the human X-exome in 56 affected males from 41 XLID families including two affected males in 15 families. Illumina sequencing libraries were generated from genomic DNA, enriched for the human X-exome using the Agilent SureSelect system, and sequenced using a HiSeq2000. This study achieved a mean coverage of 76-fold for the X-exome and 82.1% of the baits showed sufficient coverage for all samples. Approximately 1,500 variants were identified in each X-exome, 18% of which were mapped to coding regions. To enrich candidate disease-causing mutations, we first eliminated variants that were present in dbSNP and the “1000 Genomes” database. These filters effectively reduced the number of variants for validation. However, we noted that a small number of deleterious mutations in known XLID genes exist in dbSNP. We hypothesize that this is likely the result of inaccurate sample phenotyping, limited variant validation, and inclusion of sequence data from female carriers of XLID. To overcome these problems, we sequenced two affected males in some proband families to identify shared variants and eliminate variants that were present in a cohort of unrelated, affected males (n=48). This strategy achieved a significant reduction of variants for validation and successfully identified potential deleterious mutations in known XLID genes including ATRX, PLP1, OPHN1, MECP2, and MED12, as well as in novel candidate genes. We conclude that next-generation sequencing is a powerful tool for systematic characterization of known and novel disease-causing mutations in heterogeneous disorders such as intellectual disability. Current human reference databases, including dbSNP and the “1000 Genomes”, should be used with caution as automatic filters to remove non-disease-causing variants during data analysis.

2807T

Genetic studies to gain insight into the function of the MeCP2 domains in vivo. *L. Heckman[1,2], H. Zoghbi[1,2,3].* 1) Molecular and Human Genetics, Baylor College of Medicine, Houston, TX; 2) Jan and Dan Duncan Neurological Research Institute at Texas Children's Hospital, Houston, TX; 3) Howard Hughes Medical Institute.

Rett syndrome (RTT) is a debilitating neuropsychiatric disorder caused by mutations in the X-linked methyl-CpG binding protein 2 (*MECP2*) gene, which encodes for a protein with the same name (MeCP2). Many of the RTT-causing mutations create a truncating null allele, suggesting that RTT is due to a loss of MeCP2 function. Interestingly, overexpression of MeCP2 due to a duplication spanning *MECP2* also causes a progressive neurological syndrome that shares many features with RTT. Studies in mice that either lack (RTT model) or express 2X MeCP2 (duplication model) reveal that the loss and gain of MeCP2 has opposing effects on excitatory synapses and gene expression, suggesting that the duplication causes the disorder by a “hyperfunction” mechanism. However, the challenge has been to understand exactly what function of MeCP2 is being exaggerated when it is overexpressed. Traditionally, MeCP2 was believed to be a transcriptional respressor, yet gene expression data from animal models does not support this model. Newer hypotheses proposed that MeCP2 dampens transcriptional noise, such that in its absence, increased basal transcription leads to decreased expression of neuronal genes. This model, however, does not quite explain why doubling the protein will enhance expression of activity dependent neuronal genes. Another hypothesis proposes that overexpression of MeCP2 might titrate co-repressors, thus resulting in an increase of gene expression. To determine if DNA binding by the methyl-CpG-binding domain (MBD) is necessary for the neuronal dysfunction resulting from MeCP2 overexpression, I have generated mice that overexpress *MECP2* alleles with two different RTT-causing mutations: T158M, which reduces methyl-CpG binding and R111G, which abolishes methyl-CpG binding without affecting the structure of the MBD. Transgenic mouse lines that express each allele at levels similar to endogenous levels have been generated thus allowing me to study if doubling MeCP2 levels with an allele that compromises DNA binding activity is consequential. Behavioral, physiological and molecular studies of these alleles are ongoing both in the context of overexpression and in a MeCP2 null background. These models promise to provide new insight into the mechanism by which MeCP2 overexpression leads to neurological dysfunction.

2808F

Combination of linkage analysis and exome sequencing identifies a novel form of autosomal recessive syndromic intellectual deficiency linked to DNA damage response defect. *M. Langouet[1], A. Saadi[2], C. Bole-Feysot[3], P. Nitschke[4], A. Munnich[1], M. Chaouch[2], L. Colleaux[1].* 1) INSERM U781, Fondation IMAGINE, Université Paris Descartes, Hôpital Necker-Enfants Malades, Paris, France; 2) Département de Neurologie, Etablissement hospitalier spécialisé de Benaknoun, Alger, Algérie; 3) Plateforme de Génomique de la Fondation Imagine Paris, France; 4) Plateforme de Bioinformatique de l'Université Paris Descartes, Paris, France.

We ascertained a large Algerian consanguineous multiplex family in which three affected sibs presented with severe psychomotor delay, microcephaly, behavioral disorders, skeletal anomalies, and facial dysmorphism. Linkage analysis combined to exome sequencing allowed us to identify a previously unreported homozygous variation (c.1307T>A, p.I436N) in the *TTI2/C8orf41* gene. This mutation co-segregated with the disease, was not detected in 510 control chromosomes, nor was it present in an internal exome dataset of 125 individuals or among the 5,379 exomes available from the NHLBI Exome Sequencing Project. Moreover, this variant was predicted to be damaging by various in silico tools. *TTI2* encodes a 508 amino acid protein that has been recently characterized through its interactions with the TTI1 and TEL2 proteins as part of the Triple T complex. This complex plays a key role in the folding/maturation of PIKK (phosphatidylinositol 3-kinase-related kinases) proteins and is required for DNA damage signaling and ATM and ATR stability. TTI2, as TTI1 and TEL2, is essential to the stability and the function of the TTT complex and to cell survival in response to DNA damage.Western blots showed reduced amount of TEL2, TTI1 and TTI2 proteins in patient fibroblasts as compared to controls, suggesting that the p.I436N mutation results in destabilization of the whole TTT complex. Moreover, preliminary data suggest impaired ATM and ATR stability in these cells.

ID, microcephaly, developmental delay, facial dysmorphy and skeletal abnormalities are also observed in other DNA repair disorders. In addition, another missense mutation of this gene (p.P367L) has been recently identified in a second ARID family (Najmabadi et al, Nature 2011). Altogether, these data support the hypothesis that TTI2 is the disease causing gene.

Our findings further demonstrate the efficiency of high resolution sequencing strategies to unravel the molecular bases of ID syndromes. More importantly, they provide the first example of a human disorder related to a defect in TTT complex. Further experiments are now underway to evaluate DNA damage repair efficiency in patient cells and characterize in more details the effect of this TTI2 mutation on ATM/ATR signaling.

2809W

Neuronal cells generated through differentiation of stem cells isolated from patient's skin recapitulate the main features of Niemann Pick C disease. *A. Dardis[1], N. Bergamin[2], D. Cesselli[2], S. Zampieri[1], B. Bembi[1], A. Beltrami[2], CA. Beltrami[2].* 1) Regional Coordinator Center for Rare Diseases, University Hospital Santa Maria della Misericordia, Udine, Italy; 2) Department of Medical and Biological Sciences, University of Udine, Italy.

Niemann Pick C (NPC) disease is a neurovisceral lysosomal storage disorder due to mutation in NPC1 or NPC2 genes, characterized by the accumulation of endocytosed unesterified cholesterol, gangliosides and other lipids within the lysosomes/late endosomes. The analysis of the molecular pathways linking the lipid accumulation and cellular damage in the brain has been challenging due to the limited availability of human neuronal models. The aim of this study was to develop a human neuronal model of NPC disease by inducing neuronal differentiation of stem cells isolated from skin biopsies or fibroblasts in culture of NPC patients. Primitive cells from 3 NPC patients and 3 normal controls were isolated from skin biopsies and previously established skin fibroblast cultures adapting methods previously described [Beltrami et al, 2007]. After 3 passages in selective medium, cells obtained from both sources (hSKIN-MASC) displayed an antigenic pattern, determined by FACs analysis, characteristic of mesenchymal stem cells and expressed the stem cell markers Oct-4, Nanog, Sox-2 and nestin (analyzed by immmunofluorescence). hSKIN-MASC obtained from NPC patients showed a massive lysosomal accumulation of unesterified cholesterol when analyzed by filipin staining. After the induction of neural differentiation using methods previously described [Beltrami et al, 2007], remarkable morphologic changes were observed and cells became positive to markers of the neuronal linage NeuN and MAP2. Filipin staining and immunofluorescence analysis of GM2 ganglioside showed a clear lysosomal accumulation of cholesterol and GM2 only in cells from NPC patients. Differentiated NPC cells presented a morphology that was clearly different from normal cells and that closely resemble the morphology of cortical neurons of NPC patients. A quantitative analysis using the Neurite outgrowth image program, showed that differentiated cells from NPC patients presented longer neurites and a greater number of neurite extremities, segments, roots and node points, than differentiated cells from normal donors. In conclusion, we generated a human neuronal model of NPC disease by inducing neuronal differentiation of stem cells obtained from easily accessible sources, such as patient's cultured fibroblasts, without the need to express transgenes in target cells, avoiding confounding results due to the reprogramming process. The model described here recapitulates the main features of NPC disease.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2810T

Severely impaired respiratory chain causes multisystem apoptosis-driven developmental defects, a new mitochondrial phenotype in vertebrates. *A. Indrieri[1], V. van Rahden[2], V. Tiranti[3], I. Conte[1], J. Quartararo[4], M. Morleo[1], D. Iaconis[1], R. Tammaro[1], G. Chesi[1], M. Cermola[5], R. Tatè[5], I. Maystadt[6], S. Demuth[7], A. Zvulunov[8], I. D'Amato[3], P. Goffrini[4], I. Ferrero[4], P. Bovolenta[9], K. Kutsche[2], M. Zeviani[3], B. Franco[1,10].* 1) Telethon Institute of Genetics and Medicine, Naples, Italy; 2) Institut für Humangenetik, Universitätsklinikum Hamburg-Eppendorf, Hamburg, Germany; 3) Unit of Molecular Neurogenetics, The Foundation “Carlo Besta” Institute of Neurology, Milan, Italy; 4) Departments of Genetics, Biology of Microorganisms, Anthropology and Evolution, University of Parma, Italy; 5) Integrated Microscopy Facility, Institute of Genetics and Biophysics “Adriano Buzzati Traverso”, CNR, Naples, Italy; 6) Centre de Genetique Humaine, Institut de Pathologie et de Genetique, Gosselies (Charleroi), Belgium; 7) Gemeinschaftspraxis für Humangenetik, Erfurt, Germany; 8) Schneider Children's Medical Center of Israel, Faculty of Health Sciences, Medical School for International Health, Ben-Gurion University of the Negev, Beer-Sheva, Israel; 9) Centro de Biología Molecular “Severo Ochoa”, CSIC-UAM and CIBER de Enfermedades Raras (CIBERER), Madrid, Spain; 10) Medical Genetics Services, Department of Pediatrics, Federico II University, Naples, Italy.

Intrinsic (mitochondrial) programmed cell death (PCD), plays an essential homoeostatic role, by selecting bioenergetically proficient cells suitable for normal tissue and organ development. In spite of intensive investigation on cellular and animal models, the impact of this crucial execution pathway in human disease remains mechanistically undefined. In particular, the link between apoptosis and mitochondrial disorders, i.e. primary defects of oxidative phosphorylation (OXPHOS), has not been persuasively demonstrated. On the other hand, a clearly developmental phenotype, Microphthalmia with Linear Skin lesions syndrome (MLS, OMIM 309801), is associated with mutations in *HCCS*, the X-linked gene encoding Holo-Cytochrome c-type synthase, which incorporates catalytically active heme-c moieties in the mitochondrial respiratory chain (MRC). Notably, *HCCS* mutations are present in a subset of MLS patients, the remaining ones being still undefined at the molecular level. To gain mechanistic insight on the molecular pathogenesis of the developmental defect in MLS, we first showed that, similar to MLS patients, *hccs*knockdown results in eye and brain abnormalities in medaka fish (*Oryzias latipes*). Next, we demonstrated that these defects are caused by increased PCD via apoptosome-independent caspase-9 activation, triggered by MRC impairment and overproduction of reactive oxygen species (ROS). Based on these results, we then screened MRC-related genes in *HCCS*-negative MLS patients, and found deleterious mutations in a poorly characterized structural subunit of cytochrome c oxidase (COX). These data indicate an essential role for the mitochondrial respiratory chain in organogenesis and define a new group of mitochondrial diseases hallmarked by apoptosis-driven abnormal development.

2811F

Whole Exome Sequencing Reveals Novel and Known Genes Mutations in a Cohort of Centronuclear Myopathy Patients. *P. Agrawal, O. Ceyhan, K. Schmitz, E. Dechene, M. Viola, K. Markianos, A. Beggs.* Medicine, Children's Hospital Boston, Boston, MA.

Centronuclear myopathies (CNM) are a group of rare heterogeneous congenital myopathies characterized by muscle weakness and low tone of varying onset, often needing respiratory, feeding and/or ambulatory assistance and characterized by presence of increased central nuclei on muscle biopsy. CNM is caused by mutations in the *MTM1*, *DNM2*, *RYR1* and *BIN1*. *MTM1*mutations cause a form of CNM known as X-linked myotubular myopathy (XLMTM) in about 50% of CNM cases, while *DNM2* and *RYR1* account for 15% each, and *BIN1* for <5%, leaving about 20% without a known gene. We performed whole exome sequencing (WES) in 28 CNM patients to identify known and novel gene mutations responsible for the disease. These were mostly small families with a single member diagnosed with CNM and no linkage data. *MTM1* and *DNM2* mutations were ruled out in majority of these patients, while *RYR1* and *BIN1* were not sequenced because of the large size (*RYR1*) and rarity (<5%) of the mutations (*BIN1*). WES was performed by Axeq technologies using Illumina platform and 10X coverage of the target regions was in 84–88% range for these samples. One of these 28 patients was identified to have *MTM1* duplication on clinical testing. Of the remaining 27, 2 had *DNM2* and 9 had *RYR1*mutations. Interestingly, there were 4 cases with *TTN* mutations. In summary, exome sequencing has identified molecular basis for 60% of patients with CNM. Work is underway to identify novel genes in the remaining 40% exomes.

2812W

Mutations in the human isoprenoid synthase domain containing gene are a common cause of congenital and limb girdle muscular dystrophies. *S. Cirak[1], AR. Foley[2], R. Herrmann[3], T. Willer[4], M. Yau[5], S. Torelli[2], P. Vondracek[6], R. Helen[7], C. Longman[8], R. Korinthenberg[9], G. Marruso[10], P. Nurnberg[11], UK. 10K consortium[16], V. Plagnol[12], M. Hurles[13], C. Sewry[2,15], KP. Campbell[4], T. Voit[14], F. Muntoni[2].* 1) Research Center for Genetic Medicine, Children's National Medical Center, Washington DC, DC., USA; 2) UCL Institute of Child Health, Dubowitz Neuromuscular Centre, London, UK; 3) University Hospital Essen, Essen, Germany; 4) University of Iowa, Iowa City, Iowa, USA; 5) DNA Laboratory, GSTS Pathology, London, UK; 6) University Hospital Brno, Czech Republic; 7) Birmingham Heartlands Hospital, UK; 8) Yorkhill Hospital, Glasgow, UK; 9) Albert-Ludwigs University, Freiburg im Breisgau, Germany; 10) University of Cagliari, Sardinia, Italy; 11) Cologne Center for Genomics, Germany; 12) UCL Genetics Institute, UCL, London, UK; 13) Wellcome Trust Sanger Institute, UK; 14) Institut de Myologie, Paris, France; 15) Wolfson Centre for Inherited Neuromuscular Disorders, RJAH Orthopaedic Hospital, Oswestry, UK; 16) http://www.uk10k.org/consortium.html.

The secondary dystroglycanopathies are a recessively inherited diverse group of muscular dystrophies ranging in severity from congenital onset with severe cobblestone lissencephaly (Walker-Warburg Syndrome) to milder forms of limb girdle muscular dystrophy (LGMD). Their hallmark is the reduction of α-dystroglycan (ADG) glycosylation and an important part of this glycosylation is a unique O-mannosylation which regulates the interaction between ADG and extracellular matrix proteins including laminin alpha 2, perlecan and agrin. The genetic aetiologies of approximately 50% of the dystroglycanopathies have been attributed to nine genes that are involved in ADG glycosylation. Despite collaborative efforts of several international groups using traditional positional cloning, the causative genes for the unsolved dystroglycanopathy cases had escaped discovery. Recently biallelic loss of function mutations in the *isoprenoid synthase domain containing* (*ISPD*) have been described in Walker-Warburg syndrome. In a collaborative study with the UK10K Rare Disease Project from the Sanger Institute, using whole exome sequencing techniques, we have now discovered that mutations in *ISPD* are found in a significant proportion of ADG patients. The mutations identified are milder mutations, and the phenotype of affected patients ranges from muscle-eye-brain like disorders, to congenital muscular dystrophy without brain involvement and to limb girdle muscular dystrophy with or without evidence of brain involvement. While in eubacteria and in choloroblasts of higher plants ISPD orthologues are involved in the synthesis of isoprenoid precursors, this pathway is absent in eukaryotes and vertebrates and; therefore, the function of ISPD remains unknown. The mechanism of *ISPD* mutations leading to reduction of α-dystroglycan glycosylation needs to be uncovered; however, our ongoing screening shows that ISPD mutations are a common cause of dystroglycanopathies.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2813T

Prevalence of MECP2 mutations among girls presenting significant clinical overlap between Rett and Angelman syndromes. *T. M. R. de Assis[1], G. A. Molfetta[1,2], C. A. Ferreira[1], C. H. P. Grangeiro[3], W. A. Silva, Jr.[1,2].* 1) National Institute of Science and Technology in Stem Cell and Cell Therapy, Regional Blood Center of Ribeirão Preto; 2) Departament of Genetics, Medical School of Ribeirão Preto, University of São Paulo; 3) Medical Genetics Unit, Hospital of the Medical School of Ribeirao Preto, University of Sao Paulo.

Rett Syndrome (RS) is a developmental disorder caused by mutations in the MECP2 gene located on X chromosome, which affects females with an incidence of 1:10.000-1:15.000 newborns. Patients can present loss of spoken language, motor dysfunction, most notably hands-wash movements and gait disturbance, scoliosis, microcephaly, seizures, cardiac abnormalities, breathing disorder and mental retardation. These symptoms arise in 6 to 18 months old girls, after an apparently normal development. Angelman syndrome (AS) is a neurobehavioral disorder with an incidence of 1:10.000 newborns caused by defects of the UBE3A gene in chromosome 15q11–13. The syndrome is clinically characterized by central congenital hypotonia, delayed neuropsychomotor development, severe mental retardation, total lack of speech, excessive laughter, hyperactivity and dysmorphic features such as micro-brachycephaly, macrostomia, widely-spaced teeth, lingual protrusion, mandibular prognathism, strabismus and seizures. Both conditions can be difficult to diagnose in girls early in their development and the distinction from each other can be challenging due to both disorders being characterized by severe mental disability, microcephaly, speech disturbance, movement disorders with gait and/or truncal ataxia, and occasionally a similar facial appearance. This work focus on identifying and characterizing mutations along the coding region of the MECP2 gene in patients clinically diagnosed as RS or AS (biparental methylation pattern and no mutations in the UBE3A gene). DNA was extracted from blood samples after parents signing the informed consent form, and High Resolution Melting(HRM) technique and DNA sequencing were employed to screen for mutations. While searching for mutations in MECP2 exons, we have found a missense mutation (T158M) (GenBank X99686) in a patient previously diagnosed as RS and no alterations whatsoever in the group of patients clinically diagnosed as AS. The methodology is robust and efficient for detecting mutations in the MECP2 gene,confirming the diagnosis of patients suspected as RS, a subdiagnosed pathology due to its complex clinical suspicion. Furthermore, we postulate that the clinical diagnosis should be rigorously conducted in order to correctly screen patients to the appropriate laboratorial molecular approach. Genetic counseling was offered to the families carrying mutations. Financial Support: FAPESP (2011/11450-1), CTC, INCTC.

2814F

Identification of six novel *NF1* mutations in neurofibromatosis type 1 patients from the Azorean island of São Miguel (Portugal). *L. Mota-Vieira[1,2], B. Campos[3], R. Cabral[1,2], A.L. Rodrigues[4], A. Raposo[4], A. Carvalho[4].* 1) Mol Genetics & Pathol Unit, Hosp Divino Espirito Santo PD, Ponta Delgada, Azores Islands, Portugal; 2) Azores Genetics Research Group, Instituto Gulbenkian de Ciência, Lisbon, Portugal; 3) Molecular Genetic Diagnosis Centre, Bellvitge Biomedical Research Institute, Barcelona, Spain; 4) Pediatric Department, Hosp Divino Espirito Santo PD, Ponta Delgada, Azores Islands, Portugal.

Neurofibromatosis type 1 (NF1, OMIM 162200) is a common autosomal dominant complex disorder, affecting about 1 out of 3000 live births, regardless of family history, race, gender, or ethnic background. To date, 1163 different *NF1* mutations have been identified (HGMD accessed: May 31, 2012); most of them are unique to a particular family. In the present study, we describe the germline *NF1* mutation spectrum identified in a cohort of 13 patients, 11 unrelated and 2 first cousins, from the Azorean Island of São Miguel (137,830 inhabitants), located on the North Atlantic Ocean. All patients present the clinical features of neurofibromatosis type 1: café-au-lait spots (13, 100%), axillary and/or inguinal freckling (12, 92.3%), neurofibromas (6, 46.2%), optic glioma (2, 15.4 %) and/or Lisch nodules (2, 15.4 %). *NF1* mutation screening was performed using a cDNA-direct sequencing approach complemented with multiplex ligation-dependent probe amplification (MLPA) analysis for the detection of single and multiexon deletions or duplications. Until now, we identified the causative mutations in 10 patients: six are novel, and four are recurrent mutations. The novel mutations fall in the coding region (NM_000267.3): five are frameshift mutations (c.301delA in exon 4a, p.Thr101GlnfsX2; c.1498_1500delATTinsC in exon 10b, p.Ile500ProfsX10; c.2033dupC in exon 13, p.Ile679AspfsX21; c.2739dupA in exon 16, p.Arg914ThrfsX5; and c.5450_5451delGT in exon 29, p.Ser1817TyrfsX23), and one is a missense mutation (c.880A>G in exon 6, p.Met294Val that is predicted to affect neurofibromin function). The recurrent mutations reported so far consist of a complete deletion of the NF1 gene (Kayes LM et al., 1994), found in two patients, the nonsense mutation c.3721C>T in exon 22 (p.Arg1241X; Fahsold R et al., 2000) and the missense mutation c.1658A>C in exon 11 (p.His553Pro; Griffiths S et al., 2007). We are now finishing the analysis of the three other patients. If no pathogenic mutation is identified we will perform *SPRED1* mutation screening, since loss-of-function mutations in this gene cause a neurofibromatosis 1-like phenotype (Brems H et al., 2007). Finally, the epidemiological analysis revealed São Miguel Island (Azores, Portugal) has one of the highest incidence of neurofibromatosis type 1 in the world, approximately 1:2000 live births (lmotavieira@hdes.pt; Funding by the Government of the Azores).

2815W

A Pan-European study of the pathological C9orf72 hexanucleotide (G4C2) expansion associated with frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *J. van der Zee[1,2], I. Gijselinck[1,2], L. Dillen[1,2], T. Van Langenhove[1,2], M. Cruts[1,2], C. Van Broeckhoven[1,2], the Belgian Neurology consortium and the European Early-Onset Dementia consortium.* 1) Neurodegenerative Brain Diseases Group, VIB Department of Molecular Genetics, Antwerp, Belgium; 2) Institute Born-Bunge, University of Antwerp, Antwerp, Belgium.

Three studies independently identified C9orf72 on chromosome 9p21 as the causal gene associated with frontotemporal lobar degeneration (FTLD) and amyotrophic lateral sclerosis (ALS). The primary gene defect is an outsized, pathological expansion of a hexanucleotide repeat sequence, G4C2, located in the C9orf72 promotor region. We developed an alternative repeat-based PCR assay on the sense strand that eliminates interference of a low complexity sequence located adjacent to the G4C2 repeat. This reverse repeat-based PCR assay combined with a bridging STR fragment length assay allowed reliable identification of carriers of a pathological G4C2 expansion as well as sizing of the normal repeat alleles. Subsequently, we determined the geographical prevalence of the C9orf72 expansion in a sizable pan-European cohort of FTLD and ALS patients, ascertained within a Belgian Neurology consortium and in a newly founded European consortium created to study early-onset dementia, the European EOD consortium. DNA and clinical information of a total of 1489 patients was collected, consisting of 1081 FTLD, 83 FTLD-ALS and 325 ALS patients. This allowed robust estimations of the prevalence and geographical distribution of the C9orf72 expansion in Europe. We calculated overall frequencies of 6% in FTLD, 33% in FTLD-ALS, and 10% in ALS. The frequencies increased in familial patients to 10% in FTLD, 57% in FTLD-ALS, and 38% in ALS. The C9orf72 expansion showed an age-dependent penetrance reaching 90% by the age of 70 years. In conclusion, the high frequency of the pathological C9orf72 expansion in FTLD and ALS patients makes it a primary mutation for clinical genetic testing. Hereto, the combined reverse repeat-based PCR and STR fragment length assays offer a reliable screening method with high analytical and clinical validity.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2816T

Identification of novel ALS genes using linkage analysis and next-gen sequencing. *K.L. Williams[1,2], J.A. Solski[1], J.C. Durnall[1], A.D. Thoeng[1,3], S. Warraich[1,2], G. Rouleau[4], G.A. Nicholson[1,2,5], I.P. Blair[1,2].* 1) Northcott Neuroscience, ANZAC Research Institute, Concord, NSW, Australia; 2) Sydney Medical School, University of Sydney, NSW, Australia; 3) Department of Physiology, University of Sydney, NSW, Australia; 4) Ste-Justine Hospital Research Centre, Montreal, Canada; 5) Molecular Medicine Laboratory, Concord Hospital, NSW, Australia.

Amyotrophic lateral sclerosis (ALS) is a fatal neurodegenerative disorder that causes the progressive degeneration of motor neurons. Familial ALS accounts for approximately 10% of ALS cases with the remainder being sporadic. ALS is genetically heterogeneous. To date, known genes account for ~55% of familial cases. We aim to investigate known ALS genes and identify new ALS genes in a large cohort of Australian ALS families (n=187) using a combination of sequencing, traditional genetic linkage approaches and next-generation sequencing strategies. We also aim to implement a user-friendly pipeline for analysis of exome sequencing data. We analysed 187 ALS families for mutations in known ALS genes and determined that they account for 57.2% of Australian ALS families, and comprise SOD1 (13.9%), FUS (2.7%), TARDBP(2.1%), UBQLN2 (1.1%) and C9ORF72 (38.0%) mutations. To identify new loci for familial ALS, an 8cM genome-wide microsatellite linkage scan was performed on 51 individuals (affected, unaffected and obligate carriers) from two large informative families negative for all known ALS genes. Previous linkage analysis was performed on one of these families using a 10K Affymetrix SNP chip. Subsequent analyses have yielded suggestive linkage to several chromosomal regions. To reduce sequencing burden when new ALS genes are identified, a proband from each of our remaining 79 families was subjected to exome capture and sequencing (Agilent capture-SOLiD4 sequencing or Illumina TruSeq capture-HighSeq2000 sequencing). Bioinformatic analysis has required the development of a user-friendly pipeline. The genetic defects are yet to be identified among 42.2% of ALS families (79/187 families) within our cohort. The chromosomal regions implicated from our genome-wide linkage scans do not overlap previously identified loci, implicating substantial genetic heterogeneity. Linkage analysis in combination with exome capture and sequencing has allowed us a greater opportunity to identify novel ALS genes. Two candidate genes are currently being validated in extended patient and control cohorts, patient tissues and functional studies. The identification of these novel ALS genes will give insights into the biological basis of both familial and sporadic motor neuron degeneration, allow development of new disease models and provide new targets for therapeutic development.

2817F

The distributions of two SOD1 missense mutations in the pet dog population and their association with canine degenerative myelopathy, a model for amyotrophic lateral sclerosis. *R. Zeng[1], J. R. Coates[2], L. Hansen[1], G. C. Johnson[1], F. A. Wininger[2], M. L. Katz[3], G. S. Johnson[1].* 1) Department of Veterinary Pathobiology, University of Missouri College of Veterinary Medicine, Columbia, MO, USA; 2) 2Department of Veterinary Medicine and Surgery, University of Missouri College of Veterinary Medicine, Columbia, MO, USA; 3) Mason Eye Institute, University of Missouri School of Medicine, Columbia, MO, USA.

Canine Degenerative Myelopathy (DM) is a naturally occurring, progressive, adult-onset neurodegenerative disease and a potential large animal model for amyotrophic lateral sclerosis. The initial signs of DM include asymmetric general proprioceptive ataxia and spastic paresis in the pelvic limbs. Over a period of 9 to 12 months, affected dogs become nonambulatory paraplegic with proprioceptive deficits followed by lower motor neuron signs. Owners usually elect euthanasia at this stage. If euthanasia is delayed, clinical signs progress to flaccid tetraplegia and widespread muscle atrophy and may involve the brain stem with signs of dysphagia. Previously, we reported an association between DM and homozygosity for a c.118G>A missense mutation in the canine ortholog of SOD1 in five dog breeds. In addition, we have reported a case of DM in a single Bernese Mountain Dog that was homozygous for a SOD1:c.52G>T missense mutation. For the current report we used TaqMan allelic discrimination assays to estimate the breed-specific distributions of the SOD1:c.118A and SOD1:c.52T alleles in the canine pet populations. So far we have genotyped 26,898 dogs at SOD1:c.118G>A and 1,577 dogs at SOD1:c.52A>T. Representatives from 217 different dog breeds were genotyped at SOD1:c.118G>A and 115 of them were found to segregate the SOD1:c.118A allele. The Wire Fox Terrier breed had the highest SOD1:c.118A allele frequency: 0.94. Of the 60 breeds examined at SOD1:c.52A>T, the SOD1:c.52T allele was only found in members of the Bernese Mountain Dog breed, in which the allele frequency among the 884 genotyped breed members genotyped by us was 0.032. Only 2 Bernese Mountain Dogs were homozygous for the SOD1:c.52T allele: one has been euthanized because of DM, the other one is currently too young to develop signs. We also identified 24 Bernese Mountain Dogs that were SOD1:c.52A,c.118A/c.52T,c.118G compound heterozygotes. Pertinent clinical information was available for 6 of the 9 compound heterozygotes over 8 years of age. Four of the compound heterozygous Bernese Mountain Dog have developed clinical signs suggestive of DM; the other two died from unrelated diseases before their tenth birthday. In addition, we have examined 6 dogs from four different breeds that had histopathologically confirmed DM and were A/G heterozygotes at SOD1:c.118, but had no other SOD1 sequence variants. Finally, one German Shepherd Dog without any SOD1 sequence variants had histopathologically confirmed DM.

2818W

Explore the impact of parkin mutations in ER-associated protein degradation: cellular models for juvenile PD. *M. Alves-Ferreira[1], S. Morais[1], J. Sequeiros[1,2], I. Alonso[1,2].* 1) UnIGENe, IBMC, Porto, Portugal; 2) ICBAS, University of Porto, Portugal.

Parkinson disease (PD) is the second most common neurodegenerative disorder and is mainly characterized by motor dysfunction, as the result of selective loss of dopaminergic neurons. Although the aetiology and pathogenesis in PD are still poorly understood, central players causing rare genetic forms of PD should prove useful to understand idiopathic PD. Among these is an Autosomal Recessive-Juvenile Parkinson's disease (AR-JP) form, showing symptoms resembling idiopathic PD, although being characterized by a young age-at-onset (usually under 40). This disorder is caused by mutations in the PARK2 gene which encodes parkin, an E3-ubiquitin ligase that specifically recognizes its substrate protein, promoting its ubiquitination and subsequent degradation by the proteasome. Parkin misfolding seems to be the major mechanism for parkin inactivation. Recent observations support the existence of escorting machinery, which connects the ubiquitylated substrates at the ER-membrane with the 26S proteasome, in a process known as ER-associated degradation (ERAD). Also, it has been shown that misfolded proteins and aggregates, even when involving cytosolic proteins, are able to induce ER stress, probably by suppressing proteasome function, an essential component of the ERAD machinery. Our aim is to explore the role of ERAD in mutant parkin degradation by increasing ERAD capacity. To accomplish this SH-SY5Y cell lines transiently transfected with wild-type and mutant parkin as well as with or without VCP/p97, a key component of the ERAD machinery. We have selected 9 different mutations located in the different parkin domains that were introduced in a GFP-fused parkin construct. The basic characterization through florescence microscopy shows that two mutants (N52MfsX29 located in the ubiquitin-like domain and R275W located in the RING1 domain) are highly prone to aggregation. We are now exploring the impact VCP/p97 overexpression in aggregate formation/degradation. With this study we hope to provide further clues about the role of the ER in parkin-associated PD and in particular on the impact of the different mutations in the ERAD.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2819T

A nonsense mutation in DHTKD1 causes Charcot-Marie-Tooth disease type 2 in a large Chinese pedigree. *M. Gu[1], W. Xu[1], L. Sun[1], W. Guo[1], J. Ma[2], S. Chen[2], W. Huang[3], Z. Wang[1,4].* 1) Department of Medical Genetics, Shanghai Jiao Tong University School of Medicine (SJTUSM), Shanghai, China; 2) Department of Neurology and Institute of Neurology, Rui-Jin Hospital Affiliated to SJTUSM, Shanghai, China; 3) Chinese National Human Genome Centre at Shanghai, Shanghai, China; 4) Shanghai Research Centre for Model Organisms, Shanghai, China.

Charcot-Marie-Tooth (CMT) disease represents a clinically and genetically heterogeneous group of inherited neuropathies caused by molecular defects in more than 40 different genes. Based on clinical manifestation and electrophysiological properties, CMT can be classified into autosomal-dominant demyelinating (CMT1) and axonal (CMT2), X-linked (CMT1X), and autosomal-recessive neuropathies. Here, we report a five-generation family of 8 affected individuals with CMT2. Genome-wide linkage analysis showed no linkage to any known loci for CMT2. However, the disease phenotype was found to be closely linked to chromosome 10p13–14, the region spanning 5.41 Mb between D10S585 and D10S1477. DNA sequencing of mutations in 20 candidate genes in this region revealed a nonsense mutation c.1455T>G (p.Tyr485X) in exon 8 of dehydrogenase E1 and transketolase domain containing 1 (DHTKD1) in all 8 patients, but not in other unaffected individuals in this family and 250 unrelated normal persons. DHTKD1 mRNA level in peripheral blood of patients was found to be reduced to half of that in unaffecteds. In vitro studies showed that DHTKD1Tyr485X mRNA and truncated protein significantly decreased with rapid mRNA decay in transfected cells when compared to wild-type DHTKD1. Inhibition of nonsense-mediated mRNA decay (NMD) by UPF1 silencing was found to effectively rescue the decreased levels of mutant mRNA and protein. More importantly, both DHTKD1 silencing and overexpression of DHTKD1Tyr485X led to impaired energy production evidenced by decreased ATP, total NAD+ and NADH, and NADH levels. Thus, we identified a previously uncharacterized mutation in DHTKD1 as one of the causes of CMT2, implicating an important role of DHTKD1 in mitochondrial dynamics and neurological development.

2820F

Identification of the SPG46 gene by next generation sequencing in four families with autosomal recessive complicated spastic paraplegia. *G. Stevanin[1,2,3], R. Schule[4], K. Smets[5], E. Martin[1,3], S. Zuchner[6], L.J. Loureiro[7], A. Boukhris[8], A. Rastetter[1,3], T. de Coninck[5], M. Gonzalez[6], E. Mundwiller[1], F. Lamari[9], G. Gyapay[10], P. Coutinho[7], C. Mhiri[8], P. de Jongue[5], E. Kabashi[1], A. Brice[1,2].* 1) INSERM / UPMC UMR_S975 - NEB, Paris, France; 2) AP-HP, Groupe Hospitalier Pitié-Salpêtrière, Département de Génétique et Cytogénétique, Paris, France; 3) Ecole Pratique des Hautes Etudes, Paris, France; 4) University of Tuebingen, Germany; 5) University of Antwerp, Antwerpen, Belgium; 6) University of Miami, USA; 7) UnIGENe and Centro de Genética Preditiva e Preventiva, Institute for Molecular and Cellular Biology, Porto, Portugal; 8) Service de Neurologie, Hôpital Universitaire Habib Bourguiba Sfax, Tunisia; 9) APHP, Groupe Hospitalier Pitié-Salpêtrière, Service de Biochimie, Paris, France; 10) Genoscope, Evry, France.

Hereditary spastic paraplegias (HSP) are genetically and phenotypically heterogeneous disorders with various clinical profiles and modes of transmissions. The SPG46 locus was first reported to account for a rare form of complicated autosomal recessive spastic paraplegia mapped to chromosome 9. Using the next generation sequencing, we identified four different mutations in GBA2, including three truncating and one missense variants, that were found to segregate with the disease in four families. The overall phenotype was a complex HSP with mental impairment, cataracts and hypogonadism in males associated with thinning of the corpus callosum at brain MRI. Antisense morpholino oligonucleotides targeting the zebrafish GBA2 orthologous gene led to abnormal morphology and motor behaviour that were rescued by using the human wild type mRNA but not with the same mRNA carrying the missense change identified in two families. GBA2 encodes a microsomal non-lysosomal glucosylceramidase, an enzyme that catalyzes the conversion of glucosylceramide to free glucose and ceramid and the hydrolysis of bile acid 3-O-glucosides. The identification of the SPG46 gene is the starting point for the elucidation of the mechanisms underlying axonal degeneration in this complicated form of HSP and highlights lipid metabolism in these diseases.

2821W

A balanced (1q32;14q31) translocation associated with epilepsy, learning disability and leukaemia has breakpoints in developmental genes as detected by whole-genome sequencing. *M.I. Rees[1], S. Ali[1], R.H. Thomas[1], J.G. Mullins[1], M. Kerr[2], S.K. Chung[1].* 1) Neurology Research & Molecular Neuroscience, College of Medicine, Swansea, United Kingdom; 2) Centre for Learning Disability, Cardiff University, Cardiff, United Kingdom.

Next generation sequencing (NGS) technologies are rapidly evolving human genomic research and in the continual hunt for disease-causing genes. In a proof-of-principle experiment, whole-genome sequencing was used to identify the physical breakpoints in a constitutional translocation involving chromosomes 1q32 and 14q31 in an index-case with epilepsy, learning disabilities (LD) and a familial history of aggressive, treatment-resistant acute myeloid leukaemia (AML). Whole-genome sequencing revealed the exact translocation breakpoint confirming a gene-fusion between coding regions on chromosome 1 and chromosome 14 and was validated by Sanger sequencing. The gene on chromosome 1 has an unknown function and is poorly documented in the literature, whilst the gene on chromosome 14 is a neurodevelopmental regulator of the cell cycle and belongs to the FOX gene family. There is potential for the disrupted genes to have synergistic multi-system consequences and the coding regions of both genes have been used to mutation-screen panels of idiopathic epilepsy, learning disabilities, and cytogenetic-negative AML cases. We will present the exact nature of the gene fusion and the outcomes of the mutational scanning in the core phenotype presentations and the preliminary molecular characteristics of any fusion protein possibilities. These findings have predictive value for the at-risk members of the family and presented 2 new genes that may have impact in neurodevelopmental disorders and AML predisposition.

2822T

A *DYNC1H1* mutation causes a quadriceps-dominant neurogenic muscular atrophy. *S. Saitoh[1], Y. Tsurusaki[2], K. Tomizawa[3], A. Sudo[4], N. Asahina[5], H. Shiraishi[5], J. Ito[6], H. Tanaka[7], H. Doi[2], H. Saitsu[2], N. Miyake[2], N. Matsumoto[2].* 1) Dept Pediatrics, Nagoya City Univ, Nagoya, Japan; 2) Dept Human Genetics, Yokohama City Univ, Yokohama, Japan; 3) Dept Pediatrics, Nakashibetsu Town Hosp, Nakashibetsu, Japan; 4) Dept Pediatrics, Sapporo City General Hosp, Sapporo, Japan; 5) Dept Pediatrics, Hokkaido Univ, Sapporo, Japan; 6) Dept Pediatrics, Taiyo no sono, Date, Japan; 7) Dept Pediatrics, Asahikawa Habilitation Center, Asahikawa, Japan.

Exome sequencing unexpectedly identified a *DYNC1H1* mutation (c.917A>G; p.H306R) in a family presenting with a unique quadriceps-dominant form of neurogenic muscle atrophy. Two half sibs showed motor developmental delay and non-progressive muscle weakness and wasting predominantly in their quadriceps femoris muscle with different severity. Their mother was most mildly affected and only showed difficulty in squatting. Distal lower limbs were relatively preserved and upper limbs were not affected. Sensory involvement was indistinct. Muscle CT revealed muscle atrophy predominantly in quadriceps femoris muscle. Nerve conductive study revealed normal results in two sibs, while electromyography showed neurogenic change. Muscle biopsy from quadriceps femoris muscle of the elder sib showed grouping atrophy, presence of enlarged muscle fibers and fiber type grouping consistent with neurogenic etiology. The same mutation in *DYNC1H1* was recently reported in a pedigree with the axonal form of Charcot-Marie-Tooth disease, while three missense mutations (p.I584L, p.K671E, p.Y970C) were identified in familial patients with dominant spinal muscular atrophy with lower extremity predominance (SMA-LED). Our patients showed consistent clinical features with SMA-LED. Two de novo missense mutations (p.E1518K, p.H3822P) have also been identified in patients with severe intellectual disability and variable neuronal migration defects. Our study has confirmed that the same *DYNC1H1* mutation found in distal neuropathy also causes predominant motor neuron involvement, and suggests broad phenotypic effects of *DYNC1H1* mutations.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2823F

Late-onset Charcot-Marie-Tooth disease 4F caused by periaxin gene mutation. *S. Tokunaga[1], A. Hashiguchi[1], A. Yoshimura[1], K. Maeda[2], T. Suzuki[3], H. Haruki[4], T. Nakamura[1], H. Takashima[1].* 1) Dept Neurol Genetics, Kagosima university, Japanese Society of Neurology, Kagoshima, Japan; 2) Dept Neurol, National Hospital Organization Shiga hospital, Japanese Society of Neurology, Shiga, Japan; 3) Dept Neurol, joetsu General Hospital, Japanese Society of Neurology, Niigata, Japan; 4) Dept Neurol, Yamaguchi University Gtaduate School, Japanese Society of Neurology, Yamaguchi, Japan.

We identified the main features of Charcot-Marie-Tooth (CMT) disease, type 4F, caused by a periaxin gene (PRX) mutation in Japanese patients. Periaxin is known as one of the key myelination molecules, forming tight junction between myelin loop and axon. We collected 427 DNA samples from individuals with CMT or CMT-related neuropathy, negative for PMP22 duplication. We investigated PRX mutations using a purpose-built resequencing array screen during the period 2006–2012. We detected two types of PRX mutations in three patients; one patient showed a novel homozygous p.D651N mutation and the other two showed homozygous p.R1070X mutation. All PRX mutations reported so far have been of nonsense or frameshift type. In this study, we found homozygous missense mutation p.D651N. Aspartate 651 is located in a repeat domain; its position might indicate an important function. PRX mutations usually lead to early-onset, autosomal-recessive demyelinating Charcot-Marie-Tooth neuropathy 4F (CMT4F) or Dejerine-Sottas disease; their clinical phenotypes are severe. In our three patients, the onset of the disease was at the age of 27 years or later, and their clinical phenotypes were milder compared with those reported in previous studies. We showed a variation of clinical phenotypes for CMT4F caused by a novel, nonsense PRX mutation. .

2824W

Detection of slipped-DNAs at the trinucleotide repeats of the myotonic dystrophy type I disease locus in patient tissues. *M.M. Axford[1,2], Y-H. Wang[3], M. Nakamori[4], M. Zannis-Hadjopoulos[5], C. Thornton[4], C.E. Pearson[1,2].* 1) Genetics and Genome Biology, Hospital for Sick Children, Toronto, Ontario, Canada; 2) Department of Molecular Genetics, University of Toronto,Toronto, Ontario, Canada; 3) Department of Biochemistry, Wake Forest University School of Medicine, Winston-Salem, NC, USA; 4) Department of Neurology, University of Rochester School of Medicine and Dentistry, Rochester, New York, USA; 5) Goodman Cancer Research Centre and Department of Biochemistry, McGill University, Montreal, QC, Canada.

Slipped-strand DNAs, formed by out-of-register mispairing of repeat units on complementary strands, were proposed over 50 years ago as transient intermediates in repeat length mutations, hypothesized to cause at least 30 neurodegenerative diseases. While slipped-DNAs have been characterized in vitro, evidence of slipped-DNAs at an endogenous locus in biologically relevant tissues, where instability varies widely, is lacking. The largest expansions typically arise in non-mitotic tissues such as cortex and heart. Here, using an anti-DNA junction antibody and immunoprecipitation, we identify slipped-DNAs at the unstable trinucleotide repeats (CTG)n•(CAG)n of the myotonic dystrophy disease locus in patient brain, heart, muscle and other tissues. Slipped-DNAs were present in chromatinized DNA. We characterized slipped-DNAs as clusters of slip-outs along a DNA, with each slip-out having 1–100 extrahelical repeats, where their levels correlate with the degree of repeat instability between tissues. Surprisingly, this supports the formation of slipped-DNAs as persistent products of repeat instability, and not merely as transient intermediates. These findings further our understanding of genetic variation and have prognostic implications.

2825T

A lyonization study in heterozygous females for X-linked Adrenoleukodystrophy. *F.S. Pereira[1], C. Habekost[2,3], U. Matte[1,3], L.B. Jardim[2,3,4,5].* 1) Gene Therapy Center, Experimental Research Center, Hospital de Clínicas de Porto Alegre, Rio Grande do Sul, Brazil; 2) Medical Genetics Service, Hospital de Clínicas de Porto Alegre, Brazil; Post-Graduation Program in Genetics and Molecular Biology, Universidade Federal do Rio Grande do Sul, Brazil; 3) Post-Graduation Programs in Genetics and Molecular Biology, Universidade Federal do Rio Grande do Sul, Brazil; 4) Post-Graduation Programs in Medical Sciences, Universidade Federal do Rio Grande do Sul, Brazil; 5) Department of Internal Medicine, Hospital de Clínicas de Porto Alegre, Brazil.

The X-linked Adrenoleukodystrophy (X-ALD - OMIM 300100) is a genetic disorder of metabolism of peroxisomes, where the degradation of very long saturated fatty acids (VLCFA) is prevented or very limited. X-ALD is caused by a defect in the gene for the adenosine triphosphate (ATP)-binding cassette protein, subfamily D, member 1 (*ABCD1*) located on Xq28. X-ALD affects mainly the adrenal cortex, the myelin of the central nervous system and the central and peripheral axons of affected males. For unclear reasons, many heterozygous women develop a neurological condition in adulthood. The main hypothesis relates this to chromosome X lyonization. Aim: to assess the pattern of chromosome X inactivation in heterozygous women for X-ALD thereby identifying the predominantly expressed allele in women carriers. Methods: We selected heterozygous women belonging to families X-ALD accompanied in Medical Genetics Service from Hospital de Clínicas de Porto Alegre. Initially the *HUMARA* locus was amplified by PCR for identification of informative women (*HUMARA* carriers of alleles with different sizes). The informative samples were digested with *HpaII* enzyme and the chromosome X inactivation (CXI) skewness was calculated. We used a ratio of the area of the smaller peak (digested allele) to the sum of the areas of the smaller and larger peak (non digested allele). Results: Fifty-four women were included from 28 X-ALD families with at least one heterozygous woman molecularly identified. Among them, 90.7% (49/54) were informative for the *HUMARA* locus. Of the 17 women studied to date, 47% (8/17) showed a extremely skewed pattern of CXI (<0.1). Discussion: These data seem to indicate that most women heterozygous for X-ALD show a skewed pattern of lyonization. If this pattern will be related to phenotypic variability, remains to be studied.

2826F

Mutations in CUL4B, a member of the Cullin-RING ubiquitin ligase complex, cause syndromic X-linked intellectual disability associated with short stature. *A.K. Srivastava[1,3], M. Cook[1], K. Corrier[2], L. Lee[2], S.S. Cathey[1], F. Bartel[1], K.R. Holden[1], C. Skinner[1], A.D. Chaubey[1], B.R. DuPont[1], N. VerBerkmoes[2], R.J. Schroer[1].* 1) J. C. Self Research Institute of Human Genetics, Greenwood Genetic Center, Greenwood, SC; 2) Chemical Science Division, Oak Ridge National Laboratory, Oak Ridge, TN; 3) Department of Genetics & Biochemistry, Clemson University, Clemson, SC.

The ubiquitin-proteasome protein degradation pathway plays a major role in regulating expression levels of specific proteins in cells, and aberrations in this pathway have been implicated in the pathogenesis of several neurodevelopmental disorders, including intellectual disability (ID). Recently, we and others have identified mutations in CUL4B, a member of the Cullin-RING ubiquitin ligase complex, in patients with X-linked intellectual disability (XLID) syndrome. Here we report a novel CUL4B nonsense mutation (p.L163X) in two affected brothers with ID, grand mal seizures, relative macrocephaly, and short stature. In addition, we identified a novel copy number gain of approximately 2 kb involving the genomic region within the CUL4B gene in a male patient with severe cognitive impairment, short stature, behavioral issues and distinct facial features. Both alterations were inherited from phenotypically normal mothers who had highly skewed X-inactivation and expressed only the wild type CUL4B allele. The male patient with the duplication involving the CUL4B gene showed significantly reduced quantity of CUL4B transcript but no detectable wild type CUL4B protein. The CUL4B mutant allele in one of the affected male with the nonsense mutation was not completely degraded by the nonsense-mediated decay mechanism and produced a truncated CUL4B protein retaining the predicted nuclear localization signal but lacking the C-terminal cullin domain. In cultured neuronal cells, the truncated CUL4B protein was primarily localized to the nucleus and affected neurite outgrowth. Quantitative proteome analysis of lymphoblastoid cells (LCLs) from the patient with the p.L163X mutation and a normal control revealed several up-regulated proteins in the patient's LCLs that appear to be clinically relevant targets of CUL4B-mediated protein ubiquitination and degradation. Altogether, our study underlines the importance of the ubiquitin-proteasome protein degradation pathway in the pathogenesis of ID and provides insight into the underlying molecular and cellular mechanisms by which CUL4B impacts ID.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2827W

IDENTIFICATION OF NOVEL EXON SPLICE ENHANCERS (ESEs) IN THE GROWTH HORMONE GENE (*GH1*) MUTATED IN ISOLATED GH DEFICIENCY (IGHD). *D. Babu, I. Fusco, S. Mellone, M. Godi, A. Petri, F. Prodam, S. Bellone, P. Momigliano-Richiardi, G. Bona, M. Giordano.* Department of Health Sciences, University of Eastern Piedmont, Novara, Italy.

IGHD type II is may be caused by alteration of an ESE located at the first 7bp of exon 3 (ESE1), resulting in an increased level of the exon 3-skipped mRNA, normally present at low doses (about 5% relative to the wild-type mRNA). An increased level of this isoform, encoding a 17.5 kDa protein was demonstrated to exert a dominant negative effect on the secretion of the 22-kDa isoform, independently from the level of the wild-type. Another ESE (ESE2) was detected by *in vitro* mutagenesis from E3+19 to E3+32. We identified 3 IGHD patients out of 157 carrying variations in the *GH1* exon 3. Patient #1carried an already reported non-synonymous mutation E3+75G>C (Glu82Asp), patient #2 carried the two variations E3+90 C>T (Pro87Pro) and E3+101 A>T (Glu91Val) and patient #3 carried E3+90 C>T (Pro87Pro) and E3+84 G>A(Pro85Pro). These variants were absent in a panel of 400 normal chromosomes and were all originated by gene conversion events from one of the *GH1* cluster homologous genes. None of them was included in the already identified ESEs. However the *in silico* analysis suggested the presence of two other ESEs at nt E3+83-89 (ESE3:CGACACC) and E3+98-104 (ESE4:GGGAGGA). The mutations at positions E3+84 and E3+101 were predicted to completely abolish ESE3 and ESE4, respectively. We thus performed *in vitro* splicing analysis with constructs bearing the wild type *GH1*, and the *GH1* gene carrying the different mutations identified in the patients. The mRNAs from transfected GH4C1 rat pituitary cells showed the same level of the 22 kDa isoform in all the constructs. The GH-E3+75G>C mRNA splicing pattern was similar to that of the wild-type with the 17.5 kDa about 2–5% relative to the wild-type. Conversely, the GH-E3+84G>A and GH-E3+101A>T exhibited a significantly higher level of the 17.5 kDa band (about 30–35% with respect to the wild-type). Interestingly the 17.5 kDa isoform was completely absent in the GH-E3+90C>T mRNA. In conclusion we identified two novel ESEs in the *GH1* exon3 through the detection of mutations in IGHD patients. Two of these mutations E3+84 G>A and E3+101 A>T, cause an increased level of the dominant negative 17.5 kDa isoform. The analysis of the parents of the patients will reveal whether the mutations were inherited from the same parent or one from each parent or if they are *de novo* mutations.

2828T

Functional analysis of R75Q dominant mutation and in cis recessive mutation V37I of GJB2 gene causing hearing loss in a Korean Family. *J. Kim[1], J. Jung[2], Y. Kim[3], J. Choi[2], K. Lee[3].* 1) Lab Med, Yonsei University Wonju College of Medicine, Wonju, South Korea; 2) Otorhinolaryngology, Yonsei University College of Medicine, Seoul, South Korea; 3) Lab Med, Yonsei University College of Medicine, Seoul, South Korea.

Among different genes responsible for hereditary hearing loss, mutations in GJB2 gene, which encodes the gap junction protein connexin 26, are the most common cause of non-syndromic hereditary hearing loss. Interestingly, the phenotypic variation has been observed in different types of mutations. Here, we report a Korean family harboring two mutations R75Q which was reported as a dominant mutation and V37I reported as a recessive mutation in cis. Previously, R75Q has been reported in a family with hearing loss which was rescued by another recessive mutation c.35delG, which was located upstream of R75Q in cis, and the cause of the hearing loss in that family could not be identified. In our study, the propositus reported a bilateral sensorineural form of hearing loss with onset at 4 years and similar form of hearing loss was present in her father. Her grandparents and 7 siblings of her father were not affected. The sequencing analysis of GJB2 gene revealed R75Q and V37I mutations in affected individuals. The grandfather of the propositus had only heterozygote V37I and grandmother and siblings of the propositus' father revealed no mutation in GJB2 gene. Single strand DNA sequencing confirmed that R75Q and V37I resided on the same allele. Thus, R75Q arose de novo in propositus' father and transmitted to his daughter. In the biochemical permeability assay for different genotypes, we examined the cell-to-cell transfer of the impermeant dye Lucifer Yellow. Both the R75Q alone, and R75Q with V37I variant displayed highly defective intercellular dye transfer while V37I revealed a moderate deficit in GJ mediated biochemical coupling. For hemichannel functional study, the entry of PI dye through hemichannels was evaluated. As a result, cells transfected with R75Q and R75Q+V37I displayed significantly lower PI loading, indicating greatly reduced hemichannel activity. The PI loading efficiency was not significant between R75Q and R75Q+V37I. The V37I transfected cells demonstrated hemichannel activity reduced to less than 50% of the wild type. In contrast to other family harboring R75Q and c.35delG in cis, when a dominant mutation is in cis with a recessive mutation that does not cause premature protein truncation is not effected by the recessive mutation. Interpretation of mutations in GJB2 gene should be carefully reviewed and more in depth functional study is needed to help genetic counseling in families with hereditary hearing loss to avoid misguidance.

2829F

Progressive hearing impairment with homozygous deletion in GJB2 gene (Connexin-26) und normal newborn hearing screening. *N. Lublinghoff[1], E. Löhle[2], R. Birkenhäger[1].*1) University Medical Center Freiburg, Department of Otorhinolaryngology Killianstr. 5, D-79106 Freiburg, Germany (Head of Department: R. Laszig, PhD); 2) Department for Phoniatrics and Paedaudiology of the Department of Otorhinolaryngology of the University Medical Center Freiburg, Lehener Str. 88, D-79106 Freiburg, Germany (Head of Department: E. Löhle, PhD).

Hearing impairment is the most common sensorineural disease in humans. About one per 500 to 1000 neonates suffers at birth or in the first years from high-grade to severe hearing impairment. In 50% genetic alterations are the cause of prelingual hearing impairment. Another 50% of prelingual non-syndromic hearing impairment can be attributed to mutations in the GJB2 gene. The most widespread genetic change is the deletion mutation c.35delG, this means a total loss of function. In the studied patient population we have 133 patients with a homozygous deletion mutation in GJB2 gene (c.35delG) and 27 patients who are heterozygous for this mutation on one allele and heterozygous for another loss-of- function mutation in GJB2 gene. We found a normal newborn hearing screening (oto-acoustic emissions, OAE) in 8,6% and a progressive hearing impairment in 20,6%. Analysis of the literature revealed that various teams have similarly concluded that some of the patients with the homozygous gene defect c.del35G have unremarkable findings in neonate hearing screening. Cryns et al. showed in about 11% of the cases they examined that a mild to moderate hearing impairment was initially observed in patients with the homozygous gene defect c.del35G. Snoeckx et al. reported even higher results in an international multicenter study, in which about 28% of 889 patients who carried the above homozygous gene defect presented initially with mild to moderate hearing impairment. This group found similar results in other homozygous function-loss mutations. Pargarkar et al., Norris et al. and Putcha et al. also demonstrated this in a smaller percentage of their patients. Orzan et al. showed that in just over 50% of 79 children who later presented with severe hearing impairment had normal findings in the first 3–6 months of life in the neonate hearing screening. This fact suggest that homozygous deletion mutation in GJB2 gene (c.35delG) does not always contribute to an congenital hearing impairment, but to a progressive hearing loss that might develop over the first month and years of life. The molecular background is not fully understood. Possibly so called modifier genes attribute to different phenotypic characteristic in identical genotypes. Additionally, we have to reevaluate the value of OAE for newborn hearing screening, keeping in mind that one positive result is no warranty for a normal development of hearing function, but a result that should be checked in the course.

2830W

Bestrophin mutation within Ca^{2+} binding domain alters Cl^- conductance. *S. Brar[1], B. Pattnaik[1,2,3], P. Halbach[1], S. Tokarz[1], D. Pillers[1,3].* 1) Pediatrics, University of Wisconsin, Madison, WI; 2) Opthalmology, University of Wisconsin, Madison, WI; 3) Eye Research Institute, University of Wisconsin, Madison, WI.

Best Vitelliform Macular Dystrophy (Best's disease or BVMD) is an autosomal dominant juvenile-onset macular degeneration associated with more than 100 mutations in the 585 amino acid protein, bestrophin. Clinical and histological features of BVMD include vitelliform lesions in the macula. A distinct genotype-phenotype correlation based on these mutations has yet to be delineated, and may be due in part to the impact on the three dimensional structure and function of the protein in the context of the mutation. Bestrophin is a Cl- ion-channel present in the retinal pigment epithelium (RPE), and is activated by intracellular Ca2+. The RPE is a monolayer of post-mitotic cells located between the neural retina and the choroidal vasculature. Ion channels in the RPE support many functions by controlling the ion and nutrient environment of the photoreceptors, and by forming the blood-retinal barrier. We hypothesized that mutations that affect Cl- conductance are a basis for some of the retinal phenotypes associated with BVMD. We sought to physiologic phenotype disease associated mutation, Asn296His, located within the hotspot Ca2+-binding domain. Using site-directed mutagenesis, we introduced the Asn296His within human bestrophin protein (hBest1). The physiology of the wildtype and mutant protein was tested by transiently expression in Chinese Hamster Ovary (CHO) cells. Whole cell current was measured after inducing intracellular Ca2+ changes by activation of G-protein Coupled Receptor. We found that bestrophin functions as a Ca2+ activated Cl- channel. The Asn296His mutation generates a non- functional channel which does not respond to physiologic changes in intracellular Ca2+ in CHO cells. Permeability of the mutant channel to other anions is also altered. The mutation Asn296His is situated within the intracellular loop of bestrophin, which is comprised of a cluster of amino acid residues lining the Ca2+ binding pocket. We found that this mutation alters Ca2+ sensitivity of the channel which confirmed that Ca2+ binding gates this channel open. The results established a direct relationship between the Ca2+ binding site mutation and altered Cl- conductance in support of the hypothesis that hBest1 is a Ca2+-activated Cl- channel. By altering the physiological ionic homeostasis across the RPE, the photoreceptor outersegment is not metabolized resulting in the accumulation of debris between RPE and retina and thus the characteristic clinical phenotype.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2831T

Functional characterization of *SLC2A9* allelic variants in Czech population. *B. Stiburkova[1], O. Hurba[1], V. Krylov[1,2], A. Mencikova[2].* 1) Institute of Inherited Metabolic Disorders, First Faculty of Medicine, Charles University in Prague and General University Hospital in Prague, Czech Republic; 2) Department of Cell Biology, Faculty of Science, Charles University in Prague, Czech Republic.

Urate (uric acid) is the major end product of purine degradation in humans and higher primates. Serum urate concentration is determined by the balance between production and elimination. Correlation between increased serum uric acid level (hyperuricemia) and hypertension, cardiovascular disease, insuline resistance, metabolic syndrome and renal disorders has been described. Transport mechanisms for urate are localized in renal proximal tubule, where urate is secreted and extensively reabsorbed. The most causative gene influencing the level of serum uric acid is *SLC2A9* gene (transporter GLUT9). The aim of the project is to clarify frequency, role and the effect of allelic variants of urate transporter *SLC2A9* to the serum level of uric acid in statistically significant cohort of Czech population. Methods used: retrospective cohort of 869 subjects, which were already biochemically and clinically characterized; selection of 150 subjects with normouricemia and 59 subjects with pathological levels of serum uric acid. After identification of allelic variants (PCR amplification and sequencing analysis) further detailed studies using expression system of *Xenopus laevis* oocytes were performed. Analysis of *SLC2A9* revealed 14 sequence variations in the intronic regions. 12 allelic variants were found in exonic regions: six substitutions (p.S41S, p.L108L, p.T125T, p.G136G, p.I168I, p.L189L) and six transitions (p.G25R, p.T275M, p.D281H, p.V282I, p.R294H, p.P350L). cDNA mutant variants were prepared from both isoforms of wild type GLUT9 using mutagenesis kit. In vitro transcribed cRNAs were then microinjected into growing *Xenopus laevis* oocytes. The function analysis for *SLC2A9* including subcellular localization, colocalization and processing dynamics and transport of proteins are in process. Results from the study could contribute to the elucidation of genetic basis of serum uric acid concentration; clarify the interference among major urate transporter, hyperuricemia and gout. This study was supported by grant IGA MZ NT/11322-4/2010 from the Ministry of Health of the Czech Republic.

2832F

Identification of a missense mutation outside the forkhead domain of FOXL2 causing a severe form of BPES type II. *H. Haghighi-Kakhki[1], A. Haghighi[2], H. Verdin[3], N. Piri[4], N. Saleh-Gohari[5], E. De Baere[3].* 1) Faculty of Medicine, Mashhad Azad University, Mashhad, Iran; 2) Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK; 3) Center for Medical Genetics, Ghent University Hospital, Ghent, Belgium; 4) Department of Ophthalmology, Alavi Eye Hospital, Ardebil University of Medical Sciences, Ardebil, Iran; 5) Department of Genetics, Kerman University of Medical Sciences, Kerman, Iran.

Blepharophimosis-ptosis-epicanthus inversus syndrome (BPES) is a developmental disease characterized by a complex eyelid malformation associated or not with premature ovarian failure (POF). BPES is essentially an autosomal dominant disease, due to mutations in the forkhead box L2 (FOXL2) gene, encoding a forkhead transcription factor. More than one hundred unique FOXL2 mutations have been described in BPES in different populations, many of which are missense mutations in the forkhead domain. Here, we report on a very severe form of BPES resulting from a missense mutation outside the forkhead domain. A clinical and molecular genetic investigation was performed in affected and unaffected members of an Iranian family with BPES. The FOXL2 coding region was sequenced in an index case. Targeted mutation testing was performed in 8 family members. We have identified a heterozygous FOXL2 missense mutation c.650C>G (p.Ser217Cys) co-segregating with disease in members of a three-generation family with BPES type II. Only few missense mutations have been reported outside the forkhead domain so far. They were all found in mild BPES, in line with in vitro studies demonstrating mostly normal localization and normal or increased transactivation properties of the mutant proteins. Unlike previous studies reporting missense mutations outside the forkhead domain, affected members of the family studied here showed a severe BPES phenotype, with bilateral amblyopia due to uncorrected ptosis. As this was observed in all affected individuals of different generations, this might be attributed for instance to a cis-effect of regulatory variants within or outside the transcription unit modulating the expression of mutant FOXL2. This hypothesis was tested by sequencing of the UTRs and by copy number screening of the upstream regulatory domain of FOXL2, covering previously described regulatory deletions, but no variations were found. Several alternative possibilities, such as subtle sequence changes in regulatory elements, or promoter variations cannot be ruled out however. In conclusion, this is the first study demonstrating a severe BPES phenotype resulting from a FOXL2 missense mutation outside the forkhead domain, expanding our knowledge about the phenotypic consequences of missense mutations outside the forkhead domain in BPES. The phenotypic effect of the identified mutation seems to be tissue-specific, as no ovarian involvement was observed in any of the affected females here.

2833W

The UBIAD1 prenyltransferase links menaquione-4 synthesis to cholesterol metabolic enzymes. *M. Nickerson[1], A. Bosley[2], J. Weiss[3], B. Kostiha[1], W. Brandt[4], T. Andresson[2], T. Okano[5], M. Dean[1].* 1) Cancer & Inflammation Program, National Cancer Institute, Frederick, MD; 2) Laboratory of Proteomics and Analytical Technologies, SAIC-Frederick, National Cancer Institute, Frederick, MD; 3) Department of Ophthalmology, Louisiana State University, New Orleans, LA; 4) Department of Bioorganic Chemistry, Leibniz Institute of Plant Biochemistry, Halle (Saale), Germany; 5) Department of Hygienic Sciences, Kobe Pharmaceutical University, Higashinada-ku, Kobe, Japan.

Schnyder corneal dystrophy (SCD) is an autosomal dominant disease characterized by missense alterations of UBIAD1 and deposition of cholesterol in the cornea leading to progressive opacification and loss of visual acuity. UBIAD1 was recently shown to synthesize menaquinone-4 (MK-4), a form of vitamin K, but causal mechanisms of disease are unknown. Here we report novel genetic and biochemical data that clarifies the role of UBIAD1 in SCD. We report a novel G>A mutation altering UBIAD1 glycine 177 to glutamic acid (p.G177E) in six SCD families, including four families from Finland who share a likely founder mutation. We observed significantly reduced MK-4 synthesis by p.G177-mutant UBIAD1 due to disruption of transmembrane helices and active site residues. We show UBIAD1 binds to and co-localizes with HMGCR and SOAT1, enzymes catalyzing cholesterol synthesis and storage, respectively. Docking simulations indicated cholesterol bound to UBIAD1 in the substrate binding cleft and overlapped with GGPP binding, a MK-4 substrate, suggesting competition between metabolites. Impaired MK-4 synthesis is a first biochemical defect identified in SCD and suggests that topical administration of MK-4 should be investigated as a treatment for SCD patient corneas. Our data identifies a role for endogenous MK-4 in maintaining cornea health and visual acuity. These results are discussed in light of recent data showing vitamin k is an electron carrier in mitochondria able to reverse symptoms of Parkinson's disease.

2834T

Intergenic or 5'UTR deletion of Transcription factor 4 (TCF4) gene in Pitt-Hopkins syndrome. *J. Lee[1,3], S-U. Moon[1], E. Shin[2], S. Hassed[3], J. Chanloner[3], X. Wang[3], S. Li[3].* 1) Pathology, Korea University, Seoul, South Korea; 2) Genome Research Center, Neodin Medical Institute 2-3, Seoul, 133–847 Korea; 3) Department of Pediatrics, OUHSC, OK, 73014, USA.

Haploinsufficiency of transcription factor 4 (TCF4) gene, encoding a basic helix-loop-helix (bHLH) transcription factor, has been identified as the molecular mechanism underlying Pitt-Hopkins syndrome (PHS), which is characterized by mental retardation, distinctive facial features including a wide mouth, and intermittent hyperventilation followed by apnea. Recent analysis of 112 individuals with PHS delineated the TCF mutational spectrum, with 40% point mutations, 30% small deletions/insertions, and 30% microdeletions. Most of these were private mutations and generated premature stop codons. Here, we present two cases with intergenic deletion and one case with 5' UTR deletion of TCF4 gene in patients with PHS phenotypes. Reduced level of RNA expression of TCF gene was confirmed by subsequent real time quantitative PCR. In conclusion, haploinsufficiency of TCF gene due to the intergenic or 5' UTR deletions may not be rare and high density array CGH will facilitate the diagnosis of PHS.

2835F

No evidence of genetic difference for clinical severity between monozygotic twins with Rett syndrome. *T. Kubota[1], K. Miyake[1], T. Hirasawa[1], Y. Minakuchi[2], A. Toyoada[2].* 1) Department of Epigenetic Medicine, Interdisciplinary Graduate School of Medicine and Engineering, University of Yamanashi, Chuo, Yamanashi, Japan; 2) Comparative Genomic Laboratory, Center for Genetic Resource Information, National Institute of Genetics, Research Organization of Information and Systems, Mishima, Shizuoka, Japan.

Thanks to the advancement of sequencing technology, mutations have been identified in neuronal molecules as causes of neurodevelopmental disease such as autism. In order to identify a new related gene, we performed re-sequencing in discordant monozygotic twins with Rett syndrome. In this twins, twin 1 exhibited severe neurological features starting from infancy, whereas twin 2 showed no symptom until 3 years of age, although they share the same frame-shift mutation in MECP2 gene and have similar X-chromosome inactivation patterns in lymphocytes. We initially identified 28 twin 1-specific SNPs (130 billion bp, read coverage 43.3x) and 16 twin 1-specific SNPs (116 billion bp, read coverage 38.5x) by next-generation sequencing (Hiseq 2000, Solexa) in lymphocytes and differential 19 SNPs between the twins by array-based method (SNP/CNV BeadChip) in fibroblasts, but nor were any reproducible differences. Therefore, we did not find evidence of DNA-nucleotide differences that explained clinical severity. We are now currently investigating genetic differences in copy number variations and epigenetic differences in DNA methylation.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2836W

Two-nucleotide deletion of the ATP7A gene, leading not to premature termination, but to translation reinitiation, produces mild phenotype in a male Menkes disease patient. *T. Wada[1], T. Murakami[1], H. Shimbo[1], H. Osaka[1], H. Kodama[2].* 1) Pediatric Neurology, Kanagawa Children's Med Center, Yokohama, Kanagawa, Japan; 2) Dept Health Dietetics, Teikyo Heisei Univ, Tokyo, Japan.

It is important to determine the relationship between the clinical phenotype and genotype of genetic diseases in clinical practice. In general, frameshift mutations that is, any downstream change that adds or removes a non-integral number of codons, or a number of nucleotides that is not a multiple of three would be expected to produce a premature termination codon, or abnormal splicing. Nonsense-mediated decay works for these mutations, detecting mRNAs containing premature termination codons and degrading them, resulting in no expression of the mutated gene. So, in X-linked diseases, male patients with frameshift mutations are expected to have a severe or leven phenotype. Here, we report a case of an 11-month-old male patient with Menkes disease, an X-linked recessive copper metabolism disorder, caused by mutations in ATP7A (NIM300011). His genomic DNA shows two-nucleotide deletion at exon 4 out of 23 exons (c.1137_1138delTG) of the ATP7A gene, but his clinical phenotype involves unexpectedly mild symptoms. Characterization of his ATP7A mRNA shows no abnormal splicing. Analysis of ATP7A protein in cell lysate obtained from his cultured fibroblast shows a smaller form in size. Our findings suggest that translation reinitiation takes place at downstream internal codons, leading to the synthesis of N-terminally-truncated and partially functional Menkes proteins, which are probably missing of some copper-binding domains. Our results suggest that an apparent frameshift mutation cannot always lead to the production of a premature termination codon, and these patients can exhibit unexpectedly mild phenotypes.

2837T

Retinoic Acid Induced 1, RAI1: Molecular and functional studies to understand its relationship to Neurobehavioral Alterations. *K. Walz[1, 2, 3], IC. Perez[1], P. Carmona-Mora[4], C. Abad[1], CP. Canales[4], L. Cao[2], AK. Srivastava[5, 6], JI. Young[1, 2].* 1) John P. Hussman Institute for Human Genomics, Univ Miami, Miami, FL; 2) Department of Human Genetics, Miller School of Medicine, University of Miami, FL, USA; 3) Department of Medicine, Miller School of Medicine, University of Miami, FL, USA; 4) Cellular and Genetic Medicine Unit, Department of Anatomy, School of Medical Sciences, The University of New South Wales, Sydney, Australia; 5) JC Self Research Institute of Human Genetics, Greenwood Genetic Center, Greenwood, SC, USA; 6) Department of Genetics and Biochemistry, Clemson University, Clemson, SC, USA.

Smith Magenis Syndrome (SMS) is a complex disorder associated with del17p11.2. SMS clinical presentation includes behavioral, sleep, skeletal, and neurological abnormalities among others. Most SMS features are caused by haploinsufficiency of the RAI gene, whose function remains poorly understood. With the goal of determining the underlying mechanisms linking RAI1 mutations with neurobehavioral phenotypes present in SMS we compared the phenotype of Rai1+/− mice (“Rai1 specific” model of SMS) with Df(11)17/+ mice (model of SMS with deletion of 17p11.2 syntenic region). Three main phenotypic domains were tested: social interactions, communication, and repetitive behaviors. We found that Rai1+/− mice showed a similar decrease in dominant behavior and increase in preference for social novelty as the Df(11)17/+ mice, confirming Rai1 main role in social behaviors. Given this link between Rai1 haploinsufficiency and impairment in social outcomes, we focused the analysis of RAI1 mutations found in SMS patients on transcriptional regulation of its endogenous target BDNF, a mediator of neural plasticity linked to social exposure. We defined two groups of mutations: 1) early truncating mutations and 2) missense mutations mapping to the C-terminal region. The first group exhibited exclusive cytoplasmic localization (in contrast to the nuclear localization of full length RAI1) and inability to activate transcription through the BDNF enhancer. We analyzed lymphoblastoid cells from a SMS patient carrying RAI1 c.3103insC, belonging to this first group, by cell fractionation and western blot. We found a protein of the size of truncated Rai1 in the cytoplasm fraction and another of full length size in the nucleus, confirming the in vitro data of abnormal localization of early truncation mutants of Rai1. The second group did not activate BDNF enhancer mediated transcription despite having a correct nuclear localization. In summary, our data further implicate Rai1 as a main player in the neuro-behavior aspects of SMS in mice. In addition, our in vitro results indicate that the different groups of mutations have different molecular outcomes. However; a common theme appears to be a reduction of total RAI1 transcription factor activity. This may explain the phenotypic similarity between SMS patients (and mouse models) harboring big deletions vs RAI1 mutations, and certainly gives a clue to explore therapeutic interventions.

2838F

RAB11FIP1 interacts with the BLOC-1 complex to retrieve melanogenic proteins from the recycling pathway and a dominant negative mutation in *RAB11FIP1* causes Hermanksy-Pudlak Syndrome Type 10 (HPS-10). *A.R. Cullinane[1], J. Pan[1], M.A. Merideth[1], J.A. Curry[1], J.G. White[2], M. Huizing[1], W.A. Gahl[1].* 1) Medical Genetics Branch, NHGRI (NIH), Bethesda, MD; 2) Department of Laboratory Medicine, University of Minnesota, Minneapolis, MN.

Hermansky-Pudlak Syndrome (HPS) is a genetically heterogeneous disorder of lysosome-related organelle (LRO) biogenesis and is characterized by oculocutaneous albinism and a bleeding diathesis. There are currently 9 known genes that cause HPS, all of whose protein products function in the biogenesis of LROs. The Biogenesis of Lysosome related Organelle Complex 1 (BLOC-1) contains 8 subunits but relatively little is known about the intracellular function of the complex, although a role in endosomal protein sorting has been suggested. Using his-tagged BLOC-1 subunits expressed in HEK293 cells and mass spectroscopy, we discovered that RAB11FIP1 is a novel interacting protein of the BLOC-1 complex. *RAB11FIP1* encodes a RAB11A interacting protein that homo-dimerizes to interact with RAB11A. A yeast-2-hybrid assay showed that the Dysbindin subunit of BLOC-1 directly interacts with RAB11FIP1; this was confirmed by co-immunoprecipitation and confocal immunofluorescence microscopy in melanocytes. Here we report a girl who had previously been screened for mutations in HPS-1 through HPS-6 and all the genes encoding the BLOC-1 complex. No mutations were found, although the patient had typical signs and symptoms of HPS and a cellular phenotype mimicking that of BLOC-1, i.e., increased plasma membrane cycling and endosomal accumulation of a melanogenic protein, TYRP1. Whole exome sequencing revealed a de novo heterozygous frameshift mutation in *RAB11FIP1*. The short protein fragment from this allele was expressed and interacted with the full-length protein, resulting in a dominant negative effect. Known cargos of the BLOC-1 complex in melanocytes are TYRP1 and ATP7A. How these cargos traffic to LROs was unknown, but we discovered that GFP-TYRP1 traffics to the plasma membrane, is endocytosed and only then directed to LROs. We demonstrated that TYRP1 interacts with the AP-1, AP-2 and AP-3 complexes, allowing this trafficking to occur. ATP7A, however, appears to traffic directly to endocytic vesicles, where RAB11FIP1 and the BLOC-1 complex are required for retrieval to LROs. Taken together, these data suggest a function of the BLOC-1 complex in retarding protein recycling by forming a physical brake between early endosomes (through the BLOC-1 interactor, Syntaxin-13) and recycling endosomes (through the BLOC-1 interactor, RAB11FIP1). This would allow more time for proteins to be retrieved from the endosomal compartment (by the AP-3 complex) and directed to LROs.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2839W

Two novel RRM2B gene mutations in a patient with autosomal recessive progressive external ophthalmoplegia, encephalopathy and cytochrome c oxidase deficiency.N. *S. Hauser[1], R. Bai[2].* 1) Medical Genetics, Childrens Hospital Central California, Madera, CA; 2) GeneDx Inc, Gaithersburg, MD 20877.

Abstract: We describe a male child born to healthy parents at 32 weeks gestation. Baby was 5lbs 9oz. He remained in the NICU for several weeks. He had difficulty with GERD in the first months of life, but no other medical problems. The patient was the third child in a sibship of 3, the 2 siblings were healthy. Mother described him as a placid baby, never demanding, with myopathic facial features. Parents described him as “weaker” as compared to his other siblings. At 7 mos of age he was still working on head control, and had not yet achieved the ability to sit up. Five days prior to admission at 7.5mos of age, he lost head control and the ability to eat. Parents brought him to the local ER, where he was admitted. During his admission, he became progressively weaker. He eventually required a tracheostomy and g-tube. Currently, he has minimal spontaneous movement, ophthalmoplegia, ptosis and seizures. The diagnosis of cytochrome C oxidase deficiency was made from muscle biopsy. The muscle biopsy showed abnormal mitochondria with structural abnormalities, reduced cytochrome C oxidase by enzyme histochemistry and increased neutral lipid. Mitochondrial respiratory chain testing showed a reproducible deficiency in cytochrome C oxidase activity to 13% of the normal reference mean. Complexes II-III were also reduced at 16% of the normal mean. Citrate synthase activity was elevated at 234% of control, a marker for mitochondrial content. Molecular Genetic Testing: Sequencing results of the following individual genes-SCO1, SCO2, SURF1, COX10, COX15, COX6B1, and FASTKD2 were normal. Sequence analysis of a panel of 24 nuclear genes (BCS1L, COQ2, COX10, COX15, COX6B1, DGUOK, DLD, MPV17, OPA1, PDHA1, PEO1, POLG, POLG2, RRM2B, SCO1, SCO2, SLC25A4, SUCLA2, SUCLG1, SURF1, TACO1, TIMM8A, TK2, and TYMP) identified two heterozygous novel missense mutations in the RRM2B gene- G200E (c.599 G>A, GGA>GAA) and E105K (c.313 G>A, GAA>AAA), both are non-conservative amino acid changes, at highly conserved positions, not seen in the general population or any available online mutation/variant database, and are predicted to be disease-causing. No multiple mtDNA deletions were detected in the patient's muscle specimen by long range PCR. The mtDNA content of the patient's muscle biopsy was approximately 2% of age matched controls. Conclusion: We conclude the cytochrome C oxidase deficiency noted on this patients muscle biopsy is caused by these two novel mutations in the RRM2B gene.

2840T

Isogenic D4Z4 contracted and non-contracted immortal muscle cell clones as a model for FSHD. *S. van der Maarel[1], Y. Krom[1], J. Dumonceaux[2], K. Mamchoui[2], B. den Hamer[1], V. Mariot[2], E. Negroni[2], L. Geng[3], R. Tawil[4], S. Tapscott[3,5], B. van Engelen[6], V. Mouly[2], G. Butler-Browne[2].* 1) Human Genetics; S4-P, Leiden University Medical Center, Leiden, Netherlands; 2) Thérapie des maladies du muscle strié, Institut de Myologie, UM76, UPMC Université Paris 6, Paris, France; 3) Division of Human Biology, Fred Hutchinson Cancer Research Center, Seattle, Washington, United States of America; 4) Department of Neurology, University of Rochester, Rochester, New York, United States of America; 5) Department of Neurology, University of Washington, Seattle, Washington, United States of America; 6) Radboud University Nijmegen Medical Centre, Department of Neurology, Nijmegen, the Netherlands.

Facioscapulohumeral muscular dystrophy (FSHD) is associated with a contraction of the D4Z4 repeat on chromosome 4. Contracted repeats show a reduction in heterochromatic markers and derepression of the DUX4 retrogene. The complex (epi)genetic etiology of FSHD and the high clinical variability complicates investigations on the pathogenic mechanism underlying FSHD. Taking advantage of the high prevalence of somatic D4Z4 repeat array contractions in FSHD we generated a clonal myogenic cell model from a mosaic patient where each clone is genetically identical except for the D4Z4 repeat array size, being either non-contracted or contracted. D4Z4 contracted and non-contracted clones retain their myogenic properties but D4Z4 contracted clones differ from non-contracted clones by a variegated pattern of DUX4 expression in sporadic nuclei. Consistently, the consequences of DUX4 expression can only be observed in D4Z4 contracted clones, including the differential expression of previously described DUX4 targets. In immunodeficient mice, D4Z4 contracted and non-contracted clones participate equally in in vivo muscle regeneration, further expanding the potential of these clones. These cell lines can be used for pair wise comparisons to identify FSHD-specific differences and are expected to become valuable tools for for high throughput drug screens.

2841F

MEFV Second Exon Variation Analysis in Familial Mediterranean Fever. *A.K. Kirectepe[1], G. Celikyapi Erdem[1], O. Kasapcopur[2], E. Tahir Turanli[1,3].* 1) Dr. Orhan Öcalgiray Molecular Biology-Biotechnology and Genetics Research Center, Istanbul Technical University, Istanbul, Turkey; 2) Department of Pediatric Rheumatology, Cerrahpasa School of Medicine, Istanbul University, Istanbul, Turkey; 3) Molecular Biology and Genetics Department, Science and Letter Faculty, Istanbul Technical University, Istanbul, Turkey.

Objective: Familial Mediterranean Fever is an autoinflammatory disorder, which is related with MEFV gen mutations and defects of expression regulation. Previous studies showed that there are different alternatively spliced forms of MEFV gene. Our group previously showed that second exon deleted MEFV transcript (MEFV-d2) is increased in FMF patients compared to healthy controls (p= 0.026). In this study, we aimed to analyze whether the transcript difference is related to second exon variations. Methods: DNA was isolated from venous blood of FMF patients and healthy controls, which are in attack-free period (N=50). Second exon of MEFV was sequenced. Analysis of the sequence results was performed by Genious 4.8.5 and was compared to MEFV variations from literature and infevers database. Results: D102D (FMF patients=44%, healthy controls=51%), G138G (FMF patients= 66%, healthy controls=71%), A165A (FMF patients=72%, healthy controls= 71%), and R202Q (FMF patients=61%, healthy controls=50%), which are located on second exon of MEFV, are the most commonly observed variations in FMF patients and healthy controls. Student T test analysis on FMF patients with healthy controls found no significant difference in MEFV-d2 mRNA levels between samples with or without variations. Conclusion: MEFV-d2 transcription difference was not related to MEFV second exon variations in our study sample. We aim to increase the sample size to assess the significance of our results.

2842W

Role of NLRP7 in Recurrent Hydatidiform Moles and Reproductive Wastage. *R. Slim[1], C. Messaed[1], E. Akoury[1], W. Chebaro[1], W. Buckett[1], J. Arseneau[1], P. Sauthier[2].* 1) McGill University, Montreal, Canada; 2) Hôpital Notre-Dame, Gynécologie oncologique.

Hydatidiform mole is an aberrant human pregnancy with no embryo that has puzzled scientists in all civilizations. To date, two genes, NLRP7 and C6orf221, with causal roles in this condition have been identified. NLRP7 is a major gene for this condition and is mutated in 50%–80% of analyzed patients with recurrent moles, depending on family history and ethnicity. C6orf221 is a recently identified gene and available data indicate that C6orf221 plays a minor role in the causality of recurrent moles because only 5% to 15% of analyzed patients with recurrent moles, who are negative for NLRP7 mutations, have mutations in C6orf221. NLRP7 is a member of the nucleotide oligomerization domain like-receptor protein with roles in inflammation and apoptosis. We showed that cells from patients with NLRP7 mutations have normal intracellular pro-ILB synthesis and processing but defective mature IL1B secretion. Within the cells, NLRP7 localizes to the Golgi, microtubule organizing center, and associates with microtubules. The association of NLRP7 with microtubules is in line with the postzygotic cleavage abnormalities leading to complex mosaicism and aneuploidies observed in some patients with NLRP7 mutations. We believe that the defective IL1B secretion of cells from patients with NLRP7 mutations make the patient more tolerant to the growth of aberrant molar conceptions and the retention of these conceptions to later gestational stages contribute to the molar phenotype. Our data implicate NLRP7 in several forms of reproductive wastage, unravel new mechanisms for reproductive wastage, and open new avenues of research to better understand gestational choriocarcinomas, which may develop after any type of pregnancy, but most commonly after hydatidiform moles.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2843T

A novel *GUCY2D* mutation, V933A, causes central areolar choroidal dystrophy. *A.E. Hughes[1], W. Meng[1], A.J. Lotery[2], D.T. Bradley[1].* 1) Centre for Public Health, Queen's Univ Belfast, Belfast, United Kingdom; 2) Clinical Neurosciences Research Group, University of Southampton, UK.

The purpose of our study was to identify the genetic cause of central areolar choroidal dystrophy (CACD) in a large Northern Irish family. We previously reported linkage of the locus for CACD in this family to an interval of approximately 5 cM on chromosome 17p13 flanked by polymorphic markers D17S1810 and CHLC GATA7B03. We undertook sequence capture, massively-parallel sequencing and computational alignment, base-calling and annotation to identify a causative mutation. Conventional sequencing was used to confirm the results. Deep sequencing identified a single-base substitution in guanylate cyclase 2D, membrane (retina-specific) (*GUCY2D*). The novel mutation segregated with the disease phenotype and resulted in substitution of valine by alanine at position 933, within the catalytic domain of the protein. It altered a motif that is strongly conserved in a large number of distantly related proteins across several species, and was predicted to have a damaging effect on protein activity. Mutations in *GUCY2D* have previously been associated with dominant cone rod dystrophies (CORD6) and recessive forms of Leber's congenital amaurosis (LCA). This is the first report of *GUCY2D* mutation causing CACD and adds to our understanding of genotype-phenotype correlation in this heterogeneous group of choroidoretinal dystrophies.

2844F

Novel splice site mutations in NPHS1 gene in Indian congenital nephrotic syndrome patients. *S. sharma, A. Bagga, M. Kabra.* All India Institute of Medical Sciences, New Delhi, India.

Congenital nephrotic syndrome (CNS) is defined as nephrotic syndrome that manifests within the first 3 months of life. Mutations in the NPHS1 gene encoding Nephrin, are a major cause for CNS. It is found at a relatively high frequency in Finns. Following approval by the Institute Ethics Committee and parental consent, we investigated the prevalence of disease-causing mutations in NPHS1 gene in 10 Indian patients with CNS. Mutation analysis was done for all the exons and exon/intron boundaries of NPHS1 gene using PCR and direct sequencing. Two novel homozygous splice site mutations were found in 2 different patients and were absent in 25 ethnically matched controls. The first mutation (c.712+2T>G) was present at donor splice site and the second mutation (c.713-2A>G) was present at acceptor splice site. The evaluation of splice sites and potential splice sites was performed using the Genie program http://www.fruitfly.org/seq_tools/splice.html) human splicing finder (http://www.umd.be/HSF/. Both the mutations were predicted to disrupt normal splicing and insertion of intronic sequence in mRNA. Our findings indicate that, even though CNS is not very common in India, it is important to include it as a possibility in genetic counselling.

2845W

Genetic Basis of Prune Belly Syndrome: Screening the Candidate Genes. *L. Baker, S. Harrison, C. Granberg, C. Villanueva.* Pediatric Urology, Univ Texas Southwestern, Dallas, TX.

Prune-belly syndrome (PBS) is a multi-system congenital anomaly complex of unknown cause. The male predominance of PBS and families with multiple individuals affected by PBS suggest a genetic basis. Recently, two genes, hepatocyte nuclear factor-1-beta (HNF1B) and muscarinic acetylcholine receptor M3 (CHRM3), have been found to cause PBS-like syndromes that included clinical features atypical for PBS. A PBS-like disorder was reported in 2 patients with large chromosomal deletions involving HNF1B, a transcription factor important in embryonic mesodermal and endodermal development. The CHRM3 gene was implicated after one inbred kindred with 6 affected brothers were found to carry a deleterious CHRM3 homozygous mutation. Previously, we found no significant point mutations in HNF1B in 32 PBS patients. Our objective was to test the role of these genes in PBS by screening for sequence mutations in CHRM3 and for copy-number variations (CNVs) in HNF1B. With IRB approval, PBS patients have been prospectively enrolled since 2001 in our Pediatric Genitourinary DNA Repository. DNA was extracted from blood lymphocytes for genetic testing. To assess for CHRM3 mutations, the 1773bp single coding exon of CHRM3 was PCR amplified and sequenced. A multiplex ligation-dependent probe amplification (MLPA) P241-B1 MODY kit (MRC Holland) was used to screen for CNVs in HNF1B and data analysis was performed with GeneMarker (Softgenetics) while a DNA sample with a known HNF1B microdeletion served as a control. DNA samples from 42 probands (40M:2F) with typical features of PBS were tested. Sequencing of the CHRM3 gene revealed no novel non-synonymous mutations in this cohort. MLPA testing for intragenic CNV in HNF1B in this PBS cohort revealed no microdeletions or microduplications. Although HNF1B and CHRM3 are candidate PBS genes, no genetic alterations in these genes were found in our large series of patients with typical PBS features, suggesting that mutations in CHRM3 and HNF1B are not common causes of typical PBS. Thus, the search for the genetic basis of PBS continues.

2846T

DNA Copy-Number Variations in Persistent Cloaca Patients. *S. Harrison, L. Baker.* Urology, Univ Texas Southwestern, Dallas, TX.

Persistent cloaca is a severe human malformation affecting females in which the urinary, genital, and lower intestinal tracts terminally converge to share a single common outlet channel, called a cloaca. Persistent cloaca occurs in 1:50,000 live births and requires multiple corrective surgeries. Cloaca is a devastating anomaly as upper renal tract malformations are commonly associated and as many as 50% of cloaca patients develop renal failure requiring dialysis or transplantation. Currently, only a few genes have been suggested to cause persistent cloaca, however mutations in the genes have only been found in isolated cases or in patients in which the cloaca phenotype is part of a multisystem syndrome. Genomic copy-number variations (CNV) have been shown to account for previously unexplained genetic diseases by identifying candidate genes or regions involved in various disorders. The purpose of this study was to analyze the role of CNVs in persistent cloaca patients. Blood samples from 14 cases of persistent cloaca were prospectively tested by whole genome comparative genomic hybridization (CGH) using the array version current (V8.1 or V8.3) at the time of ascertainment (Baylor Medical Genetics Laboratory, Houston TX). CNV screening by CGH identified 5 patients with abnormal results (~36%). Three cases had 0.240–1.113Mb duplications on 3p26.3, 16p11.206, and 21q22.3, all of which are likely benign. More importantly, two patients had possible cloaca-causing CNVs. The first case carries a 0.517Mb duplication on 16p13.2 encompassing a single gene. This CNV is paternally inherited, suggesting a possible cause of persistent cloaca as this phenotype only affects females. The second case of significance carries a de novo 7.341Mb deletion on 1q32.1q32.3 encompassing 47 genes. The fact that these CNVs are either paternally inherited or de novo suggests both regions are candidate regions for cloaca. Further investigations within these CNVs may lead to identification of causal genes for human persistent cloaca.

2847F

Whole genome sequencing reveals a deletion of the last exon of *FAN1* in Basenji dogs with adult-onset Fanconi syndrome. *G. S. Johnson[1], F. H. Farias[1], T. Mhlanga-Mutangadura[1], J. F. Taylor[2], D. P. O'Brien[3], R. D. Schnabel[2].* 1) Department of Veterinary Pathobiology, University of Missouri, Columbia, MO; 2) Division of Animal Sciences, University of Missouri, Columbia, MO; 3) Department of Veterinary Medicine and Surgery, University of Missouri, Columbia, MO.

Fanconi Syndrome (FS) is a generalized proximal tubule reabsorption deficiency, characterized by metabolic acidosis, amino aciduria, glucosuria, and phosphaturia. An autosomal recessive, adult-onset FS occurs in Basenji dogs. Linkage analysis and homozygosity mapping restricted the FS locus to 2.7 Mb of CFA3 containing 16 genes. We used NextGen sequencing technology (Illumina GAIIx and Hi-Seq 2000) to generate a 12.4-fold coverage whole genome sequence of a Basenji affected with FS. The reads were aligned to the canine genome reference sequence (build 2.1) with NextGENe software (SoftGenetics). None of the 244 annotated exons within the targeted region contained credible causal sequence variants; however, 16 exons fell partially or completely within gaps in the aligned basenji sequence, mostly associated with high G-C content. We therefore compared the depth-of-coverage profiles for the affected basenji sequence to profiles from similarly generated whole-genome sequences from 3 other dogs with different heritable diseases. A unique 334 bp gap was detected in the Basenji sequence. Sanger sequencing across the gap revealed a 321 bp deletion which included the last exon of *FAN1*. Homozygosity for the deletion allele was strongly associated with FS in a cohort of basenjis. The *FAN1* gene symbol stands for Fanconi-anemia Associated Nuclease 1, so named because its encoded protein, FAN1, interacts with components of the Fanconi anemia (FA) pathway for DNA-crosslink repair. FA and FS are two seemingly distinct diseases referred to with eponyms honoring Swiss pediatrician Guido Fanconi. The last exon of *FAN1*encodes part of a VRR-NUC nuclease domain, believed to be of functional importance. We here report that canine FAN1 deficiency causes FS, not FA as predicted by others based on previous cell culture experiments. If FAN1 deficiency is found to cause FS in additional species, the name of the gene can be changed to Fanconi-syndrome Associated Nuclease 1 without a gene symbol change.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2848W

Formation of a novel N-glycosylation motif in integrin α3 due to a rare *ITGA3* gene polymorphism causes congenital nephrotic syndrome and interstitial lung disease. *K.Y. Renkema[1], N. Nicolaou[1], C. Margadant[2], S.H. Kevelam[1], M.R. Lilien[3], M.J.S. Oosterveld[3], M. Kreft[2], A.M. van Eerde[1], R. Pfundt[4], P.A. Terhal[1], B. van der Zwaag[1], N. Sachs[2], R. Goldschmeding[5], N.V.A.M. Knoers[1], A. Sonnenberg[2].* 1) Medical Genetics, University Medical Center Utrecht, Utrecht, Netherlands; 2) Cell Biology, Netherlands Cancer Institute, Amsterdam, Netherlands; 3) Pediatric Nephrology, Wilhelmina Pediatric Hospital, University Medical Center Utrecht, Netherlands; 4) Human Genetics, Radboud University Nijmegen Medical Centre, Nijmegen, Netherlands; 5) Pathology, University Medical Center Utrecht, Netherlands.

Congenital nephrotic syndrome and interstitial lung disease is a rare multiorgan disorder, characterized by disrupted basement membrane structures. Causal *ITGA3* gene mutations were recently identified, but the underlying disease mechanism remained poorly understood. We describe a patient who presented with glomerulosclerosis, proteinuria, and neonatal respiratory distress, who died 7 months after birth due to respiratory insufficiency. A genome-wide screening for structural variations revealed a homozygous region of 19.2 Mb on chromosome 17 that included *ITGA3*. Sanger sequencing of the *ITGA3* gene that encodes integrin α3, revealed a homozygous missense mutation c.1045G>T; p.A349S, which was functionally characterized. Immunostaining in a transfected cell model showed appropriate cell surface localization for the wild-type but no surface localization for the mutant integrin α3. This was consistent with the lack of integrin α3 protein expression in the kidneys of our patient. Moreover, the mutation resulted in a *de novo* formation of an N-glycosylation motif, causing hyperglycosylation of integrinα3. Normally, integrin α3 forms heterodimers with integrin β1, though conformation of the mutant form was disrupted, preventing it from associating with β1 integrin. Increased ubiquitination supports that the mutant integrin α3 gets targeted for degradation before exiting the endoplasmatic reticulum. Our findings underscore the role of integrin α3β1 as the main regulator of basement membrane integrity in kidney tissue. We suggest that hyperglycosylation of integrin α3 finally resulting in complete lack of α3β1 expression on the basement membrane, is a new pathogenic mechanism causal for nephrotic syndrome and interstitial lung disease. Importantly, *ITGA3* gene screening was implemented in DNA diagnostics at our institute, having important consequences for genetic counseling of patients and their relatives.

2849T

Defining the phenotypic specificity of Schimke immuno-osseous dysplasia. *M. Morimoto, C. Myung, K. Beirnes, D. Leung, H.K. Fam, K. Choi, Y. Huang, C.F. Boerkoel.* Department of Medical Genetics, Child and Family Research Institute, University of British Columbia, Vancouver, BC, Canada.

Background: Schimke immuno-osseous dysplasia (SIOD), a rare autosomal recessive multisystem disorder, is caused by mutations in the *SMARCAL1* gene. The SMARCAL1 enzyme diffusely binds open chromatin and globally affects transcription. We hypothesize therefore that SMARCAL1 gives rise to the pathognomonic features of skeletal dysplasia, renal failure, immunodeficiency and vascular disease by pathologically altering the expression of specific genes in the affected tissues.

Methods: We assessed the genetic interactions of *SMARCAL1* with the Notch, Epidermal Growth Factor, Wingless, Decapentaplegic and Hedgehog developmental pathways through an F_1 genetic screen in *Drosophila melanogaster* to measure the suppression or enhancement of ectopic wing veins induced by the overexpression of SMARCAL1. We also assessed the expression of key developmental pathways in SIOD autopsy tissue through gene expression arrays, quantitative reverse transcriptase PCR and immunofluorescence.

Results: Mutations of the Notch pathway strongly enhanced, mutations of the Wingless and Epidermal Growth Factor pathways strongly suppressed, and mutations of the Decapentaplegic and Hedgehog pathways had no effect or weakly suppressed the ectopic wing vein phenotype of SMARCAL1 overexpression. Consistent with these observations, SMARCAL1-deficient patient kidney showed increased expression of several Notch signaling pathway members and targets.

Conclusion: We conclude that SMARCAL1 deficiency preferentially alters expression of specific genes in different tissues and thereby causes the pathognomonic features of SIOD. Specifically, we postulate that SMARCAL1 deficiency increases Notch signaling in the kidney to induce the focal segmental glomerulosclerosis of SIOD.

2850F

High-throughput mutation analysis in patients with a nephronophthisis-associated ciliopathy applying multiplexed barcoded array-based PCR amplification and next-generation sequencing. *E.A. Otto[1], J. Halbritter[1], K. Diaz[1], B. Tarrier[2], M. Chaki[1], C. Fu[1], J.L. Innis[1], S.J. Allen[1], R.H. Lyons[2], C.J. Stefanidis[3], H. Omran[4], N.A. Soliman[5,6].* 1) Dept Ped, Univ Michigan, Ann Arbor, MI; 2) DNA Sequencing Core, Univ Michigan, Ann Arbor, Michigan; 3) Pediatric Nephrology, "A. and P. Kyriakou" Children's Hospital, Athens, Greece; 4) Dept Ped and Adolescent Medicine, Univ Hospital Münster, Münster, Germany; 5) Center of Pediatric Nephrology & Transplantation, Cairo Univ, Cairo, Egypt; 6) Egyptian Group for Orphan Renal Diseases (EGORD), Cairo, Egypt.

Objective: To identify disease causing mutations within coding regions of 11 known NPHP genes (*NPHP1-NPHP11*) in a cohort of 192 patients diagnosed with a nephronophthisis-associated ciliopathy, at low cost. **Methods:** Mutation analysis was carried out using PCR-based 48.48 Access Array microfluidic technology (Fluidigm) with consecutive next generation sequencing. We applied a 10 fold primer multiplexing approach allowing PCR-based amplification of 480 amplicons (251 exons) for 48 DNA samples simultaneously. After 4 rounds of amplification followed by indexing all 192 patient-derived products with different barcodes in a subsequent PCR, 2×100 paired end sequencing was performed on one lane of a HiSeq2000 instrument (Illumina). Bioinformatics analysis was performed using "CLC Genomics Workbench" software. Potential mutations were confirmed by Sanger sequencing, shown to segregate with the affected status, and were absent from 96 control individuals. **Results:** NGS sequencing and bioinformatics analysis resulted in 62.1 million mapped reads with a median exon coverage depth of 449× per patient. Furthermore, 168/192 DNA samples (87.5%) and 134/251 exons (93.2%) consistently showed sufficient coverage depth above 30×, required for reliable heterozygous mutation calls. We identified pathogenic mutations in 35/192 patients (18%) and discovered 23 novel mutations in the genes *NPHP3* (7), *NPHP4* (3), *IQCB1* (4), *CEP290* (5),*RPGRIP1L* (1), *SDCCAG8* (2) and *TMEM67* (1). Additionally, we found 40 different single heterozygous missense variants of unknown significance. **Conclusions:** We conclude that the combined approach of array-based multiplexed PCR-amplification on a Fluidigm Access Array platform followed by NGS is highly cost-efficient and strongly facilitates diagnostic mutation analysis in broadly heterogeneous Mendelian disorders.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2851W

Founder mutation in *RPSPH4A* identified in patients of Hispanic descent with Primary Ciliary Dyskinesia. *M.A. Zariwala[1], M.W. Leigh[2], S.D. Davis[2], M.C. Armstrong[3], M.L. Daniels[3], J.L. Carson[3], M. Hazucha[3], S.S. Dell[4], M.R. Knowles[3], Genetic Disorders of Mucociliary Clearance Consortium.* 1) Department of Pathology & Laboratory Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA; 2) Department of Pediatrics, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA; 3) UNC School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA; 4) The Hospital for Sick Children, Toronto, Ontario, Canada.

Primary ciliary dyskinesia (PCD) is an autosomal recessive, genetically heterogeneous disorder manifested by neonatal respiratory distress, otitis media, sinusitis, situs abnormalities (in ~50% term Kartagener syndrome), bronchiectasis, and male infertility. The diagnosis of PCD has traditionally been based on defining ciliary ultrastructural defects. Establishing a diagnosis is challenging as at least one-third of PCD patients are now recognized to have normal ciliary ultrastructure and/or questionable central apparatus (CA) defects (apparent in only a small portion of cilia). In PCD, nasal nitric oxide (nNO) levels are low (<100nl/min versus ~300nl/min for controls). Genetic testing is emerging as a diagnostic tool, and currently 14 genes are known to harbor mutations that account for two-thirds of PCD cases. Many of these genes have not been adequately tested in the large cohort of patients. *RSPH4A* encodes radial spoke head protein 4A, and mutations are reported in few PCD patients with CA defects. We performed mutation profiling in 32 unrelated patients who had a clinical phenotype of PCD (including situs abnormalities in 7 patients) and low nNO levels (<100 nl/min). Ciliary ultrastructure was variable in the 32 patients; 17 normal, 5 with CA defects and 10 with questionable CA defects. Sequencing of all 6 exons and splice-sites identified novel biallelic mutations in 4 patients of Hispanic descent of which 3 were homozygous for c.921-3_6delAAGT mutation and one was compound heterozygous for c.921-3_6delAAGT and c.1732_1733delG mutations. All of these patients had questionable CA defect. One patient of Caucasian descent was heterozygous for novel p.S39X mutation. None of the patients harboring mutations had situs abnormalities. Interestingly, 7 of the 9 mutant alleles were identical and present in patients of Hispanic descent, suggesting this may be a founder mutation. Reverse transcriptase PCR analysis on total RNA from lymphoblastoid cell lines from a patient harboring homozygous c.921-3_6delAAGT mutation revealed out-of-frame deletion of exon 2 (r.687_921del) leading to premature translation termination signal. In conclusion, IVS2-3_6delAAGT splice-site mutation in *RSPH4A* is a common cause of PCD in patients of Hispanic descent with questionable ciliary ultrastructure, and genetic testing for this mutation can ascertain PCD. This abstract was funded by MO1RR00046, UL 1 RR025747, R01HL071798, U54 HL096458-06, R01HL094976-02 and 5UC2HL102923-02.

2852T

Leri's pleonosteosis results from defective SMAD signaling. *S. Banka[1], S. Cain[2], J. Urquhart[1], S. Carim[1], S. Daly[1], G. Erdem[1], M. Bottomley[1], D. Donnai[1], B. Kerr[1], H. Kingston[1], C. Kielty[2], W.G. Newman[1].* 1) Department of Genetic Medicine, University of Manchester, Manchester, United Kingdom; 2) Faculty of Life Sciences, Wellcome Trust Centre for Cell-Matrix Research, Manchester, United Kingdom.

We report the novel molecular and phenotypic characterization of Leri's Pleonosteosis [MIM 151200]. This rare autosomal dominant condition is characterized by flexion contractures of the interphalangeal joints, limited motion of multiple joints, chronic joint pains, facial dysmorphism, bony overgrowths, short broad metacarpals, metatarsals and phalanges, short stature and scleroderma-like thickening of skin. We genotyped the two most distantly related individuals in a multigenerational family with Leri's pleonosteosis by Affymetrix SNP6.0 array. Copy number analysis revealed a shared, approximately, 1Mb duplication of chromosome 8q22.1. The tandem duplication segregated with the phenotype within the family, as confirmed by QPCR and FISH. This duplication was not present in copy number polymorphism databases or in over 500 controls. We identified an overlapping 8q22.1 duplication in an unrelated patient with Leri's pleonosteosis to confirm the causal relationship. The minimum critical region consists of six genes, including GDF6 (growth/differentiation factor 6) and SDC2 (syndecan-2). GDF6 belongs to a subgroup of BMP signaling molecules that include GDF5 and 7, which are important for joint formation during embryonic development. Syndecan-2 is a transmembrane heparan sulfate proteoglycan with a known role in linking cellular interactions with extra-cellular matrix. We show that the causal duplication in Leri's pleonosteosis results in dysregulated SMAD signaling in dermal fibroblasts. The phenotypic descriptions of Leri's pleonosteosis have been inconsistent and our findings demonstrate, for the first time, a single underlying molecular basis for the phenotypic spectrum due to dysregulated SMAD/TGF-β signaling. Previously, loss of function mutations of GDF6 were described in Klippel-Feil syndrome. Our results demonstrate the importance of the balance between BMP signaling during joint development and role of syndecan-2 in this pathway. Leri's pleonosteosis can now be added to a growing list of disorders, including stiff skin syndrome, Myhre syndrome, acromicric dysplasia and geleophysic dysplasia, with overlapping phenotypes and dysregulated SMAD/TGF-β signalling pathways. These findings provide an opportunity to trial TGF-β inhibitors and provide further insight into the role of TGF-β in connective tissue ageing and in arthritis and scleroderma.

2853F

Abnormal Type I Collagen Folding and Matrix Deposition in a Cyclophilin B KO Mouse Model of Recessive Osteogenesis Imperfecta. *W.A. Cabral[1], E. Makareeva[2], M. Weis[3], S. Leikin[2], D.R. Eyre[3], J.C. Marini[1].* 1) Bone & Extracellular Matrix Branch, NICHD, NIH, Bethesda, MD; 2) Section on Physical Biochemistry, NICHD, NIH, Bethesda, MD; 3) Orthopaedic Research Laboratories, University of Washington, Seattle, WA.

Osteogenesis Imperfecta (OI) is a heritable disorder characterized by bone fragility and growth deficiency. Recessive OI is due to mutations in genes encoding proteins involved in collagen post-translational interactions. The most frequent cases (Types VII-IX OI) involve the components of the collagen prolyl 3-hydroxylation complex, CRTAP, prolyl 3-hydroxylase 1 (*LEPRE1*) and cyclophilin B (CYPB; *PPIB*). The function of α1(I) Pro986 3-hydroxylation is unknown, but a role in alignment and cross-linking of fibrils in matrix is speculated. PPIB, a prolyl cis-trans isomerase, is thought to be the major PPIase responsible for the rate-limiting step in collagen helix formation. To further characterize the role of 3-hydroxylation complex components in collagen folding, *Ppib*-null mice were generated using a gene-trap ES cell clone with a β-geo reporter construct inserted in the first intron of *Ppib*. The heterozygous clone was shown by real-time RT-PCR to have half-normal *Ppib* expression. Homozygous *Ppib* knockout mice were verified by RT-PCR to have complete absence of *Ppib* transcripts in skin, fibroblasts, femurs and calvarial osteoblasts. In growth curves from 2 wks-6 mos of age, *Ppib*-/- mice weigh 33% less than WT and Het littermates. Ppib protein was absent and P3h1 reduced 50% reduction on western blots of -/- cultured fibroblast and osteoblast lysates, with half-normal amounts of Ppib in +/− cells. As expected, α1(I) P986 3-hydroxylation of fibroblast and osteoblast collagen was severely reduced (11 and 5% of WT, respectively). In agreement with previously described patients with decreased P986 3-hydroxylation, collagen from *Ppib*-/- cells had delayed helix formation in direct intracellular folding assays with the expected retarded electrophoretic mobility on SDS-Urea PAGE. However, both thermal stability and 5-lysyl and 4-prolyl hydroxylation of α1(I) chains were normal, suggesting collagen post-translational glycosylation might be altered. In *Ppib*-/- fibroblast and osteoblast cultures, acid-labile and pepsin-soluble collagen deposition was decreased to 20–30% of WT cultures and cannot be attributed to reduced secretion, which was only moderately delayed on pulse-chase analyses of secretion kinetics. These data suggest unique roles for Ppib in collagen post-translational processing, trafficking and extracellular matrix incorporation in addition to a role as a folding chaperone.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2854W

Differential contributions of P3H1 to collagen processing in the pathogenesis of Osteogenesis Imperfecta. *EP. Homan[1], C. Lietman[1], I. Grafe[1], J. Lennington[1], R. Morello[2], WR. Hogue[2], D. Napierala[3], MM. Jiang[1,4], B. Dawson[1,4], TK. Bertin[1], Y. Chen[1], R. Lua[1], O. Lichtarge[1], J. Hicks[5], MA. Weis[6], DR. Eyre[6], B. Lee[1,4].* 1) Molecular and Human Genetics, Baylor College of Medicine, Houston, TX; 2) Physiology and Biophysics, University of Arkansas for Medical Sciences, Little Rock, AR; 3) Oral and Maxillofacial Surgery, School of Dentistry, University of Alabama at Birmingham, Birmingham, AL; 4) Howard Hughes Medical Institute, Baylor Collage of Medicine, One Baylor Plaza, Houston, TX; 5) Pathology, Texas Children's Hospital, Baylor College of Medicine, Houston, TX; 6) Orthopædics and Sports Medicine, University of Washington, Seattle, WA.

Mutations in Cartilage Associated Protein (CRTAP), prolyl 3-hydroxylase 1 (P3H1), and Cyclophilin B (CypB) have been identified as causative of Osteogenesis Imperfecta (OI). These proteins form a complex that functions to 3-hydroxylate a single proline residue on the alpha 1 chain of type I collagen and acts as a cis/trans isomerase (PPIase) activity, a function essential for proper collagen folding. This 3-prolyl hydroxylation may not be integral to the structural integrity of collagen; however, its absence may disrupt protein-protein interactions integral for proper collagen folding. P3H1 and CRTAP stabilize each other and when one is absent, the other is degraded. Mutations in this complex result in hypomorphic or loss of function (LOF) alleles. It is unclear what roles the 3-hydroxylation and/or the PPIase play in the pathogenesis of OI. This project studies the contribution of P3H1 hydroxylation and how it pertains to the pathogenesis of OI by inactivating the hydroxylase domain while maintaining the protein and complex structure. We identified a residue that is considered critical for the hydroxylase function and generated a mutant P3H1 cDNA construct containing an alanine substitution at this residue. We demonstrated by immunofluorescence and Western blot that we rescued the stability of CRTAP in the LOF cells transduced with WT or mutant P3H1 cDNA. Analysis of prolyl 3-hydroxylation by collagen mass spectrometry indicates that the enzymatic activity of P3H1 is abolished. These data suggest that the mutant can maintain the stability of the P3H1 complex, thus maintaining the PPIase function, while abolishing enzymatic activity. To understand how absent 3-hydroxylation pertains to the pathogenesis of OI, recombineering was utilized to introduce the same alanine substitution into the P3H1 locus. In our knock in (KI) mice, we observed normal modification, which is in contrast to the P3H1 KO mice. Compared to control mice and in contrast to P3H1 null mice, the KI mice are normal in appearance and growth. However, analysis of the trabecular bone shows reduced BMD with decreased Tb Th, increased Tb Sp, and decreased BV/TV, whereas cortical bone parameters are normal. We have established a novel mouse model that recapitulates the bone phenotype but not the cartilage, growth, and connective tissue phenotype of recessive OI. These observations support differential tissue contributions of P3H1 hydoxylation to collagen processing in the pathogenesis of OI.

2855T

Characterization of transient receptor potential vanilloid channel 4 (TRPV4) in metatropic dysplasia primary cell lines. *L.M. Hurd[1, 2], S.M. Kirwin[1], K.M. Vinette[1], M.B. Bober[3], W.G. Mackenzie[4], R.L. Duncan[2], V.L. Funanage[1].* 1) Molecular Diagnostics Laboratory, Nemours/Alfred I. duPont Hospital for Children, Wilmington, DE; 2) Biological Sciences, University of Delaware, Newark, DE; 3) Division of Genetics, Department of Pediatrics, Nemours/Alfred I. duPont Hospital for Children, Wilmington, DE; 4) Department of Orthopedics, Nemours/Alfred I. duPont Hospital for Children, Wilmington DE.

Transient receptor potential vanilloid 4 (TRPV4) is a non selective calcium permeable ion channel that is an important mediator of both chemical and physical stimuli. Autosomal dominant mutations in TRPV4 have been shown to cause a spectrum of skeletal dysplasias that includes brachyolmia, spondylometaphyseal dysplasia, Kozlowski type (SMDK) and metatropic dysplasia. Metatropic dysplasia is one of the more severe TRPV4 dysplasias and is characterized by short stature, metaphyseal widening, marked platyspondyly, and progressive kyphoscoliosis. Electrophysiologic studies have shown the channel to be constitutively active, resulting in increased baseline calcium in cells transfected with TRPV4 mutations. However, to date there have not been any published results on the activity of the channel in primary cell lines from patients with metatropic dysplasia. We have a cohort of patients diagnosed with metatropic dysplasia who carry TRPV4 mutations, and have isolated cells from skin, bone and cartilage tissues of these patients. Using conventional PCR genetic sequencing analysis, we have identified a novel mutation, a G800D mutation in exon 15, codon 800 (G>A) resulting in a non-synonymous change from glycine to aspartic acid. Using quantitative multiplex fluorescent PCR, we have also identified a novel duplication of TRPV4 in which the duplication encompasses the entire gene. Protein and mRNA isolated from biopsies has confirmed TRPV4 is expressed in cells derived from skin, bone and cartilage using western blot and rt-PCR analysis. Intracellular calcium imaging with fura-2 showed that TRPV4 mutant cells have increased baseline calcium concentrations when compared to controls. Furthermore, TRPV4 mutant cells exhibit a decreased response to the TRPV4 agonist 4 alpha-phorbol 12, 13-didecanoate (4αPDD) when compared to control cells ($p<0.01$). Additionally, we would like to further characterize channel activity with patch clamping studies. Characterizing channel activity in primary cells from patients with metatropic dysplasia provides insight into the cellular mechanisms associated with this pathology. Once endogenous channel activity is defined, the mechanism by which aberrant channel activity contributes to metatropic dysplasia can be further elucidated and provide potential therapeutic targets.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2856F

Exome sequencing identifies a *de novo* missense mutation in *TGF-β1* in a sporadic patient with craniotubular bone dysplasia. *D. Li[1,8], A. Haghighi[2,8], F. Salehzade[3], L. Tian[1], H. Haghighi-Kakhki[4], C. Kim[1], E. Frackelton[1], F. Otieno[1], H.-D. Nah[5,6,9], H. Hakonarson[1,7,9].* 1) The Center for Applied Genomics, Abramson Research Center, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA; 2) The Wellcome Trust Center for Human Genetics, University of Oxford, Oxford, UK; 3) Department of Pediatrics, Ardabil University of Medical Sciences, Ardebil, Iran; 4) Faculty of Medicine, Mashhad Azad University, Mashhad, Iran; 5) Division of Plastic and Reconstructive Surgery, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA; 6) Division of Plastic Surgery, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania, USA; 7) Department of Pediatrics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania, USA; 8) These authors contributed equally and considered joint first authors; 9) Equal contribution.

Autosomal dominant craniometaphyseal dysplasia (CMD, [OMIM 123000]) is a rare genetic disorder characterized by early progressive hyperostosis and sclerosis of cranial bones associated with abnormal modeling of the metaphyses. Generally radiographic examination in patients with CMD manifests diffuse hyperostosis of the cranial base, cranial vault, facial bones, and mandible and osseous obliteration of paransal sinuses and mastoid air cells. Additionally, patients may have other features, including facial paralysis and blindness resulting from cranial nerve compression, and hearing loss. Several mutations in *ANKH*, which is the only gene known to be associated with CMD so far, were found in the patients that have an affected parent. However, it's likely that the simplex cases have de novo mutations in *ANKH* and other genes. We captured exomes of a patient with the initial diagnosis of CMD based on typical metaphyseal widening feature and skull bone sclerosis, on which mutation analysis in *ANKH* was negative, together with the unaffected parents and brother, and subsequently performed sequencing analysis by Hiseq2000 with an effective mean coverage of 65 folds. The variants being studied were filtered against the 1000 Genomes Project, 5400 exomes from NHLBI GO Exome Sequencing Project and Complete Genomics 69 whole genome panel. We obtained 12 *de novo* variants in the proband, who doesn't have a family history of the disease. Of those variants, the only fitting disease-causing candidate for this case is Transforming Growth Factor-β1 (*TGF-β1*), which is a multi-functional protein acting on cell growth, differentiation, and morphogenesis of many different cell types. The novel mutation in *TGF-β1* affected a highly conserved residue of predicted damaging as determined by various algorithms, including both SIFT and Polyphen. Follow-up Sanger sequencing confirmed the heterozygous state of the mutation only in proband. Interestingly, different mutations in *TGF-β1* are known to cause the Camurati-Engelmann disease (CED), progressive diaphyseal dysplasia diagnosed by hyperostosis and sclerosis of the diaphyses of long bones. As the CED progresses, the metaphses may be also affected. A closer scrutiny of the phenotype of this proband is under way as a potential form of atypical presentation of CED. We are currently conducting follow-up functional studies to elucidate the disease causing mechanisms that will be presented together with detailed phenotypic characterization.

2857W

An Intracellular Role for FGFR2 in Bent Bone Dysplasia Syndrome. *C.L. Neben[1,2], B. Idoni[1,2], A. Sarukhanov[3], D. Krakow[3,4], A.E. Merrill[1,2].* 1) CCMB, Ostrow School of Dentistry, USC, Los Angeles, CA; 2) Department of Biochemistry, Keck School of Medicine, USC, Los Angeles, CA; 3) Department of Orthopaedic Surgery, David Geffen School of Medicine, UCLA, Los Angeles, CA; 4) Department of Human Genetics, David Geffen School of Medicine, UCLA, Los Angeles, CA.

Fibroblast Growth Factor Receptor 2 (FGFR2) is a crucial regulator of bone formation during development. While it is known that FGFR2 coordinates osteoprogenitor cell proliferation and differentiation, it remains unclear how the receptor couples these distinct cellular processes. Most FGF ligands mediate their biological effects extracellularly by activating membrane bound FGFRs. Upon activation, FGFRs autophosphorylate intracellular tyrosine residues and recruit adapter proteins juxtamembrane to launch ERK1,2, PLCγ, and PI3K/AKT cascades that subsequently regulate cell growth and differentiation. Some FGFs, such as FGF2, lack a signal sequence and can reside within the cell to regulate proliferation through nuclear routes presumed to be FGFR independent. Through an analysis of the pathophysiology of the recently described lethal skeletal disorder Bent Bone Dysplasia-*FGFR2* type (BBD), we have identified a new activity of FGFR2: our data suggests that the receptor acts within the nucleus to regulate bone formation. We previously showed that the dominant *FGFR2* mutations in BBD reduce receptor levels at the plasma membrane and markedly diminish responsiveness to extracellular FGF2. In this study we find that the mutations concomitantly enhance receptor localization to the nucleus and alter responsiveness to intracellular FGF2. We define how dual activities of FGFR2 at the plasma membrane and in the nucleus regulate bone formation by employing the mutant receptor in both an *in vitro* and *in vivo* approach. Using an osteoblast culture assay, we determine how FGFR2 activation of the ERK1,2, PLCγ, and PI3K/AKT cascades are differentially regulated by the mutant receptor. Utilizing the chick embryonic system, we uncover the influence of the mutant FGFR2 on skeletal development by driving targeted overexpression in the osteoprogenitor cells through *in ovo* electroporation. Together our findings support a distinct role for nuclear FGFR2 during the proliferation and differentiation of osteoprogenitor cells during bone formation.

2858T

ZDHHC13 is a Key Regulator in Postnatal Epiphyseal Development. *I.W. Song[1,2,3], W.R. Li[1,2], J.Y. Wu[1,2], M.T. Lee[1,2], Y.T. Chen[1,2].* 1) Institute of Biomedical Sciences, Academia Sinica,Taipei, Taiwan; 2) National Center for Genome Medicine, Academia Sinica,Taipei, Taiwan; 3) Graduate Institute of Life Sciences, National Defense Medical Center, Taipei, Taiwan.

Palmitoylation is a reversible post-translational 16-carbon fatty acid modification of the protein and is involved in a variety of physiological processes. It is mediated by DHHC-containing palmitoyl acyl-transferases (PATs) family of enzymes. We have previously shown that Zdhhc13 is a PAT and mice carry recessive Zdhhc13 mutation results in alopecia, amyloidosis and osteoporosis. Here we investigated the pathogenic mechanism of the severe osteoporosis and growth retardation in these mice. Histology of Zdhhc13 mutant mice at 2 weeks of age displayed disorganization in growth plate structure caused by narrowed hypertrophic zone and reduced proliferating chondrocyte and defective ecdochondral ossification and significant delay in the development of secondary ossification center (SOC). In addition, we demonstrated that Zdhhc13 expressed in and regulated the differentiation and function of osteoclasts, osteoblasts and chondrocytes. This regulation was most evidently at 2 weeks of age. In summary, we demonstrated that Zdhhc13 is a novel regulator in postnatal skeletal development. Moreover, we provided the first in vivo and in vitro evidences that palmitoylation acts as a key temporal regulator in skeletal development. This finding also reveals a new aspect in osteoporosis pathogenesis.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2859F

Identification of novel genes downstream of FGFR2 contributing to coronal suture synostosis in a murine model for Crouzon syndrome. *S. Kumar, E. Peskett, J.A. Britto, E. Pauws.* Neural Development Unit, Institute of Child Health, London, London, United Kingdom.

The C342Y mutation in the 3rd Ig-like domain of FGFR2 is the most frequent genetic cause of Crouzon syndrome(CS). Patients present with a combination of coronal suture synostosis and midfacial hypoplasia that frequently requires complex craniofacial surgery. Although much is known about FGF signalling in developing cranial sutures and the biochemical effect of FGFR2 mutations causing CS, little is known about genes or complex pathways occuring downstream of aberrant FGFR2 signalling directly responsible for the process of synostosis. By studying the phenotypic serial progression from E16.5 till P21 in a CS mouse model carrying the orthologous C342Y mutation, we found a divergence in phenotype between wild-type (WT) and $Fgfr2^{C342Y/+}$ (Mut) calvaria after E17.5. We hypothesized that Mut E17.5 coronal sutures would be enriched for genes playing a role in the pathogenesis of coronal suture synostosis.Bilateral micro-dissection of WT and Mut mouse calvaria isolated the dorso-medial aspect of the coronal suture with some flanking frontal and parietal bone. RNA was extracted and analysed using an Affymetrix microarray. Initial expression data analysis identified 51 candidate genes that were regulated by at least 1.5-fold ($p<0.05$). We further shortlisted five novel candidate targets (Dpt, Osr1, Nov, Dlk1 and Kera) which were downregulated up to 3.5-fold in mutant coronal sutures. Knowledge of these genes being previously reported as regulatory factors during osteoblast and chondrocyte differentiation made us suspect them to be putative downstream effectors following constitutive activation of the FGFR2 signalling pathway and play a role in the development and maturation of the coronal suture. These findings were independently validated using Taqman qPCR assays. In-situ hybridizations for Dpt and Osr1 were then performed in sagittal sections of E17.5 WT and mutant heads. Osr1 was found in both fronto-parietal osteogenic fronts outlining frontal and parietal bones, periosteum and skin. Dpt expression was also found in both dermis and outline of periosteum in the WT frontal bone additionally being within the substance of periosteum overlying frontal bone. Further spatio-temporal expression analysis will confirm the precise role these genes may play during the ontogeny of sutural mesenchyme. Our preliminary findings set the stage for directed investigation and analysis of putative effectors downstream to FGFR2 responsible for craniosynostosis in CS.

2860W

Identification of copy number variants in 413 isolated talipes equinovarus patients: role of transcriptional regulators of limb development. *D.M. Alvarado[1], J.G. Buchan[4], S.L. Frick[6], J.E. Herzenberg[7], M.B. Dobbs[1,5], C.A. Gurnett[1,2,3].* 1) Departments of Orthopaedic Surgery; 2) Pediatrics; 3) Neurology; 4) Division of Biology and Biomedical Sciences, Washington University, St. Louis, MO; 5) St. Louis Shriners Hospital for Children St. Louis, MO; 6) Levine Children's Hospital, Charlotte, NC; 7) Sinai Hospital, Baltimore, MD.

Clubfoot is one of the most common congenital musculoskeletal anomalies and has a worldwide incidence of 1 in 1000 births. A genetic predisposition to clubfoot is evidenced by the high concordance rate in twin studies and the increased risk to first-degree relatives. Despite the frequency of isolated clubfoot and the strong evidence of a genetic basis for the disorder, few causative genes have been identified. In this study, we hypothesize that rare and/or recurrent copy number variants are associated with isolated clubfoot and that these copy number variants are likely to contain genes that are important for limb development. We performed a genome-wide screen for rare deletions and duplications in 413 isolated clubfoot patients using the Affymetrix 6.0 array. Segregation analysis and gene expression in mouse E12.5 limb buds were used to determine the significance of copy number variants. We identified 74 rare, gene-containing copy number variants that were present in clubfoot probands and not present in 759 controls or in the Database of Genomic Variants. The overall frequency of copy number variants was similar between clubfoot patients compared to controls. Twelve rare copy number variants segregate with clubfoot in multiplex pedigrees, and contain the developmentally expressed transcription factors and transcriptional regulators *PITX1, TBX4, HOXC13, UTX, CHD1,* and *RIPPLY2*. Our results suggest an important role for genes involved in early embryonic patterning in clubfoot etiology that can now be tested with large scale sequencing methods.

2861T

Mutation of *B3GAT3* causes skeletal dysplasia in a consanguineous clan from Nias. *B.S. Budde[1], S. Mizumoto[2], R. Kogawa[2], C. Becker[1], J. Altmüller[1], H. Thiele[1], P. Frommolt[1,3], M.R. Toliat[1], J.M. Hämmerle[4], W. Höhne[5], K. Sugahar[2], P. Nürnberg[1,3], I. Kennerknecht[6].* 1) Cologne Center for Genomics, Universität zu Köln, Cologne, Germany; 2) Laboratory of Proteoglycan Signaling and Therapeutics, Faculty of Advanced Life Science, Graduate School of Life Science, Hokkaido University, Sapporo, Japan; 3) Cologne Cluster of Excellence on Cellular Stress Responses in Aging-Associated Diseases (CECAD), Universität zu Köln, Cologne, Germany; 4) Yayasan Pusaka Nias, Gunung Sitoli, Nias, Sumatra Utara, Indonesia; 5) Institute for Biochemistry, Charité University Hospital, Berlin, Germany; 6) Institute of Human Genetics, Westfälische Wilhelms Universität, Münster, Germany.

Nias is a unique island belonging to Indonesia. The people of Nias are still organized into patrilineal descent groups or clans. In contrast to other Southeast Asia island populations, studies on genetic diversity in Nias revealed severe bottleneck or founder events in its history giving rise to a high incidence of endemic syndromes. We describe one family with disproportionate short stature and bone dysplasia in two branches. Interestingly, the affected individuals of one of the two branches displayed a milder phenotype than the other. We conducted a genome-wide linkage scan in the branch with the more severe phenotype and determined a candidate region of 10.2 cM on chromosome 11. Sequencing of individual candidate genes did not reveal the disease-causing variant. We then performed whole-exome sequencing in one affected child. By scrutinizing variants in the linkage interval on chromosome 11 we noticed a possibly damaging homozygous missense variant in the gene *B3GAT3*. All affected individuals of the Nias clan are homozygous for the variation that could not be detected in 350 chromosomes of ethnically matched controls. Recently, Baasanjav et al. published their results on a family characterized by joint dislocations and heart defects caused by a different mutation of the same gene, c.830G>A (p.Arg277Gln). Here we describe a second mutation of *B3GAT3* associated with skeletal dysplasia, c.419C>T (p.Pro140Leu). No heart involvement was known but examination of one affected individual is scheduled to figure out if a heart phenotype is present in our family as well. *B3GAT3* encodes β-1,3-glucuronyltransferase (GlcAT-I). GlcAT-I catalyzes an initial step of proteoglucan synthesis by transferring glucuronic acid onto the trisaccharide-protein linkage structure of proteoglycans. The mutation p.Pro140Leu lies within the donor substrate binding subdomain of the catalytic domain of GlcAT-I. A proline-to-leucine exchange could influence the correct folding of the protein since the affected proline residue seems to be located on the surface according to crystallographic data. A multiple alignment shows that p.Pro140 is highly conserved among mammals, amphibians and insects. Functional studies are in progress in order to analyze the degree of damage caused by the mutation at the cellular level. Preliminary results show that the recombinant GlcAT-I (p.Pro140Leu) exhibited a significant reduction in the enzymatic activity compared to the wild-type enzyme.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2862F

Mutation analysis of the *BIN1* gene performed to evaluate the cause of centronuclear myopathy in two fetuses of consanguineous parents from Sudan. *N. Dohrn[1], A. Petersen[2], P. Skovbo[3], I.S. Pedersen[4], A. Ernst[4], H. Krarup[4], M.B. Petersen[1].* 1) Department of Clinical Genetics, Aalborg Hospital, Aalborg, Denmark; 2) Institute of Pathology, Aalborg Hospital, Aalborg, Denmark; 3) Department of Obstetrics and Gynecology, Aalborg Hospital, Aalborg, Denmark; 4) Section of Molecular Diagnostics, Aalborg Hospital, Aalborg, Denmark.

Arthrogryposis multiplexa congenita (AMC) is a rare congenital disorder characterized by multiple joint contractures. The condition can be caused by both genetic and environmental factors. Genetic causes include single gene defects (AD, AR and X-linked), chromosomal disorders and mitochondrial defects. The underlying condition can be a neuropathic abnormality, muscle abnormality, connective tissue abnormality, dermal abnormality, space limitations in uterus, maternal disease or impaired vascularity. The common pathway is the lack of fetal joint movement. Most commonly abortion is performed before exact clinical diagnosis can be made and the etiology is often only detected at autopsy. We present a consanguineous Nubian couple from Wadi Halfa, Sudan. The woman's great grandfather was the grandfather of the husband. The woman was 24 years old at the first pregnancy where ultrasound (US) was used in the 13thgestational week to diagnose AMC, leading to a termination of the pregnancy in week 14+2. Array-CGH was made on the aborted tissue and showed a normal result. In the second pregnancy increased nuchal translucency and AMC were observed during an US scan in the 13th gestational week and pregnancy was terminated in week 13+5. Karyotype was normal 46,XX. A full autopsy including CNS together with muscle histopathological examination of the two foetuses revealed identical malformations including a greater variation in muscle fibre diameter and more centrally located myocyte nuclei than expected for the gestational age, and a diagnosis of centronuclear myopathy, CNM, was proposed. Due to the observation of identical affection of two female foetuses of consanguineous healthy parents, we hypothesize that we are dealing with a case of AR CNM due to a mutation in the *BIN1* gene. Around five previous cases with a histologically and molecular proven diagnosis of CNM with *BIN1* mutation have been reported. None of the reported cases were diagnosed in pregnancy indicating a more severe phenotype in our family and perhaps due to a novel *BIN1* mutation. The *BIN1* gene consists of 20 exons and some of which are alternatively spliced. Currently 14 different transcripts are listed in the ENSEMBL database. We are in the process of carrying out bidirectional sequencing of all 20 exons and exon/intron boundaries. DNA extracted from peripheral blood samples and cell cultures are available from parents and aborted foetuses, respectively.

2863W

Novel ROR2 mutations in patients with autosomal recessive Robinow syndrome. *J. Mazzeu[1, 2], H. Kayserilli[3], C.A. Kim[4], D. Brunoni[5], P.C. Pieri[4], A. Richieri-Costa[6], H. van-Bokhoven[7], H.G Brunner[7], A.M. Vianna-Morgante[2].* 1) Programa de Pós graduação em Ciências Genômicas e Biotecnologia, Universidade Católica de Brasilia, Brasilia, Brazil; 2) Departamento de Genética e Biologia Evolutiva, Instituto de Biociências, Universidade de São Paulo, São Paulo, Brazil; 3) Medical Genetics Department, Istanbul Medical Faculty, Istanbul University, 34093, I{{av}}stanbul, Turkey; 4) Departamento de Genetica, Instituto da Criança, Universidade de São Paulo, São Paulo, Brazil; 5) Departamento de Genética Médica, Universidade Federal de São Paulo, São Paulo, Brazil; 6) Departamento de Genética, Hospital de Reabilitação de Anomalias Crânio-faciais, Universidade de São Paulo, Bauru, Brazil; 7) Radboud University Nijmegen Medical Center Department of Human Genetics 855, Nijmegen, The Netherlands.

Robinow syndrome (RS) is a genetically heterogeneous condition characterized by mesomelic limb shortening associated with distinct facial features and genital anomalies that can be inherited as autosomal dominant or autosomal recessive. The clinical picture in recessive RS is more severe, with more marked limb shortening and costovertebral segmentation defects with rib fusion. Recessive RS is caused by mutations in the *ROR2* gene, mapped to 9q22, which encodes an orphan receptor with tyrosine-kinase activity. We screened exons 1–9 of the *ROR2*gene by Sanger sequencing in five unrelated probands. Three probands had affected sibs. Three novel mutations have been identified: c.323G>A (p.R108Q) in exon 3, which changes an amino acid at the Ig-like domain (two families); c.675delG (p.E293fs*444) in exon 6 and c.2074C>A (p.P692T) in exon 9, both affecting the tyrosine-kinase domain. The two families with the same mutation in exon 3 are allegedly unrelated and a common ancestor could not be identified based on the family history, but a founder effect cannot be excluded since they come from the same region. In the fifth patient we did not identify any mutations by sequencing but we failed to amplify exons 4 and 5. We then performed MLPA for *ROR2* (P179 - MRC Holland) that confirmed homozygosity for a deletion including both exons. This study contributes to the characterization of novel *ROR2* mutations causing recessive RS, and describes for the second time a multiple exon deletion suggesting that MLPA should be included as a screening method for confirmation of a clinical diagnosis of RRS. Financial support: FAPESP (CEPID 98/14254-2), CNPq.

2864T

Whole Exome Sequencing Identifies Mutations in the Nucleoside Transporter Gene *SLC29A3* in Dysosteosclerosis, a Form of Osteopetrosis. *P. Campeau[1], J.T. Lu[2,3], G. Sule[1], M.-M. Jiang[1], Y. Bae[1], S. Madan[1], W. Högler[4], N.J. Shaw[4], S. Mumm[5,6], R.A. Gibbs[1,2], M.P. Whyte[5,6], B.H. Lee[1,7].* 1) Molecular & Human Genetics, Baylor College of Medicine, Houston, TX; 2) Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX, USA; 3) Department of Structural and Computational Biology & Molecular Biophysics, Baylor College of Medicine, Houston, TX, USA; 4) Department of Endocrinology, Birmingham Children's Hospital, Birmingham, UK; 5) Center for Metabolic Bone Disease and Molecular Research, Shriners Hospital for Children, St-Louis, MO, USA; 6) Division of Bone and Mineral Diseases, Washington University School of Medicine at Barnes-Jewish Hospital, St Louis, MO, USA; 7) Howard Hughes Medical Institute, Houston, TX, USA.

Dysosteosclerosis (DSS) is a form of osteopetrosis which distinguishes itself from other forms by the presence of skin findings, platyspondyly, and metaphyseal osteosclerosis with relative lucency of widened submetaphyseal portions of long bones. In two patients with DSS, we identified homozygous or compound heterozygous missense mutations in SLC29A3 by whole exome sequencing. This gene encodes a nucleoside transporter implicated notably in pigmented hypertrichosis with insulin-dependent diabetes (PHID) and H syndrome, conditions with little or no skeletal involvement. This transporter is essential for lysosomal function in mice. We demonstrate the expression of *SLC29A3* in osteoclasts in vivo, and reduced osteoclast differentiation and function in patients with DSS. The marked pleomorphic nature of this nucleoside transport defect suggests a new mechanism for the control of osteoclast differentiation and function.

2865F

Polymorphisms in osteoporosis and their associations with bone metabolism. *R. Richterová[1], P. Cibulková[1], A. Bóday[1], D. Stejskal[1], P. Novosad[2], P. Fojtík[3], P. Hrdý[2].* 1) P&R LAB a.s., Laboratory of molecular biology, Nový Jic{{ak}}ín, Czech Republic; 2) Mediekos Labor s.r.o. Osteological centre Zlín; 3) Centre of Gastrointestinal Care, Vítkovická nemocnice a.s., Ostrava - Vítkovice.

Background: Osteoporosis is a disease of bones that leads to an increased risk of fracture, the bone mineral density (BMD) is reduced, bone microarchitecture deteriorates, and the amount and variety of proteins in bone is altered. Osteoporosis in the Czech Republic is affected approximately 500 000 women and 200 000 men. It is a polygenic disease and subtle genetic polymorphisms might have biological effects in its pathogenesis. Epidemiological studies have shown that the variance in BMD is geneticaly determined. Vitamin D receptor (VDR) BsmI polymorphism, estrogen receptor alpha (ESR1) polymorphism PvuII and LDL receptor-related protein 5 (LRP5) polymorphisms Val667Met and Ala1330Val genes are the most frequently analyzed osteoporosis. From anamnestic indicators a frequency of bone fractures and smoking were followed. Materials and Methods: Patients with osteoporosis was screened for 4 SNPs in 3 genes mentioned above. DNA was isolated from whole blood, and amplified real time PCR using hydrolysation and FRET probes. Osteological parameters were then compared with genotype. Results: A group of 1051 patients from Central Moravia region was analysed. We followed a frequency of polymorphisms occurrence and its relation to individual parameters of bone phenotype, bone mineral density in femoral neck and lumbar spine area, markers of bone remodelation (osteocalcin; C-terminal peptide collagen I, beta crosslaps in patients sera) and risk of fractures were analysed. Demographic characteristics of individual genotype were with no significant difference. The acquired results show significantly lower BMD in homozygotes BB in Femoral Neck and Lumbar Spine area and a significantly lower occurrence of fractures in patients of genotype bb. Relation among polymorphism BsmI and bone turnover markers were not proved. In observed groups we analyzed phenotypes selected bones and discovered statistical significant difference between tested polymorphism. Conclusion: Our findings have proven a similar association among genotype BB of gene polymorphism BsmI for VDR and selected bone phenotypes. In particular, we managed to prove a statistically significant correlation with BMD in Femoral Neck area and Lumbar Spine area. Further on, we have proven a connection with a higher frequency of fractures with this genotype in a population of patients with postmenopausal osteoporosis.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2866W

Combination of Whole-Genone cytogenetics array, High-throughput SNP genotyping, homozygosity mapping and Sanger sequencing identify FXN gene expansion leading to Axonal peripheral neuropathy and ataxia phenotype in two consanguineous families from North West Africa. *H. Azzedine[1], B. Hubert[2], W. Amer[3], F. Ferrat[3], M. Chaouch[3], F. Lagace[4, 5], R. Fetni[4, 5], C. Paisan-Ruiz[6].* 1) Medical genetics, DGM Faculty of Biology and Medicine, CHUV-Unil Lausanne, Switzerland; 2) Department of Neurology, Mount Sinai School of Medicine; 3) service de neurologie, Hopital de Ben Aknoun, Alger, Algeria; 4) PROCREA Cliniques, Montréal, (QC) Canada; 5) CHU saint Justin research center, Montréal, (QC) Canada; 6) Departments of Neurology, Psychiatry, Genetics and Genomic Sciences, Friedman Brain Institute, Mount Sinai School of Medicine.

Friedreich Ataxia (FRDA) is severe autosomal recessive neurodegenerative disorder due to the abnormal expansion of GAA repeats in the FXN gene, leading to heterochromatin gene silencing. The consequence of this mechanism is the very low level of FXN mRNA. In normal conditions, FXN gene product is a small protein (frataxin) with a large specter tissue expression. It's involved in mitochondrial iron metabolism via the biosynthesis of iron-sulfur clusters (ISCs). Iron is necessary for energy production, DNA and neurotransmitters synthesis, phospholipid metabolism and myelination. Loss of myelinated axons in peripheral neurons, the degeneration of posterior columns of the spinal cord and the loss of peripheral sensory fibers constitute some major pathological development of FRDA. Other cardiac, skeletal and sensory abnormalities are also observed. In this study we present 2 consanguineous North African families with a combined phenotype of axonal neuropathy and ataxia as an autosomal recessive trait. We performed both Affymetrix Cytogenetics Whole-Genome 2.7M Array and High-throughput SNP genotyping. The later was carried out using HumanOmniExpress beadchips, containing 733,202 genome-wide tag SNPs, and the HiScanSQ system (Illumina). All genotyping quality assessments were undertaken according to the appropriate options within the Genome studio program (GS; Illumina). Data analysis revealed four regions of homozygosity were identified in each family. However, only one region, in chromosome 9, was shared by both. This region encompasses the FXN gene. By sequencing of all the FXN exons, no mutation was found in any patient. Sequencing of the first intron of the gene revealed an abnormal homozygous expansion of 500 GAA repeats in the 2 affected individuals of the first family. Interestingly, in the second family, one patient had 800 repeats in one chromosome and only 100 in the other, while he is homozygous for the SNP haplotype in the region of interest. His affected sibling is homozygous for a 500 triplet repeats. The parents of each family and their healthy children were in the normal allele expansion range ($8 \leq GAA \leq 20$). We did not investigate the FRDA locus as a primary genetic testing intention because the clinical data did not well overlap with the phenotype of Friedreich Ataxia. More clinical investigations are ongoing to reevaluate the scale of the disease in all the patients, especially this one with “heterozygous” expansion.

2867T

Exome sequencing in small families segregating autosomal dominant non-syndromic hearing loss. *H. Azaiez[1,2], E. Shearer[1], K. Booth[1], M. Hildebrand[3], N. Meyer[1], RJH. Smith[1].*1) Department of Otolaryngology, University of Iowa, Iowa City, IA, USA; 2) Institut Pasteur de Tunis, Tunis, Tunisia; 3) Austin Health, Department of Medicine, University of Melbourne.

In small families segregating autosomal dominant non-syndromic hearing loss (ADNSHL), the OtoSCOPE platform, which includes all genes known to be implicated in ADNSHL, can be used to identify possible deafness-causing mutations. The `solve' rate for these families with OtoSCOPE is 52%, reflecting the fact that deafness is extremely heterogeneous and that there are many mapped loci for which causal genes remain unknown. Here we show that in OtoSCOPE-negative families with insufficient meioses for classical linkage analysis, a combination of segregation analysis-exome sequencing approach can reduce the number of variants of unknown significance (VUSs) to an actionable number. In the family described here, DNA samples were available for six informative meioses making this family too small for classical linkage analysis. After excluding variants in known ADNSHL genes using OtoSCOPE, whole exome sequencing was completed on three patients. The unfiltered analysis of the whole exome data using our Galaxy bioinformatics platform identified 822 non-synonymous, splice site or coding indel variants that were on the `short' list of VUSs for further study, a number too large to consider. However filtering by regions-of-interest based on segregation analysis reduced the possible ADNSHL-causing variants to only five. Sanger sequencing of these five variants identified p.R185P in Homer2 as the only variant that segregated with hearing loss phenotype. The p.R185P variant is located in exon 6, with R185 being highly conserved in vertebrates. Polyphen2 predicts this change to be probably damaging with a score of 0.996, and Mutation Taster predicts the change to be disease causing with a score of 0.999. This case study shows the power of combining a platform like OtoSCOPE to rule out known-genes followed by novel gene discovery coupling segregation analysis with exome sequencing.

2868F

Application of Massive Parallel Sequencing for Mutation Discovery and Genetic Diagnosis of Hereditary Hearing Loss in Chinese DFNA families. *H. Yuan[1], J. Cheng[1], Y. Lu[1], X. Zhou[2], X. Zhang[2], R. Choy[3].* 1) Inst of Otolaryngology, Chinese PLA General Hospital, Beijing, China; 2) Bioinformatics Division, Tsinghua National Laboratory of Information Science and Technology, and Department of Automation, Tsinghua University, China; 3) Department of Obstetrics and Gynaecology, Chinese University of HongKong, HongKong.

Hearing loss is the most common birth defect and sensorineural hearing loss (SNHL) results from malfunction of inner ear. Because of the hereditary features of SNHL, genetic diagnosis has exceptionally high value in the medical practice for detecting hereditary hearing loss. Giving that there are more than 200 known genes associated with SNHL, a method with high throughput diagnostic capability but greatly reduced cost such as massively parallel sequencing (MPS) may replace Sanger sequencing and array in clinical application to identify the genetic cause of SNHL. In this study, we developed a target enrichment method that enables the capture of known regions of the human genome (252 genes) associated with SNHL on a single reaction for high throughput mutation identification by MPS platform. We designed target intervals with Agilent SureSelect platform and ran 12 samples with the MPS kit to find their unknown mutations. We use an in-house built computation pipeline based on GATK to align reads to the human reference genome and call variants from the alignments. For confirming the QC filter, all candidate causal variants were closer examined with IGV software and tested by Sanger sequencing. Although barcode pooling often generates unequal amount of total sequences, we observed per-target read-depths were highly correlated between pooled pairs in the same dilution unit. We observed that mitochondrial DNA which designed as a single target turned out to be over 100 fold deeper than non-MT targets. For majority of our samples, about 40%~50% of the total bases map to MT. For the six DFNA families (2 samples in each family), mutations in TECTA, COCH and MYO6 are shared in 3 families, respectively. Linkage and Sanger sequencing data supported co-segregation of these mutations with the phenotypes in these DFNA families. Three DFNA families did not find the causal mutations by this MPS approach. Four possibilities might be account for the failure of mutation identification in these DFNA families: (1) causal variants on known genes were missed due to incomplete coverage; (2) hearing loss within the same family may be caused by more than one causal variants of known genes; (3) patients with phenocopy denied the real positive variants; (4) novel DFNA genes underlying the hearing loss in these families.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2869W

A new gene for autosomal recessive congenital ichthyosis identified by exome sequencing demonstrates impact of ceramide synthesis and acyl chain length for keratinocyte terminal differentiation. *K. Eckl[1,2,3], R. Tidhar[4], M.L. Preil[5], H. Thiele[6], I. Hausser[7], V. Oji[8], A. Önal-Akan[1], S. Brodesser[9,10], F. Stock[11], R. Casper[1], J. Altmüller[6], P. Nürnberg[6,10], H. Traupe[8], A.H. Futerman[4], H.C. Hennies[1,2,3,10].* 1) Div. of Dermatogenetics, CCG, University of Cologne, Cologne, Germany; 2) Div. of Human Genetics, Innsbruck Medical University, Innsbruck, Austria; 3) Dept. of Dermatology, Innsbruck Medical University, Innsbruck, Austria; 4) Dept. of Biological Chemistry, Weizmann Institute of Science, Rehovot, Israel; 5) Practice for Dermatology Dres. Krnjaic, Merk und Schäfer, Ansbach, Germany; 6) Cologne Center for Genomics, University of Cologne, Cologne, Germany; 7) Dept. of Dermatology, University Hospital of Heidelberg, Heidelberg, Germany; 8) Dept. of Dermatology, University Hospital of Münster, Münster, Germany; 9) Inst. of Med. Microbiology, Immunology and Hygiene, University of Cologne, Cologne, Germany; 10) Cluster of Excellence on Cellular Stress Responses in Aging-associated Diseases, University of Cologne, Cologne, Germany; 11) Inst. of Human Genetics, University Hospital of Leipzig, Leipzig, Germany.

Autosomal recessive congenital ichthyosis (ARCI) is a severe genodermatosis characterized by extensive scaling of the entire skin, variable erythema, and often a collodion encasement at birth. ARCI has been suggested as a disorder characterized by disturbance of the cornified lipid envelope of terminally differentiated keratinocytes. Using a combined strategy of linkage analysis in an extended pedigree of German descent with multiple consanguinity and whole-exome sequencing of samples from two patients, we identified one single homozygous mutation in the candidate region with a >100-fold coverage. It is a missense mutation in *CERS3*, the gene for ceramide synthase 3, which completely co-segregated in the pedigree. Ceramide synthases are a family of tissue and acyl chain length-specific N-acyl transferases. *CERS3* is mainly expressed in skin and testis and synthesizes ceramides of C24 and longer acyl chains. In order to analyse the importance of long chain ceramides for epidermal barrier function, we have modelled the disease in vitro using 3D full-skin equivalents with either patient keratinocytes or keratinocytes with siRNA-mediated knockdown of *CERS3*. Disease models demonstrated only slightly reduced epidermal barrier function compared to wild-type models using Lucifer Yellow penetration assays. This finding is in line with the patient phenotype characterized by generalized superficial, whitish scaling and without any atopic manifestations. Morphological analysis of skin biopsy specimens showed mild acanthosis and orthohyperkeratosis with a moderate thickening of the stratum corneum as well as cleft-like inclusions, partly containing melanosomal remnants, in horny lamellae and irregular vesicular structures in the granular layer. The ceramide profile of patient keratinocytes showed a specific loss of C26 NS and ND ceramides and an increase in C16 ceramides. We have found loss of C26 ceramide synthesis in cells overexpressing mutant *CERS3*, compared to cells transfected with wild-type *CERS3*, which demonstrates that loss of CerS3 activity is the underlying cause for these changes. Taken together, we have identified *CERS3* as a new gene for ARCI. Our findings point out that synthesis of long acyl chain ceramides through CerS3 is a crucial first step for formation of the cornified lipid envelope, which then involves other enzymes mutated in forms of ARCI such as 12R-lipoxygenase, epidermal lipoxygenase-3, and transglutaminase-1.

2870T

The NIH Undiagnosed Diseases Program: Defining Pathogenicity for Personalized Medicine. *C.F. Boerkoel III, M.C. Malicdan, D. Adams, C. Toro, C. Tifft, W.A. Gahl, T. Markello.*NIH Undiagnosed Diseases Program, NIH/NHGRI, Bethesda, MD.

Human disease arises by maladaptation of humans to their ecological niche. Rare diseases, which affect ~8% of the population, frequently arise from strong genetic and epigenetic mutations causing maladaptation within a stable ecological niche. Within this context, a precise diagnosis is the first step to understanding illness and defining appropriate therapies. The NIH Undiagnosed Diseases Program uses genomics to approach diagnoses for ill individuals whose diseases have otherwise eluded identification. Even extending the genomic analysis to the community of the family, we confront the difficulty of defining what sequence variants are responsible for disease. In the absence of an epidemiological association, we consider four levels of evidence for pathogenicity: 1) statistical likelihood that the variant will occur in healthy individuals; 2) empiric evidence for pathological consequences of the variant on protein function; 3) associative evidence from systems biology considerations; and 4) recapitulation of disease in a model organism. Based on a study of a rare distal myopathy identified in a single family, we conclude that, in some instances, a systems biology approach garners sufficient evidence to prove disease causation and thereby abrogates the need for epidemiological association or attempt at full recapitulation of the disease in model organisms.

2871F

Efficient search for allelically heterogeneous disease genes in whole genome or exome datasets. *K. D'Aco, M.A. Aquino, N.M. Pearson.* Knome, Inc., Cambridge, MA.

Intro: Searches for disease genes often focus on single kindreds, which implicate one causal variant at a time; or large case/control cohorts. which are hard to gather for rare diseases, and may incur noise from heterogeneous and/or weakly penetrant etiology. Rare diseases that trace to diverse penetrant inherited or de novo variants in one or a few genes, however, can be powerfully studied in modest-sized samples of multiple unrelated probands, often with closely related controls. But without strong prior information (e.g., from linkage mapping), such studies require robust formal methods for ranking candidate genes with etiology too varied for single-kindred methods, and too penetrant to be efficiently discovered by conventional rare variant burden tests. Method: We developed a candidate gene search method that, unlike other rare variant search methods, leverages the fact that, in diseases tracing to rare penetrant variants in one or a few genes, no case is likely to carry more than one (if dominant/recessive) or two (if compound heterozygous) causal variants. We presume that, among a set of k probands, a causal gene harbors a set of j<=k (dominant/recessive) or j<=2k (compound heterozygous) functionally suspect variants, where each proband — but no control - carries a minimal set of one or two such variants at appropriate zygosity. We then search for genes that best fit those criteria among cases but not controls, scoring each gene by how closely it satisfies them, and by a permutation-based p-value to exclude genes that likewise satisfy random case/control distributions. Genes are then jointly ranked by these numbers to shortlist candidates for further investigation. Discussion: Using this method we successfully shortlisted the causal gene in published exome data from ten Kabuki syndrome patients. Our method readily incorporates variant weights and filters, sample kinship, and recessive/compound heterozygous searches. And it can take as input arbitrary groupings of variants, such as multi-gene sets or regulatory regions.

2872W

From large-scale clinical recruitment to identification of a novel mutation in ADAMTSL2 in a family with Ehler-Danlos Syndrome. *A.S. Desai[1], D. Hadley[1], L. Tian[1], C. Kim[1], E. Frankelton[1], R. Chiavacci[1], F.G. Otieno[1], F. Mentch[1], H. Hakonarson[1], K. Wang[2], S. Pannosian[1].* 1) The Center for applied Genomics, CHOP, Philadelphia, PA. 19104, USA; 2) University of Southern California, Los Angeles, CA, 90007, USA.

The Center for Applied Genomics at The Children's Hospital of Philadelphia conducts large-scale clinical recruitment of subjects for genomic analysis. Biological samples are drawn from children and clinical surveys and electronic medical records (EMRs) are linked to those collected samples. Clinical surveys and samples from parents and siblings are also collected when possible. For > 85% of subjects recruited consent to re-contact is given. Through systematic data mining of >80K EMRs from subjects, we identified samples for a family with Ehler-Danlos Syndrome (EDS) with permission to re-contact. EDS is a rare dominant connective tissue disorder that primarily affects skin, joints, muscle and blood vessels. To date, only nine mutations for EDS are known with limited available clinical molecular tests. Furthermore, EDS is known to be associated with osteogenesis imperfecta (OI). To identify the underlying genetics of EDS in the family we identified (negative for known mutations), the parents were re-contacted to verify the family's affected status. The family consisted of an affected male proband, with an affected father, unaffected mother, unaffected sister, and unaffected maternal half brother. We performed exome sequencing on all family members and used ANNOVAR to identify non-synonymous variants that fit a dominant inheritance pattern. We identified 50 of these variants absent from public databases, one of which was a novel missense mutation within the ADAMTS like 2 (ADAMTSL2) gene. Known mutations for the rare type 7C form of EDS includes ADAMTS2 which belongs to the same protein family as the defective ADAMTSL2 gene we found. Sanger sequencing of 9 other EDS samples in our biobank showed the novel heterozygous mutation was unique to the affected family. Our work is significant not only because it illuminates the pathophysiology of EDS for the affected family and the field at large, but also because it demonstrates the utility of unbiased large-scale recruitment and biobanking in understanding rare Mendelian diseases. Families with most other Mendelian disorders for which genetic variants have been found were biasedly recruited by clinical experts; but the majority of rare Mendelian diseases may go unsequenced under such a recruitment model. With unbiased large-scale clinical recruitment we strive to sequence as many rare Mendelian diseases as possible, and this work in EDS serves as a successful proof of concept to that effect.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2873T

Exome sequencing for the knowledge-driven and unbiased disease-mutation search of mitochondrial respiratory chain disorder. *M. Kohda[1], Y. Tokuzawa[2], Y. Moriyama[2], H. Kato[3], N. Uehara[2], Y. Nakachi[1], N. Matoba[1], S. Tamaru[4], Y. Kishita[2], T. Yamazaki[5], M. Mori[6], K. Murayama[7], Y. Mizuno[2], A. Ohtake[5], Y. Okazaki[1,2].* 1) Saitama Medical University, Hidaka, JapanDivision of Translational Research, Research Center for Genomic Medicine, Saitama Medical University, Hidaka, Saitama, Japan; 2) Division of Functional Genomics & Systems Medicine, Research Center for Genomic Medicine, Saitama Medical University, Hidaka, Saitama, Japan; 3) Division of Developmental Biology, Research Center for Genomic Medicine, Saitama Medical University, Hidaka, Saitama, Japan; 4) Department of Obstetrics and Gynecology, Saitama Medical University Hospital, Moroyama, Saitama, Japan; 5) Department of Pediatrics, Saitama Medical University, Moroyama, Saitama, Japan; 6) Department of Pediatrics, Jichi Medical University; 7) Department of Metabolism, Chiba Children's Hospital.

Mitochondria are small organelles and serve as the powerhouse of the living cells because they generate vital energy in their respiratory chain system consisting of protein complexes I to V. Mitochondrial respiratory chain disorder (MRCD) is an intractable disease that develops in childhood. It is a highly frequent inborn error of metabolism that occurs in one out of every 7,000 births. Prominent symptoms develop in such organs as the brain, heart, and muscles, where a great deal of energy is required. In most cases, effective treatment has not been established yet. It is known that various gene abnormalities cause defects of the protein complexes of the respiratory chain, which results in mitochondrial dysfunction. However, identification of the causative gene and the understanding of pathogenic mechanism of MRCD remain largely unsolved. We have started a project to reveal the pathological mechanism of MRCD at the neural cell level using samples obtained from patients with MRCD. In this study, various data and findings such as information obtained using high-throughput sequencers, knowledge of genome informatics and cell biology, and clinical observations will be utilized to understand why various symptoms develop in a tissue-specific manner in patients with MRCD, while mitochondria are ubiquitously existent in all cells. We report current progress and strategy of this project, which uses not only knowledge-driven approach but unbiased *ab initio* approach with high-throughput exome sequencing followed by experimental confirmation. Our study will illustrate how large-scale sequencing, coupled with functional prediction and experimental validation, can identify disease mutations in single patients.

2874F

Exome Sequencing for Disease Gene Discovery in Jeune's Asphyxiating Thoracic Dystrophy. *P. Taylor[1], S. Wu[2], SF. Nelson[1,3], DH. Cohn[1,2], D. Krakow[1,2].* 1) Department of Human Genetics, University of California, Los Angeles, California, USA; 2) Department of Orthopaedic Surgery, University of California Los Angeles, California, USA; 3) Department of Pathology and Laboratory Medicine, University of California, Los Angeles, California, USA.

Jeune's Asphyxiating Thoracic Dystrophy (JATD) is a clinically heterogeneous, autosomal recessive skeletal dysplasia characterized primarily by an abnormally narrow thorax with short ribs and short limbs. It has an estimated incidence of ~1 out of every 126,000 live births. Most patients die in early infancy and childhood from respiratory insufficiency due to a small thorax. To date, four genes and one additional locus have been found to cause JATD in an autosomal recessive fashion. All four known genes are essential for normal primary cilia function. We employed the use of exome sequencing of seven unrelated patients affected with JATD to identify candidate mutations responsible for or contributing to the disorder. To maximize the success of gene discovery, we used Agilent's SureSelect Exome Pull Down capturing 50 Mb of the protein-coding exome. Each sample library was sequenced using a single lane of Illumina HiSeq. We generated 252Gb of paired-end sequence data with a mean coverage of 115X per sample. The sequencing data was aligned using Novoalign and variants were called using the Genome Analysis Toolkit. Variants were filtered against dbSNP135, 69 complete Genomics samples, 95 NIEHS EGP exome samples, 5400 exomes from the NHLBI Exome Sequencing Project, and 40 in-house exome samples. Only mutations that were present in the above controls in the homozygous state were removed from the candidate list. After filtering, potential disease-causing mutations representing two known genes and four novel genes were found in all seven individuals. In particular, one individual had two novel, predicted-damaging mutations in the gene *TRAF3IP1* in the compound heterozygous state. *TRAF3IP1* is a known cilia gene that has been found to play an essential role in ciliogenesis and embryonic development (Berbari et al. 2011. Dev Biol 360:66–76). We are currently characterizing these mutations in *Trypanosoma brucei*.

2875W

Seeking causative genes for human congenital general anosmia in multiply-affected Israeli families. *A. Alkelai[1], T. Olender[1], D. Oz-Levi[1], I. Keydar[1], R. Milgrom[1], E. Feldmesser[1], E. Ben-Asher[1], E.K Ruzzo[2], D.B. Goldstein[2], E. Pras[3,4], D. Lancet[1].* 1) Dept. of Molecular Genetics, The Weizmann Institute of Science, Rehovot, Israel; 2) Center for Human Genome Variation, Duke University School of Medicine, Durham, North Carolina 27708, USA; 3) The Danek Gertner Institute of Human Genetics, Sheba Medical Center, Ramat Gan, Israel; 4) The Sackler School of Medicine, Tel Aviv University, Tel Aviv, Israel.

General anosmia (the inability to sense odors) has negative effects on life quality, including impaired food and beverage enjoyment, social interactions and avoidance of poison and fire. An estimated 2.5–5% of the general population suffers from acquired general anosmia or hyposmia. In contrast, congenital general anosmia (CGA) affects merely <0.1%, and appears in isolated or syndromic forms. In specific cases, family members of patients with syndromic anosmia (Kallmann syndrome) have been found to suffer from the isolated form of CGA. While for an appreciable number of syndromic CGA types a causative gene has been identified, none of the isolated CGA instances have been genetically deciphered. Most of the reported isolated CGA cases are sporadic, and family-based studies are relatively rare, primarily because samples are difficult to ascertain. In our previous report (Feldmesser et al, Chem Senses 32:21–30, 2007) we have recruited 66 families of Jewish origin and performed whole-genome linkage analysis for selected families with multiple affected individuals with suspected founder variation. This study did not reveal significant linkage, probably due to small family size. We now report whole-exome next-generation sequencing in 9 individuals from three of these families. Each family was separately analyzed under the appropriate mode of inheritance (dominant, recessive and dominant with partial penetrance). We focused on family-shared variants present at minor allele frequency <0.05 in the 1000 Genomes Project, in 5000 exomes of the NHLBI database and in our own 50 exomes of Jewish individuals. These variants involved functionally reasonable genes, including such appearing in our recently constructed database of CGA candidate genes. Among these was neuropilin 2 (*NRP2*), a modulator of semaphoring 3A (*SAMA3A*) reportedly deleted in a form of Kallmann's syndrome. All implicated variants are now being further validated, examined for co-segregation in additional family members, and screened for low frequency in ethnically matched healthy controls. The identification of specific pathogenic functional CGA variants will help elucidate the molecular basis of general olfactory sensitivity.

2876T

Studying a locus for bilateral congenital perisylvian polymicrogyria: Linkage analysis, Cytoscan HD Array and Target Enrichment/Sequencing approaches. *F.R. Torres[1], S.S. Tsuneda[1], E.H. Yamamoto[1], R. Secolin[1], M.G. Borges[1], W.T.S.T. Ide[1], C.S. Rocha[1], F.M. Artiguenave[1], F. Cendes[2], M. Guerreiro[2], I. Lopes-Cendes[1].* 1) Department of Medical Genetics, UNICAMP, Campinas, SP, Brazil; 2) Department of Neurology, UNICAMP, Campinas, Brazil, SP, Brazil.

Bilateral congenital perisylvian polymicrogyria (BCPP) is the most common form of polymicrogyria. Familial recurrence was described for BCPP and different patterns of inheritance have been identified. Recently, chromosome Xq28 has been identified as a candidate locus for BCPP. This is a highly repetitive region, making it prone to non-homologous recombination events. We have identified a large family with an X-linked dominantly inherited BCPP. The aim of this study is to identify the deleterious event involved leading to BCPP in this family. To identify the candidate locus, linkage analysis was performed using 18 microsatellite DNA markers located in Xq27-q28 region. Subsequently, the candidate locus identified was captured with SureSelectXT technology (Agilent Technologies) and sequenced in a high-performance Hiseq Illumina sequencing machine (Illumina). In addition, in order to assess whether a submicroscopic structural rearrangement is involved in BCPP, three patients carrying the same haplotype were submitted to HD Cytoscan Array (Affymetrix). Linkage analysis identified a 13Mb candidate region on Xq27.2-q27.3. This locus is approximately 16Mb more centromeric than the first reported Xq28 locus. The Xq27.2-q27.3 region was captured and sequenced with 50x coverage. Approximately 3402 SNPs present in several genes were indentified. HD Cytoscan Array analysis did not reveal structural chromosome rearrangements in patients carrying the affected haplotype. We conclude that gross chromosomal abnormalities are not involved with BCCP etiology in this family, thus point mutations or small deletions are likely to be the event leading to this syndrome. Bioinformatics analysis of high performance sequencing data revealed two genes as potential candidates for BCPP etiology in the family studied. Putative roles related to development and/or cell division pathways make the two candidate genes identified relevant for BCPP.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2877F

Targeted Exome Sequencing for Molecular Diagnostics of Heterogeneous Genetic Disorders: Primary Ciliary Dyskinesia. *J.-L. Blouin[1,2], J. Bevillard[2], M. Guipponi[1,2], C. Gehrig[1], A. Vannier[2], F. Santoni[2], S.E. Antonarakis[1,2], F. Fellmann[3], S. Fokstuen[1], R. Lazor[4].* 1) Dept Genetic Medicine and Laboratory, Univ Hosps Geneva, Geneva, Switzerland; 2) Dept Genetic Medicine and Development, University of Geneva School of Medicine, Geneva, Switzerland; 3) Medical Genetics, CHUV, Lausanne, Switzerland; 4) Dept of Respiratory Medicine, CHUV, Lausanne, Switzerland.

Monogenic diseases with extensive genetic heterogeneity are laborious and expensive to molecularly diagnose by conventional sequencing or other indirect mutation screening due to the large number of genes. This had been partially overcome by microarray to either detect known mutations or resequencing a few genes. Today wide availability and decreasing cost of high throughput sequencing (HTS) as well as various DNA selection (DS) methods allow the analysis from dozens to thousands of genes even entire exome in a single experiment. Although this is now widely used in research, translating these techniques in molecular diagnostics is a challenge. We tested here a combination of targeted exome DS and HTS in patients with Primary Ciliary Dyskinesia (PCD) in order to develop a comprehensive molecular diagnostic method, and also a research tool to identify novel causative genes. PCD is a recessively inherited respiratory ciliopathy, caused sofar by 14 genes. Despite tremendous efforts in the past 15 years, genetic etiology remains unknown in more than 60% of cases. Liquid phase capturing was designed for all coding exons of known and candidate genes for PCD as well as other ciliopathies for a total of 164 genes including 3298 exons. Patient DNA libraries after DS are analyzed in a single lane (multiplexing indexed libraries) of a HT sequencer. Data were analyzed using a locally developed pipeline using BWA for alignment, SAMtools for variant calling, and ANNOVAR for annotation of all sequence changes, including functional prediction, allele frequency (dbSNP, 1000-genomes). We prioritized non-synonymous, frameshift, splicing, and indel variants according to novelty, quality score, and putative pathogenicity. We have provided sofar diagnostics in 2 of 5 patients showing compound heterozygous mutations, both in gene DNAH11, the most frequent and large PCD gene with 79 exons. Patient 1 had a nonsense (p.Arg3666*) and a frameshift mutation (p.Pro3606Hisfs*23), both confirmed by Sanger sequencing to be in trans. Patient 2 carried 2 deletions creating premature termination of the protein as p.[Arg761Serfs*10];[Val2957Trpfs*9]. Analysis of a number of additional patients is currently ongoing. Targeted exome sequencing of candidate genes presents advantage of lower analysis cost than entire exome and avoiding incidental findings. It holds considerable promises as next generation diagnostics of such highly heterogeneous monogenic disorders in clinical practice.

2878W

Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. *N. Matsumoto, Y. Tsurusaki, N. Miyake.* Department of Human Genetics, Yokohama City Univ Grad Sch of Med, Yokohama, Japan.

Coffin-Siris syndrome (CSS; MIM 135900) is a rare congenital anomaly syndrome characterized by growth deficiency, intellectual disability, microcephaly, coarse facial features and hypoplastic nail of the fifth finger and/or toe. The majority of patients are sporadic, being compatible with autosomal dominant inheritance with unknown genetic cause(s). By exome sequencing, we found de novo SMARCB1 mutations in two of five typical patients with Coffin-Siris syndrome, a rare autosomal dominant anomaly syndrome. SMARCB1 is a subunit of switching defective/sucrose nonfermenting (SWI/SNF) complexes. The SWI/SNF complexes play important roles in lineage specification, maintenance of stem cell pluripotency and tumorigenesis, known as one of chromatin-remodelers regulating the accessibility to target genes and gene expression by dynamic alterations of the chromatin structure. Since SMARCB1 is a subunit of SWI/SNF complexes, we screened 15 other subunits of SWI/SNF complexes in 22 Coffin-Siris syndrome patients. Nineteen patients (87%) showed a germline mutation in one of five SWI/SNF subunits: SMARCB1, SMARCA4, SMARCE1, ARID1A or ARID1B. The various types of mutations in different SWI/SNF components resulted in similar CSS phenotypes. This suggests that these complexes coordinately regulate chromatin structure and gene expression. This is the first report of germline mutations in SWI/SNF complexes associated with a multiple congenital anomaly syndrome, highlighting novel biological aspects of SWI/SNF complexes in humans. Aknowledgments: Drs. Okamoto N, Ohashi H, Kosho T, Imai Y, Hibi-Ko Y, Kaname T, Naritomi K, Kawame H, Wakui K, Fukushima Y, Homma T, Kato M, Hiraki Y, Yamagata T, Yano S, Mizuno S, Sakazume S, Ishii T, Nagai T, Shiina M, Ogata K, Ohta T, Niikawa N, Miyatake S, Okada I, Mizuguchi T, Doi H, and Saitsu H are highly appreciated for contributing to this work.

2879T

Identifying Mutations Causing Atypical Progeria by Exome Sequencing and CNV Analysis. *K.M. McSweeney[1], L.B. Gordon[2,3,4], J. Xu[1], W. Ping[1], M.F. Arlt[1], J. Li[1], T. Glover[1].* 1) Dept. Human Genetics, University of Michigan, Ann Arbor, MI; 2) The Progeria Research Foundation, Boston, MA; 3) Dept. of Pediatrics, Brown University, Providence, RI; 4) Dept. of Anesthesia, Harvard Medical School, Boston, MA.

Progeria is a rare genetic disease characterized by a rapid aging phenotype. The classically defined form of progeria is Hutchinson-Gilford Progeria Syndrome (HGPS), an autosomal dominant form of progeria characterized by growth delay, alopecia, loss of subcutaneous fat, stiff joints, and vascular disease, but normal intelligence. HGPS is caused by a point mutation (c. 1824 C>T; G608G) that creates a cryptic splice site 150nt upstream of the 3' end of exon 11. This results in a 50 amino acid deletion, including an endoproteolytic cleavage site for the ZYMPSTE24 protease. The resulting disease-causing protein is an immature form of the lamin A protein called “progerin”. Progerin remains localized at the nuclear lamina resulting in nuclear blebbing and nuclear envelope instability, which is believed to be one cause of the premature aging phenotype. Other progeroid diseases are caused by mutations involving lamin A or its processing factors, or mutations in DNA damage response genes such as the WRN helicase. However, the genes responsible for many cases of progeroid diseases are unknown. We performed whole exon sequencing on nine patients with “atypical” progeria who show some phenotypes similar to classical HGPS but lack mutations in *LMNA* and *ZPMTSE24*. SNP array analysis for CNV and homozygosity detection was also performed on all nine patients and these data were considered together with the exome sequences. We obtained an average of 174 million sequencing reads and a range of 150- to 237-fold genome coverage for each patient sample. The data were analyzed and filtered using a number of state of the art algorithms and revealed 5–10 deleterious mutations in each patient in candidate genes. These genes were chosen for their involvement in nuclear membrane structure/function, DNA damage response, lipid metabolism, and other functions related to the progeria phenotype. Currently, parental DNA is being analyzed to determine the mode of inheritance of each mutation in each unrelated patient. Identification of causative genes will not only be of direct importance to affected patients and their families, but may have great impact on the understanding of the aging process and its associated disorders.

2880F

Diagnostic Sequencing Applications Using the Ion Torrent. *D. Muzny[1], X. Wang[1], C. Buhay[1], M. Wang[1], H. Dinh[1], J. Reid[1], D. Wheeler[1], L. Lotta[2], E. Boerwinkle[1,3], R. Gibbs[1].* 1) Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX; 2) 2Angelo Bianchi Bonomi Hemophilia and Thrombosis Center, Milan, Italy; 3) 3University of Texas Health Science Center at Houston, School of Public Health, Houston, TX 77030.

Since its introduction to the Human Genome Sequencing Center (HGSC) in January 2011, the Ion Torrent PGM platform has demonstrated a tremendous capacity growth, as well as great flexibility. The platform has been ideal for amplicon experiments, including AmpliSeq designs. These have been largely focused on validation of putative variants discovered on other platforms. The amplicon sequencing to characterize more than 5,000 variant sites from different cancer and genetic discovery projects has been straightforward and has only been limited by the capacity to produce the multiple PCR products. To extend the utility of the device we have now performed targeted gene panel capture sequencing on the PGM platform. This required modifications to the library construction process to improve complexity, incorporate multiplexing and allow for the use of sub-microgram DNA amounts in library and hybridization. Capture variant discovery has been demonstrated using a custom design (~1Mb), targeting exons from all 167 known retinal disease genes. The procedure was calibrated using a well characterized (hapmap) sample, yielding 99.59% (1181/1185) overall accuracy, with a total of just four erroneous SNP calls identified. PGM target sequencing was then applied to diagnosis of two retinitis pigmentosa (RP) patient families. Disease causing mutations in Ush2A have been identified and validated in both RP families representing the first reported human patient cases diagnosed on the PGM. In addition, eight regional capture designs as well as exome sequencing have now been evaluated using the PGM pipeline ranging in size from 0.1 to 43Mb in target regions. Thrombosis and X Chromosome capture designs as well as whole exome sequencing have resulted in the identification of putative and causative variants. These experiments have been supported by advances in a highly configurable variant caller specifically for Ion Torrent validation applications. The platform is now demonstrated to be suitable for a clinical setting where accuracy and rapid cycle time are essential.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2881W

Next Generation Sequencing Reveals an Allelic Variant of GLI2 in a Turkish Pedigree with a unique clinical phenotype. *D. Pehlivan[1], A. Yesilyurt[2], C.G. Gonzaga-Jauregui[1], E. Karaca[1], Y. Pehlivan[3], S. Guran[4], R.A. Gibbs[1,5], J.R. Lupski[1,6,7].* 1) Molecular and Human Genetics, Baylor College of Medicine, Houston, TX; 2) Department of Medical Genetics, Dr. Zekai Tahir Burak Women Health Training and Research Hospital, Ankara, Turkey; 3) Department of Internal Medicine, University of Gaziantep, School of Medicine, Gaziantep, Turkey; 4) Department of Medical Genetics, Gulhane Military Medical Academy, Ankara, Turkey; 5) Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX, USA; 6) Department of Pediatrics, Baylor College of Medicine, Houston, TX; 7) Texas Children's Hospital, Houston, TX, USA.

The GLI2 gene was discovered and so-named for its amplification in gliomas of the brain. Beside its overexpression in tumoral tissues, simple mutations (small deletions, missense mutations) in this gene were reported in association with holoprosencephaly sequence. GLI2 is one of three vertebrate transcription factors implicated as obligatory mediators of Sonic hedgehog (SHH) signal transduction. Diminished SHH signaling is associated with the most common forebrain defect in humans, holoprosencephaly, which includes cyclopia, a phenotype also seen in mice and other vertebrates with defective Shh signal transduction. We report a homozygous missense mutation in GLI2 in a consanguineous family with a unique phenotype. A 22 year-old male patient, born from first degree cousin marriage, was referred to medical genetics due to hypernasal speech. Physical examination revealed additional dysmorphic facial features including triangular and elongated face, frontal upsweep, narrow forehead, bitemporal narrowing, thick and medially-flared eyebrows, prominent tubular nose, prominent columella, thin upper lips, retromicrognathia, partial cutaneous syndactyly on both feet, and an inability to move the tongue. To the best of our knowledge, this phenotype is unique and there is no similar reported case in the published literature. Family history showed two additional affected sisters with similar findings and two healthy siblings. Homozygosity mapping and whole-exome sequencing of DNA from two affected individuals revealed the same candidate genes. Segregation analysis showed that only the GLI2 variant is segregating in the family. The mutation is confirmed with Sanger sequencing. This result indicates that exome sequencing can be important to elucidating the molecular etiology of extremely rare, novel clinical phenotypes and to explore allelic variants of known genes.

2882T

Rapid and efficient mutation detection in the hundreds of target genes by bench-top next generation sequencer with custom target capture method. *A. Shimizu[1], C. Torii[2], N. Suzuki[3], H. Mutai[3], J. Kudoh[4], R. Kosaki[5], T. Matsunaga[3], K. Kosaki[2].* 1) Department of Molecular Biology, Keio University School of Medicine, Tokyo, Japan; 2) Center for Medical Genetics, Keio University School of Medicine, Tokyo, Japan; 3) Laboratory of Auditory Disorders, National Institute of Sensory Organs, National Tokyo Medical Center, Tokyo, Japan; 4) Laboratory of Gene Medicine, Keio University School of Medicine, Tokyo, Japan; 5) Division of Medical Genetics, National Center for Child Health and Development, Tokyo, Japan.

In combination with massively parallel sequencing using next-generation sequencers and custom sequence capture methods for restricted target genomic regions is possible and accessible to identify the mutation from hundreds of known disease genes at a lower cost per sample. We focused on studies of patients with hearing loss and some other congenital disorders. Our regions of the interest genes spanning 860 kb and including 5,477 coding exons of 309 genes. With eArray, online custom design tool (Agilent Technologies, Inc.), we have designed and manufactured more than 29,000 capture probes that targeted the genomic region of interest. We failed to design probes for only 31 kb (3.6%) of target coding regions. The fragmented genomic DNA samples from patients and their family members were captured by our custom probes and sequenced using MiSeq sequencer (Illumina, Inc.) that generated 150 bp (six samples in one lane) paired-end reads. The sequencer generated more than 2 million reads with 600 Mb sequences per sample. These sequences were aligned to the human genome reference sequence using BWA (version 0.6.1). Over 99.7% of the targeted regions were sequenced, 75.6% of sequence reads were on the target region and the average depth of target coverage was x167, with 96% of bases covered by ≥10 reads. Genome analysis toolkit package was used to perform local realignment, map quality score recalibration and SNPs/indels calls for each individual. We removed potential PCR duplicates by using Picard. We identified and labeled the identified SNPs and indels within the targeted coding regions and that overlapped with known SNPs that were reported in the dbSNP version 135, the 1000 Genomes Project and our in-house Japanese SNPs dataset. We annotated the novel variants and known variants with clinical significance using snpEff, ANNOVAR and a custom analysis pipeline. We also used commercial available software packages Avadis NGS (Agilent Technologies, Inc.), CLC Genomics Workbench (CLC bio) and NextGENe (Next-Gen, Inc) for comparing to the detected mutations. We have identified the potentially deleterious mutations in half of the patients, and validation of the identified mutations is currently underway using Sanger sequencing. Our study indicated that the combination of bench-top type next-generation sequencer and target sequence capture methods provide cost effective, accurate and rapid diagnosis of congenital disorders.

2883F

Identification of candidate genes associated with autosomal dominant pulmonary fibrosis by whole-exome sequencing. *W. Wiszniewski[1], C.G. Gonzaga-Jauregui[1], T.J. Vece[2], D.M. Muzny[3], M.N. Bainbridge[3], J.G. Reid[3], R.A. Gibbs[3], J.R. Lupski[1], Baylor-Hopkins Center for Mendelian Genomics.* 1) Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX; 2) Department of Pediatrics, Baylor College of Medicine, Houston, TX; 3) Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX.

Familial pulmonary fibrosis (OMIM#178500) is a rare and progressive lung disease that presents initially with shortness of breath and later progresses to end stage lung disease. The histopathological analysis of lung tissue shows varying degrees of inflammatory and fibrotic changes. Most cases are sporadic, however it has been estimated that familial form is responsible for 0.5–2% of cases of idiopathic pulmonary fibrosis. Susceptibility to pulmonary fibrosis has been associated with sequence variants in 6 genes including TERC, TERT, SFTPA1, SFTPA2 and MUC5B. We have identified two large families with familial pulmonary fibrosis and autosomal dominant mode of inheritance. In the first family, the disease was diagnosed in 8 cases in 4 generations, in the second family in 5 cases in 2 generations. We have selected two affected individuals from each family for whole-exome sequencing studies. We have identified deleterious variants in candidate genes that are potentially associated with the observed phenotype.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2884W

Family matters: Exome sequencing can identify causal variants in isolated probands through family-based studies. *H. YU[1], C.R Coughlin[1], E.A. Geiger[1], M. Friederich[1], L. Medne[2], J.E. Ming[2], E.H. Zackai[2], J.L. Van Hove[1,3], X. Gai[4], G.H. Scharer[1,3], T.H. Shaikh[1,3].* 1) Section of Clinical Genetics and Metabolism, Department of Pediatrics, University of Colorado School of Medicine, Aurora, CO; 2) Division of Human Genetics, The Children's Hospital of Philadelphia, Philadelphia, PA; 3) Intellectual and Developmental Disabilities Research Center, University of Colorado School of Medicine, Aurora, CO; 4) Department of Molecular Pharmacology and Therapeutics and Center for Biomedical Informatics, Loyola University Chicago Stritch School of Medicine, Maywood, IL.

Exome sequencing has proven to be a powerful tool for gene discovery especially when applied to multiple related or unrelated individuals with a similar phenotype. This paradigm limits the application of exome sequencing, as the vast majority of individuals with multiple anomalies do not have a specific genetic diagnosis or present with phenotypes that are unique to the patient. Our work demonstrates successful strategies for novel gene discovery as well as establishment of causality in single probands with a variable spectrum of multiple congenital anomalies or inborn errors of energy metabolism. Exome sequencing was performed in six probands and their unaffected parents to obtain >50X average coverage per sample. Following standard bioinformatics for variant calling and annotation, candidate genes were highlighted using parental data and testing all possible modes of inheritance. This resulted in <5 candidate genes per proband, which were further examined in the context of the patient's phenotype. In Patient 1, who had blepharophimosis ptosis and epicanthus inversus syndrome (BPES) plus developmental delay, cryptorchidism and digital abnormalities, we identified a de novo mutation in *KAT6B*, a histone acetyltransferase gene. *KAT6B* has recently been implicated in related but phenotypically variable genetic disorders. In Patient 2, who had progressive neurodegeneration, intractable seizures, myoclonus, extrapyramidal movement and biochemical evidence for a mitochondrial disorder, we identified compound heterozygous, missense mutations in a very strong candidate gene. This candidate, a nuclear gene required in mitochondrial function, has not been previously associated with mitochondrial disease. Single, very strong candidate genes were identified in Patients 3–5. The candidate gene in Patient 3 may be a novel gene in the methylmalonic aciduria and homocystinuria pathway, as the patient had biochemical findings associated with Cobalamin C deficiency. In Patient 6 we had more than one candidate gene. For novel genes identified, proof of causality was pursued using animal models or transfection studies to show complementation. We demonstrate that causal gene mutations can be identified in a single proband with the use of appropriate family-based studies followed by confirmatory functional analyses to establish causality.

2885T

Whole Exomic Triangulation and Systems Biology Identify Novel Pathogenic Substrates for Autosomal Dominant Long QT Syndrome. *N.J. Boczek[1], DJ. Tester[2], JR. Giudicessi[1, 3], TJ. Kamp[4], MJ. Ackerman[2,5,6].* 1) Mayo Graduate School, Mayo Clinic, Rochester, MN; 2) Department Molecular Pharmacology and Experimental Therapeutics, Windland Smith Rice Sudden Death Genomics Laboratory, Mayo Clinic, Rochester, MN; 3) Mayo Medical School, Mayo Clinic, Rochester, MN; 4) Deparment of Medicine, Cardiovascular Section, and the Cellular and Molecular Arrhythmias Research Program and Inherited Arrhythmias Clinic, University of Wisconsin, Madison, WI; 5) Department of Medicine, Division of Cardiovascular Diseases, Mayo Clinic, Rochester, MN; 6) Department of Pediatrics, Division of Pediatric Cardiology, Mayo Clinic, Rochester, MN.

Long QT syndrome (LQTS) is the most common cardiac channelopathy with 13 elucidated LQTS-susceptibility genes. An estimated 20% of LQTS cases remain genetically elusive. We combined whole exome sequencing (WES), genomic triangulation, and bioinformatic/systems biology approaches to identify the pathogenic substrate responsible for non-syndromic, LQT1–13 genotype negative, autosomal dominant LQTS in a multigenerational pedigree and established the spectrum and prevalence of the elucidated gene among a large cohort of unrelated patients with genotype negative/phenotype positive LQTS. WES was completed on the symptomatic index case (QTc = 498ms), unaffected father (QTc = 383ms), and affected maternal aunt (QTc = 479ms). Genomic triangulation (removal of variants shared between the index case and unaffected father, and unshared variants between the index case and affected maternal aunt) in combination with bioinformatic tools including the published LQTS interactome and 3 disease network analysis gene prioritization ranking algorithms (ToppGene, Endeavor, and SUSPECTS), were used to identify potential disease-causing candidate variants. The index case's exome contained 46,307 variants. Of the 698 rare, nonsynonymous, single nucleotide variants and insertions/deletions, 65 variants remained after genomic triangulation. Only 6 of these variants resided within the published 1629 node LQTS interactome and P857R-CACNA1C was ranked number 1 by all 3 disease network algorithms. P857R-CACNA1C co-segregated with the disease in all affected individuals of the multigenerational pedigree making it the most probable pathogenic mutation. The *CACNA1C* encoded L-type calcium channel (Cav1.2) α-subunit missense mutation involves a highly conserved residue localizing to a critical PEST domain in the II-III linker. Subsequent mutational analysis identified 4 additional mutations within *CACNA1C* in ~4% of our larger LQTS genotype negative/phenotype positive cohort. Two of these mutations also involved highly conserved residues within Cav1.2's PEST domain. This study provides evidence that coupling WES and bioinformatic/systems biology is an effective strategy for the identification of potential disease causing mutations. The identification of a *CACNA1C* mutation co-segregating with disease in a single pedigree and multiple mutations residing in the PEST domain suggests that *CACNA1C* perturbations are *bona fide* pathogenic substrates for classical, autosomal dominant LQTS.

2886F

X-exome sequencing identifies the causal variant in a large pedigree with X-linked intellectual disability, truncal obesity, gynecomastia, hypogonadism and unusual face. *J.K.P. van Amstel[1], M. Harakalova[1], M.J. van den Boogaard[1], R.J. Sinke[2], S. van Lieshout[1], M. van Tuil[1], K. Duran[1], I. Renkens[1], P.A. Terhal[1], C.G.F. de Kovel[1], I.J. Nijman[1], M.M. van Haelst[1], N.V.A.M. Knoers[1], G. van Haaften[1], W.P. Kloosterman[1], R.C.M. Hennekam[3], E. Cuppen[1].* 1) Dept of Medical Genetics, Univ Medical Ctr (UMCU), Utrecht, The Netherlands; 2) Dept of Genetics, Univ Medical Ctr (UMCG), Groningen, The Netherlands; 3) Dept of Pediatrics, Acad Medical Ctr (AMC), Amsterdam, The Netherlands.

We characterized a large Dutch family with 7 males affected by a rare syndrome of X-linked intellectual disability (XLID), hypogonadism, gynaecomastia, truncal obesity, short stature, small head, short ears and recognizable craniofacial deformities. Eight females in this pedigree showed a much milder expression of the phenotype, comprising learning disorder and recognizable facial features. We performed X chromosome exome (X-exome) sequencing in five individuals from this family and identified a novel intronic variant in a histone deacetylation family member. The gene is located in the region for which linkage was established in this family with a maximum LOD-score of 4.93. The variant was shown to affect normal splicing, resulting in exon skipping and introduction of a premature stop. The mutation completely co-segregates with the phenotype in this family and is absent in Dutch controls or available exome databases. Interestingly, affected female carriers show a markedly skewed X-inactivation pattern in blood where all cells show the mutated chromosome X completely inactive. The mutated gene is important for epigenetic control of developmental processes and a knock-out model supports the facial phenotype present in our patients. In summary, we provide genetic evidence for the involvement of the histone deacetylation pathway in a syndromic form of XLID.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2887W

Completing the BBS jigsaw by exome sequencing and phenotype-genotype correlations: LZTFL1 a BBSome related gene associated with developmental features is BBS17. *H. Dollfus[1,4], F. Stutzmann[1], M. Gerard[2], A. Verloes[3], E. Schaefer[1,4], C. Stoetzel[1], V. Marion[1].* 1) Laboratoire de génétique médicale, AVENIR INSERM- Université de Strasbourg, Strasbourg, France; 2) CHU de Caen servicede génétique de Caen, France; 3) Unité de génétique clinique, Hôpital Robert Debré, Paris, France; 4) CARGO, Centre de référence génétqiue ophtalmologqiue, Hôpitaux Universitaires de Strasbourg, France.

Introduction: Bardet-Biedl syndrome is an emblematic recessive and genetically heterogeneous ciliopathy characterized by polydactyly, retinitis pigmentosa, obesity, cognitive impairment and kidney dysfunction. To date, 20% of the BBS patients have no mutation identified in the 16 known BBS genes. Exome sequencing has proven to be a valuable approach to identify new ciliopathy genes. Material and methods: Single nucleotide polymorphism analysis followed by exome sequencing was performed on a consanguineous family diagnosed with Bardet-Bield syndrome with additional unusual developmental features namely insertional polydactyly and situs inversus. Results: A homozygous 5-bp deletion was identified in LZTFL1 encoding a protein known to be related to the BBSome as an important regulator of BBSome ciliary trafficking and Shh signaling. No transcript was found in the fibroblast cells of one affected individual. Sonic Hedgehog (SHH) signalling cascade analysis on the patient's fibroblasts showed significant increase for Smo. Patched and Gli2 indicated substantial activation of Shh signalling. Conclusions: We show that mutation in LZTLF1 (or BBS17), leads to a BBS phenotype with an enhanced developmental phenotype with Shh signaling abnormalities.

2888T

Simultaneous identification of single nucleotide mutations and DNA copy number aberrations in neurodevelopmental disorders by exome sequencing. *AM. Sulonen[1, 2], H. Almusa[1], H. Edgren[1], P. Ellonen[1], L. Forsstrom[3], T. Linnankivi[4], S. Knuutila[3], J. Saarela[1].* 1) Institute for Molecular Medicine Finland, FIMM, 00290 Helsinki, Finland; 2) National Institute for Health and Welfare, 00290 Helsinki, Finland; 3) Department of Pathology, Haartman Institute and HUSLAB, University of Helsinki and Helsinki University Central Hospital, 00290 Helsinki, Finland; 4) Department of Child Neurology, Children's Hospital, University of Helsinki and Helsinki University Central Hospital, 00290 Helsinki, Finland.

Neurodevelopmental disorders include well known syndromes, but also a wide variety of yet uncharacterized syndrome-like phenotypes. If acquired etiology cannot be indicated, genetic causes are considered. However, genetic background is multiform; *de novo*-mutations and familial variants are typical, which makes genetic testing complicated with conventional, single mutation or gene targeted methods. We show here, how next-generation exome sequencing can be used to identify both single nucleotide variants (SNV) and copy number variations (CNV), thereby providing a comprehensive genetic profile of patients with neurodevelopmental disorders for the use of clinical diagnostics.

Three families with multiple affected children were selected for the study. Array comparative genomic hybridization (aCGH) was previously performed for the families: Family 1 shared a heterozygous deletion of *RTN4RL1* in the patients, which has not been previously associated with developmental disorders. Patients of family 2 had a heterozygous 5.3 Mb deletion in 9q22, which was also carried by their cognitively normal father. Family 3 had no anomalous findings in aCGH. Two patients from each family and two healthy individuals were captured with NimbleGen SeqCap v3.0 exome capture kit and sequenced on Illumina HiSeq2000. SNVs were called with an in-house built variant identification and annotation pipeline VCP. For CNV identification we developed a method, Exo-CNV, which is based on copy number ratios from test vs. reference samples, derived from RPKM-normalized exon coverages. Mean exon coverages of other than familial samples were used as reference.

We identified 67, 111 and no novel, non-synonymous SNVs and 48, 64 and 36 CNVs shared in the patients of families 1, 2 and 3, respectively. When SNPs seen in the 1000Genomes were also included, six non-synonymous shared variants were identified in the family 3. CNVs identified with aCGH were verified with Exo-CNV. Detailed analyses of the genes affected by identified mutations and CNVs, and genotyping of potential disease causing variants in all family members are ongoing.We developed a pipeline that efficiently and comprehensively identifies family specific mutations possibly underlying neurodevelopmental disorders. Combination of SNV and CNV identification further enlightens the complexity of familial disorders.

2889F

***USP9X* mutation in a child with developmental delay and multiple congenital anomalies.** *M.S. Brett[1], A.H.M. Lai[2], E.C.P Lim[1], J.R. McPherson[3], B.W.M Cham[2], Z.J. Zang[4], P. Tan[4], S. Rozen[3], E.C. Tan[1].* 1) KK Research Centre, KK Women's & Children's Hospital, Singapore; 2) Department of Pediatrics, KK Women's & Children's Hospital, Singapore; 3) Neuroscience and Behavioral Disorders, Duke-NUS Graduate Medical School, Singapore; 4) Cancer and Stem Cell Biology Program, Duke-NUS Graduate Medical School, Singapore.

We describe a patient with intellectual disability and developmental delay (ID/DD) and multiple congenital anomalies (MCA) including facial dysmorphisms, Dandy-Walker malformation, ventricular septal defect, anal stenosis and bilateral post-axial polydactyly. She is the first child of healthy, unrelated parents of Chinese ancestry with no significant family history. Her karyotype was normal and array CGH with an Agilent 400K array did not detect any clinically significant aberrations. Massively parallel sequencing with a selected panel of 355 genes that have been associated with ID/DD and MCA was carried out on the child. Targeted exon capture was performed using the Agilent SureSelect system and sequenced on an Illumina HiSeq. The mean sequencing depth was 798 and 85% of the targeted regions were covered >10X. After filtering out the variants present in dbSNP and the 1000 Genome project there were a total of 107 single nucleotide variants (SNVs) in the targeted genes. The SNVs were further filtered by severity of predicted effect, gene function, evolutionary conservation, presence in locus-specific mutation databases and correlation with patient phenotype. SNVs were selected for validation by Sanger sequencing followed by testing of parental DNA. Sequencing identified a heterozygous *de novo* canonical splice site mutation c.1986-1G>T in the *USP9X* (Ubiquitin specific peptidase 9, X-linked) gene at chrX: 41025124 (hg19) which is predicted to lead to the deletion of exon 15 in the RNA splicing process. The gene codes for a deubiquitylating enzyme and has an important regulatory role at the level of protein turnover. It has been shown to be an essential component of the TGFβ/BMP signalling cascade and specifically deubiquinates SMAD4. A truncating mutation in *USP9X* has been identified in a family with XLMR but no other reports have associated *USP9X* mutations with ID/DD and MCA. In conclusion, targeted gene sequencing has identified a canonical splice site mutation in the *USP9X* gene that could be causative of some or all of the phenotypes in this patient. Identification of further patients with mutations in the *USP9X* gene will be needed to confirm the association with ID/DD and MCA.

2890W

Congenital cataracts, hypomyelination, progressive microcephaly, digital anomalies: a new recessive syndrome due to mutations in a nucleoporin gene? *E. Geraghty[1], G. Carvill[1], I. Glass[1,2], R. Kapur[3], D. Doherty[1,2], H. Mefford[1,2].* 1) Dept. of Pediatrics, Div. of Genetic Medicine, University of Washington, Seattle, WA; 2) Seattle Children's Hospital, Seattle, WA; 3) Dept. of Laboratories, Seattle Children's Hospital, Seattle, WA.

We report two affected siblings with congenital cataracts, congenital heart defects, prenatal onset ventriculomegaly, progressive microcephaly, severely delayed myelination, thin corpus callosum, hypotonia, and hand and foot dysmorphia. The first girl presented with hypotonia, congenital cataracts, unilateral preaxial polydactyly and gracile fingers. Echocardiogram showed bicuspid aortic valve. She developed central hypoventilation and died of respiratory failure at 5 months. She had extensive genetic testing with normal blood and skin karyotype, Affymetrix 6.0 SNP array, plasma amino acids, urine organic acids, acylcarnitine profile, very long chain fatty acids, 7-dehydroxycholesterol, *BCOR* sequencing, chromosome breakage studies, transferrin isoelectric focusing and TORCH titers. Her sister was identically affected except that she had partial anomalous pulmonary venous return (PAPVR) and a normal aortic valve and did not have polydactyly. She died at five weeks of age. Autopsy in the second sibling revealed gliosis of cerebrum, cerebellum, brainstem and spinal cord, as well as streak ovaries and polysplenia. Given the similar findings in both girls, we suspected a new autosomal recessive (AR) syndrome. We performed exome sequencing in the first sibling and filtered the data to identify nonsynonymous variants present in <1% of 1537 subjects from the Exome Sequencing Project (ESP; http://snp.gs.washington.edu/EVS/). Filtering for genes with rare homozygous or compound heterozygous variants identified 19 candidate genes. We excluded 17 candidates because variants were not confirmed by Sanger sequencing or segregation was not consistent with AR disease. *CCDC99*has a homozygous missense change in both girls that is heterozygous in 0.5–1% of controls. *NUP188* harbors a 4-bp deletion causing a frameshift and premature stop on the maternal allele and a nonsense mutation on the paternal allele. Neither is seen in controls, and no truncating or splice site variants in *NUP188* are present in 5400 ESP exomes. NUP188 is a component of the highly conserved Nup93 nucleoporin subcomplex and may help restrict the passage of transmembrane proteins into the inner nuclear envelope. Left-right patterning defects caused by*NUP188* knockdown in *X. tropicalis* are intriguing, since one sibling had PAPVR and polysplenia. Future work will focus on identifying additional patients and evaluating patient fibroblasts for defects in *NUP188* expression, localization and function.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2891T

Trio-exome sequencing identifies mutations of the gene encoding the histone acetyltransferase KAT6B/MYST4 in individuals with the Young-Simpson syndrome. *K. Kurosawa[1], K. Enomoto[1], T. Kondoh[2], S. Mizuno[3], M. Adachi[4], K. Muroya[4], Y. Yamanouchi[6], T. Nishikawa[1], N. Furuya[1], M. Tominaga[1], T. Naruto[1], K. Ida[1], Y. Kuroda[1], T. Sengstag[5], R. Manabe[5], M. Masuno[6].* 1) Division of Medical Genetics, Kanagawa Children's Med Ctr, Yokohama, Japan; 2) Division of Developmental Disabilities, The Misakae-nosono Mutsumi Developmental, Medical and Welfare Center, Isahaya, Japan; 3) Department of Pediatrics, Central Hospital, Aichi Human Service Center, Kasugai, Aichi, Japan; 4) Division of Endocrinology and Metabolism, Kanagawa Children's Med Ctr, Yokohama, Japan; 5) Omics Science Center, RIKEN Yokohama Institute, Yokohama, Japan; 6) Genetic Counseling Program, Graduate School of Health and Welfare, Kawasaki University of Medical Welfare, Kurashiki, Japan.

Young-Simpson syndrome (YSS), also known as Say-Barber-Biesecker variant of Ohdo syndrome or Say-Barber-Biesecker-Young-Simpson syndrome (SBBYSS), is a pleiotropic disorder characterized by skeletal, endocrine, and neurological abnormalities. To identify the molecular basis of YSS, we performed whole-exome sequencing of one proband-parent trio and one independent case. Two different protein-truncating mutations in the highly conserved histone acetyltransferase gene KAT6B/MYST4 were found to be heterozygous in the cases, respectively. The mutations were confirmed by Sanger sequencing. Sequencing of the KAT6B/MYST4 in 4 additional affected individuals identified disease-causing mutations in all probands, including two recurrent de novo alterations. 5 out of the 6 individuals had the protein-truncating mutations, but one patient, who represented mild phenotype, had a mutation of three-bases deletion resulting in a truncated product for only one amino acid residue in the glutamic acid-rich region. Our results confirmed that trio-based exome sequencing is a powerful approach to discover genes causing sporadic developmental disorders. These findings indicated that histone acetylation has a crucial role in early human development.

2892F

Exome sequencing identifies a missense mutation of ANKRD26 in autosomal dominant Thrombocytopenia. *C. Kim[1], P.M.A. Sleiman[1,2,3], L. Tien[1], S. Panossian[1], K. Xu[1], B. Tweddale[1], N. Abdel-Magid[1], E. Frackelton[1], Y. Housawi[4], H. Hakonarson[1,2,3].* 1) Ctr Applied Genomics, CHOP, Philadelphia, PA; 2) Dept Pediatrics, Perelman School of Medicine, Univ of Pennsylvania, Philadelphia PA; 3) Division of Human Genetics, Children's Hosp Philadelphia, Philadelphia, PA; 4) King Fahad Specialist Hospital-Dammam, Saudi Arabia.

Thrombocytopenia-2 is an autosomal dominant nonsyndromic disorder characterized by decreased numbers of normal platelets, resulting in a mild bleeding tendency. Laboratory studies show no defects in platelet function or morphology, and bone marrow examination shows normal numbers of megakaryocytes and normal maturation stages, suggesting defective platelet production or release. Recently, Pippucci et al. (2011) identified 6 different heterozygous mutations in the 5-prime promoter region of the ANKRD26 gene in 9 of 20 unrelated families with THC2 subsequently expanding that to 21 families (Noris et al, 2011). All of the mutations occurred in a highly conserved 19-nucleotide stretch. In vitro functional expression assays in Dami human megakaryoblastic cells indicated that the mutations resulted in increased expression, particularly when the cells were stimulated toward maturation. The findings suggested a gain-of-function effect. We carried out exome sequencing in two affected members, one affected parent and an affected child, of a family with presumed dominant thrombocytopenia that had screened negative for all known genetic mutations. Coding sequences were enriched using the Agilent SureSelect 50Mb kit and the resultant libraries sequenced on an Illumina HiSeq 2000 to a mean depth of ~100X. Following variant annotation and filtration with ANNOVAR, 42 heterozygous variants that were shared between the two affected individuals and not present in a large number of control chromosomes remained. The 42 variants included a D158G missense mutation in the ANKRD26 gene. The mutation was confirmed by sanger sequencing and shown to segregate with the phenotype in the family. Gain of function mutations in the 5'-UTR of ANKRD26 have recently been reported to be causal of THC2, however, no missense mutations have yet been reported in this gene. Functional characterization of the D158G missense mutation is ongoing to determine if it confers a gain of function effect on the ANKRD26 protein. In combination with the 5' UTR mutations described by Pippucci et al. (2011) and Noris et al (2011) our finding that missense mutations within the open reading frame of ANKRD26 can result in thrombocytopenia could place ANKRD26 as a significant cause of thrombocytopenia.

2893W

Holoprosencephaly (HPE) : Identification of new candidate genes by high throughput sequencing of family trios. *V. David[1,2], C. Dubourg[1,2], M. De Tayrac[1], V. Dupé[1], M. Aubry[5], L. Ratié[1], S. Mercier[1,3], J. Mosser[1,4], S. Odent[1,3].* 1) Faculté de Médecine, UMR 6290 CNRS/IGDR,Université Rennes1, Rennes, France; 2) Génétique Moléculaire, CHU Pontchaillou, Rennes, France; 3) Génétique Clinique, Hopital sud, Rennes, France; 4) Génomique médicale, CHU Pontchaillou, Rennes, France; 5) Plateforme Biogenouest, Université Rennes1, Rennes, France.

Holoprosencephaly (HPE), the most common developmental defect (1/250 conceptuses, 1/10000 newborns), is caused by a failure to form the midline of the forebrain and midface. The etiology of HPE has pointed to a strong genetic component complicated by locus heterogeneity. Genetic anomalies in the four main genes (SHH, ZIC2, SIX3, TGIF) were identified in only 30% in the largest European series of 645 HPE probands. Clinical feature is extremely variable and many mutation carriers are unaffected. Except for ZIC2 mutations, genetic anomalies are inherited in 70% of the cases from one of the parents,. These latter are asymptomatic or carrying a microform of HPE, suggesting the implication of a second anomaly arising de novo or inherited. Our objective is thus to identify new genes involved in HPE. Recent findings and literature reports showed that exome sequencing of family trios could lead to the identification of new mutations in Mendelian diseases We thus have selected 10 well characterized family trios where one of the parents (asymptomatic or with a microform) was carrying a heterozygous mutation in one of the major HPE genes SHH, SIX3 or TGIF, transmitted to the offspring with a very severe HPE form (alobar or semilobar HPE). We postulated that this phenotype was due to a second de novo or inherited event. Each member of the family trios has been studied for the four main HPE genes and for copy number variations by CGH array analysis (Agilent, 180K) and only one of the parents carried a mutation in SHH, SIX3 or TGIF. Exome sequencing of these trios are pending and the results will be presented. They should seek a second mutation in new genes in these severe HPE cases. This will lead to a better knowledge of the physiopathology and improve genetic counseling of the disease.

2894T

Exome sequencing reveals a novel de novo GNAS gene mutation in a boy with usual presentation of pseudohypoparathroidism. *M. J. Dasouki[1], S. E. Soden[2,3,5], E. G. Farrow[2,3,4,5], C. J. Saunders[2,3,4,5], D. L. Dinwiddie[2,3,4,5], N. A. Miller[2,3,5], S. F. Kingsmore[2,3,4,5].* 1) Departments of Pedaitrics & Internal Medicine, University of kansas Medical Center, Kansas City, KS, 66160, USA; 2) Center for Pediatric Genomic Medicine, Children's Mercy Hospital, Kansas City, Missouri 64108, USA; 3) Department of Pathology, Children's Mercy Hospital, Kansas City, Missouri 64108, USA; 4) Department of Pediatrics, Center for Pediatric Genomic Medicine, and Department of Pathology, Children's Mercy Hospital, Kansas City, Missouri 64108, USA; 5) School of Medicine, University of Missouri-Kansas City, Kansas City, Missouri 64108, USA.

The bi-parentally imprinted GNAS complex locus (GNAS) is known to be associated with multiple endocrine, skeletal and neurologic abnormalities including pseudohypoparathyroidism (OMIM#103580, #603233, #612462), progressive osseous heteroplasia (OMIM#166350) and Albright Hereditary Osteodystrophy (OMIM#174800). In a 26 month old Caucasian boy with congenital hypothyroidism, growth hormone deficiency, vascular & hyperostotic skull, metopic craniosynostosis, and developmental delay, serum PTH & FGF23 levels where elevated. Blood chromosomes and oligonucleotide Chromosomal Microarray Analysis were normal. Exome sequence analysis revealed a novel de novo, inactivating GNAS p.Leu164Pro (c.491T>C) variant that was predicted to be damaging, and was heterozygous in the affected child and absent in both parents. The unusual clinical presentation in this child coupled with the overlapping phenotype reported by Graul-Neumann et al (AJMG, 2009) expand the clinical spectrum in individuals with mutations in GNAS. Exome sequencing with CLIA-compliant confirmation enables diagnosis of inherited diseases that feature clinical or genetic heterogeneity that obfuscates clinical diagnosis by expert physicians or conventional, targeted testing.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2895F

Clinical Diagnostic Exome identifies novel mutations in the *C20orf7* gene, a mitochondrial complex I assembly factor, in a Taiwanese infant with Leigh's disease. *J.J. Wei[1], K.D. Gonzalez[1], X. Li[1], H.M. Lu[1], H. Lu[1], R. Chang[2], R. Wang[2], J. Abdenur[2], W. Zeng[1].* 1) Ambry Genetics, Aliso Viejo, CA; 2) Division of Metabolic Disorders, Children's Hospital of Orange County, Orange, CA.

We report a 14 month-old Taiwanese girl of non-consanguineous union with failure to thrive, gross motor delay, nystagmus, and dysphagia. Laboratory findings are significant for lactic acidosis, ketosis, and abnormal levels of urine organic acids; brain magnetic resonance imaging (MRI) studies demonstrate partial agenesis of the corpus callosum, as well as progressive T2 weighted hyper-intensities in the brain stem, thalamus, and frontal lobes. The family history was negative for similar phenotypes. Additional molecular testing of the mitochondrial genome,*SURF1*, and *POLG* did not reveal any pathogenic mutations. Due to the heterogeneous nature of this patient's symptoms, clinical diagnostic exome sequencing was performed and revealed that the proband has compound heterozygosity for two novel missense mutations: c.155A>C (p.K52T) in exon 1 and c.836T>G (p.M279R) in exon 9 of the *C20orf7* gene (also known as*NDUFAF5*). Co-segregation analysis using Sanger sequencing confirmed that the mother and father each carried one mutation. These two amino acids are completely conserved throughout vertebrates. Furthermore, both alterations are predicted to be possibly damaging and deleterious by Polyphen and SIFT in silico analyses, respectively. *C20orf7* encodes a complex I assembly factor with a predicted N-terminal mitochondrial leader sequence and is imported into the mitochondria; C20ORF7 dysfunction is consistent with the patient's clinical diagnosis of Leigh's disease. Mutations in *C20orf7* have been associated with mitochondrial complex I deficiency, but they have been previously reported in two Ashkenazi Jewish families, as well as consanguineous Moroccan and Egyptian families, with symptoms ranging from neonatal-onset, fatal Leigh's disease to a childhood-onset syndrome of progressive spasticity and choreoathetosis. The proband shares the findings of nystagmus, dysphagia, hypotonia, and MRI T2-hyperintense brain lesions with the previously reported patients. In addition, this report demonstrates that *C20orf7*-related Leigh's disease is a pan-ethnic disease not limited to individuals of North African or Ashkenazi Jewish descent.

2896W

De Novo Mutations in MLL Cause Wiedemann-Steiner Syndrome. *W.D. Jones[1,2,3*], D. Dafou[4], M. McEntagart[1], W.J. Woollard[4], F.V. Elmslie[1], M. Holder-Espinasse[5,6], M. Irving[6], M. Lees[2], A.K. Saggar[1], S. Smithson[7], R.C. Trembath[8], C. Deshpande[6], M.A. Simpson[4].* 1) SW Thames Regional Genetics Service, St George's, University of London, London, SW17 0RE, United Kingdom; 2) NE Thames Regional Genetics Service, Great Ormond Street Hospital for Children, Great Ormond Street, London, WC1N 3JH, United Kingdom; 3) University of Cambridge, Cambridge, United Kingdom, *from 1st August 2012; 4) Division of Genetics and Molecular Medicine, King's College London School of Medicine, Guy's Hospital, London, SE1 9RT, United Kingdom; 5) Service de Génétique Clinique, Hôpital Jeanne de Flandre, Lille, 59037, France; 6) Clinical Genetics, Guy's and St Thomas' NHS Foundation Trust, Guy's Hospital, London, SE1 9RT, United Kingdom; 7) Department of Clinical Genetics, St Michael's Hospital, Bristol, BS2 8EG, United Kingdom; 8) Queen Mary University of London, Bart's and The London School of Medicine and Dentistry, Charterhouse Square, London, EC1M 6QB, United Kingdom.

Excessive growth of terminal hair around the elbows (hypertrichosis cubiti) has been reported both in isolation and in association with a variable spectrum of associated phenotypic features. Hypertrichosis cubiti is present in some, but not all, individuals reported under the clinical classification of Wiedemann Steiner syndrome, a congenital multiple anomaly syndrome associated with short stature, a distinctive facial appearance, and intellectual disability. We identified six individuals with hypertrichosis cubiti associated with short stature, a distinctive facial appearance, and intellectual disability, consistent with a diagnosis of Wiedemann-Steiner syndrome. Utilising a whole exome sequencing approach, we identified de novo mutations in MLL in five of the six individuals. Each of the five mutations is predicted to result in premature termination of the protein product. Expression studies demonstrated that transcripts arising from the mutant alleles are subject to nonsense-mediated decay. MLL encodes a histone methyltransferase, which regulates chromatin-mediated transcription, through the catalysis of methylation of histone H3K4. Our findings define the genetic basis of Wiedemann-Steiner syndrome, provide further evidence for the role of haploinsufficency of histone modification enzymes in congenital multiple anomaly syndromes and further illustrate the importance of the regulation of histone modification in development.

2897T

Targeted resequencing of known and candidate epilepsy genes in 500 patients with epileptic encephalopathies. *G. Carvill[1], B.J. O'Roak[2], S.C. Yendle[3], J. Cook[1], N. Krumm[2], S.F. Berkovic[3], J. Shendure[2], I.E. Scheffer[3], H.C. Mefford[1].* 1) Department of Pediatrics, Division of Genetic Medicine, University of Washington, Seattle, WA; 2) Department of Genome Sciences, University of Washington, Seattle, WA; 3) Epilepsy Research Center and Department of Medicine, University of Melbourne, Austin Health, Australia.

Epilepsy is one of the most common neurological disorders, with a lifetime incidence of 3%. The epileptic encephalopathies (EE) are the most severe of all the epilepsies. Patients typically present with refractory epilepsy with multiple seizure types, cognitive arrest or regression, and have a poor prognosis. Several EE genes and EE-associated CNVs have been described, but a systematic mutation screening approach across multiple EE has not been undertaken and most cases have an unknown etiology. Using molecular inversion probes for targeted capture followed by multiplexed next generation sequencing, we developed a protocol for high-throughput resequencing and copy number variant (CNV) detection of 32 known and 40 candidate genes in a cohort of 518 EE patients. To date, we have analyzed data from 365 individuals for 13 known and 19 candidate EE genes with 93% of targeted nucleotides sequenced at >25 fold coverage. At this coverage threshold we detected 100% of known single nucleotide variants from 20 individuals with variant calls from exome sequencing (n= 609), and 13 individuals with known disease-causing mutations. Additionally, using read depth data, normalization and a standard score analysis we were able to detect 86% of known CNVs from 15 individuals. Segregation analysis of rare, possibly damaging variants in known EE genes detected in 317 EE probands revealed pathogenic changes in *SCN1A*, *GABRG2*, *CDKL5*, *GRIN2A* and *PCDH19*in five families and possibly pathogenic inherited changes in an additional five families. The clinical manifestations of mutation carriers often differed from those previously described, broadening the clinical spectrum of phenotypes associated with known EE genes. We also observed intra-familial variability and incomplete penetrance. Rare variants within our candidate genes, comprising glutamate transporters, the NA+/K+ transporter ATPase-interacting proteins and synaptic vesicle exocytosis proteins is ongoing. Putative CNVs will be validated using a custom oligonucleotide array with high-density coverage of all target genes. In conclusion, we present a cost-effective, efficient method of screening multiple EE genes in large cohorts that will transform molecular testing in EE, enabling rapid early diagnosis, informing prognosis, treatment and genetic counseling.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

† Trainee Award Finalist

2898F

Deciphering the genetic basis of progressive myoclonus epilepsies by exome sequencing. *M. Muona[1,2,3], S.F. Berkovic[4], L. Dibbens[5], K.L. Oliver[4], T. Joensuu[6,7,8], E. Hämäläinen[1,2], A. Palotie[1,2,9], A.-E. Lehesjoki[6,7,8].* 1) Institute for Molecular Medicine Finland, University of Helsinki, Helsinki, Finland; 2) Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton CB10 1SA, UK; 3) Public Health Genomics Unit, National Institute for Health and Welfare, Helsinki, Finland; 4) Epilepsy Research Center, Department of Medicine, University of Melbourne, Austin Health, West Heidelberg, Victoria 3081, Australia; 5) School of Pharmacy and Medical Sciences, University of South Australia, Adelaide, Australia; 6) Folkhälsan Institute of Genetics, University of Helsinki, Helsinki, Finland; 7) Haartman Institute, Department of Medical Genetics and Research Program's Unit, Molecular Medicine, University of Helsinki, Helsinki, Finland; 8) Neuroscience Center, University of Helsinki, Helsinki, Finland; 9) Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA.

Progressive myoclonus epilepsies (PMEs) are a clinically and genetically heterogeneous group of rare, usually recessive, inherited disorders. In a subset of PMEs cognition is essentially preserved and the clinical features are confined predominantly to action myoclonus, tonic-clonic seizures, and ataxia. The paradigmatic form of this is Unverricht-Lundborg disease (ULD), which is caused by mutations in the *CSTB* gene. Other genes for ULD-like phenotypes include *GOSR2* and *SCARB2*. However, a significant number of PME patients do not have mutations in these genes. The identification and characterization of additional PME genes are prerequisites for improved diagnostics of the disease, which relies strongly on genetic testing of the patients. Moreover, it potentially helps to understand the underlying pathogenesis, which is currently largely unknown.

Here, we aimed to identify novel genes underlying patients with a ULD-like phenotype by exome sequencing. We studied 38 patients who have been collected worldwide and excluded for mutations in *CSTB*, *GOSR2*, and *SCARB2* by Sanger sequencing. Based on the clinical variation in the phenotypes we hypothesize that they represent more than one molecular entity. The cases are mostly singletons and a proportion is from consanguineous marriages. We carried out exome sequencing by Illumina Hi-Seq 2000 and selected potentially deleterious variants - either homozygous or compound heterozygous - of essential splice site, nonsense, frameshift, and nonsynonymous types. Additional filtering criteria were based on variant population frequencies, i.e., variants with a 1000 Genomes minor allele frequency of 1% or higher were excluded. Preliminary analysis of the sequencing data confirmed the absence of disease causing mutations in the three known genes for ULD-like phenotypes. Next, we analyzed genes previously linked to other PMEs or epilepsy. Results from this analysis suggest that a small subset of patients may carry causal variants in known epilepsy or PME genes (*EFHC1*, *CLN6*, *EPM2B*). This indicates the clinical diagnosis for these individuals might need re-evaluation or that one gene can contribute to more than one subtype of PME. Finally, we searched for novel PME genes by looking for predicted deleterious variants shared by multiple patients. Potential variants in genes without a previous connection to epilepsy have emerged. All the preliminary results require technical and functional validation.

2899W

Mutations in a Zinc Finger Protein cause a Novel Autosomal Recessive Mental Retardation Syndrome Identified Through Homozygosity Mapping & Whole Exome Sequencing of a Single Affected Individual. *M. Kambouris[1,2], T. BenOmran[3,4], Y. AlSarraj[1], R. Ali[3], M. Almureikhi[3], K. Erraffi[1], H. ElShanti[1,5].* 1) Molecular Genetics & DNA Diagnostics, Shafallah Medical Genetics Center, Doha, Qatar; 2) Yale University School of Medicine, Genetics, New Haven, CT, USA; 3) Clinical & Metabolic Genetics, Department of Pediatrics, Hamad Medical Corporation; 4) Weill Cornell Medical College, Doha, Qatar; 5) University of Iowa, Pediatrics, Iowa City, IA, USA.

A consanguineous Arab family affected by an autosomal recessive disorder characterized by severe mental retardation and failure to thrive, was studied by Illumina's 700K SNP genotyping, candidate gene mutation screening and whole exome sequencing for one affected member. Clinical findings include ptosis, bilateral epicanthic folds, striking midface hypoplasia, downturned mouth corners, thin upper vermillion, prominent ears, bilaterally short 4th metatarsal bone, bilateral 5th finger camptodactyly, mildly limited mobility in both knees and hypotonia. The gene was mapped by homozygosity mapping to 4 possible intervals, as the number of family members was insufficient for a significant LOD score. The positional candidate MAP2K1 was sequenced. No pathogenic mutations identified. Whole exome target enrichment sequencing was performed on ABI SOLiD4 and Illumina HiSeq platforms for a single-affected individual. Six non-synonymous variants were positioned within the homozygosity intervals. Two affected evolutionary highly conserved amino acids with damaging effects according to PolyPhen and SIFT protein-modeling software, were validated by Sanger sequencing and co-segregate with the disease phenotype. The variants were absent in 188 ethnically matched control chromosomes. One gene shows very limited expression in brain while the other, a Zinc Finger Protein, appears to be the disease causing mutation. Zinc Finger Proteins -a family of DNA & RNA binding proteins- are transcriptional regulators controlling developmental cascades of gene expression especially during fetal brain development. Mutations in Zink Finger domains interfere with normal brain development, are associated with non-syndromic X-linked mental retardation with impairments in adaptive behavior and manifest during the developmental period causing severe mental retardation. The c.C5054G [p.S1685W] mutation affects 2 of the 3 ZNF407 isoforms, is located in the last 1/3rd Zink finger domains and affects a serine residue in the alpha-helical part adjoining two zinc-finger domains probably eliminating the functionality of downstream domains and interfering with expression of various genes under ZNF407 control during fetal brain development. Homozygosity mapping and whole exome sequencing of a single affected individual was most effective, least labor demanding and most economical approach in identifying the mutation for this novel autosomal recessive Zinc Finger Protein mental retardation syndrome.

2900T

Exome sequencing of seven cases affected by a congenital muscular dystrophy with hyperlaxity (CMDH) belonging to five French-Canadian families. *M. Tetreault[1,2], G. D'Anjou[3], J. Mathieu[4], M. Vanasse[3], B. Brais[1].* 1) Neurogenetics of motion laboratory, Department of Neurology and Neurosurgery, Montreal Neurological Institute, McGill University, Montreal, Quebec, Canada; 2) Laboratoire de neurogénétique de la motricité, Centre d'Excellence en Neuromique de l'Université de Montréal, CRCHUM, Montréal, Québec, Canada; 3) Clinique des maladies neuromusculaires, Centre de réadaptation Marie-Enfant, Hôpital Sainte-Justine Hospital, Montréal, Québec, Canada; 4) Clinique des maladies neuromusculaires, Carrefour de Santé de Jonquière, Saguenay, Québec, Canada.

Over the past two decades, there have been major advances in defining the genetic bases of congenital muscular dystrophies (CMD). Genetic research has allowed the identification of more than 14 genes responsible for various forms of CMDs. Despite the great progress in this field, there are still a significant percentage of cases for which the mutated gene is unknown. This is particularly the case for milder forms. Novel genomic techniques like Next-Generation DNA Sequencing (NGS) open new avenues in the elucidation of genetic defects causing monogenic disorders such as CMD. We recruited a French Canadian (FC) cohort (27 cases, 24 families) from Southwestern Quebec affected by CMDH. Homozygosity mapping using SNPs genotyping (Illumina OmniExpress) on an informative consanguineous family have allowed us to identify six homozygous region shared by the two affected members. The two affected cases and the mother from the consanguineous family as well as five other affected individuals from three FC families were sent for exome sequencing at Perkin Elmer service lab. The Agilent SureSelect 50Mb kit was used for exons capture and the sequencing was performed on an Illumina HiSeq. A list of rare candidate variants shared by affected individuals that lie in the homozygote regions are being investigated as possibly causing this milder form of CMD. The uncovering of the genetic bases of muscular dystrophies serve as the essential original building blocks on which successful new therapeutic approaches can be designed.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2901F

Diagnosis of syndromic autism using a 62-gene next generation sequencing panel: experience from a clinical diagnostic laboratory. *J.R. Jones, M.J. Basehore, S. McGee, K.J. Champion, M.J. Friez.* Molecular Diagnostic Laboratory, Greenwood Genetic Center, Greenwood, SC.

Autism spectrum disorders (ASD) are complex neurobehavioral conditions that become evident in childhood and persist as lifelong disabilities. These common disorders affect as many as one in 88 children. Coexisting genetic disorders are found in a subgroup of patients with autism (estimated to be between 11 and 37 percent). Many of these genetic disorders have overlapping clinical features, and often numerous single gene sequencing tests are ordered before a pathogenic alteration is identified in a patient. Analyzing these genes one at a time by Sanger sequencing can be costly and time-consuming and may even lead to a delay in diagnosis. In an effort to provide comprehensive testing for this subgroup of patients with syndromic autism in a more rapid and cost-effective manner, the Molecular Diagnostic Laboratory of the Greenwood Genetic Center (GGC) collaborated with investigators from the Emory Genetics Laboratory and RainDance Technologies™ to design a Next Generation Sequencing (NGS) Panel for 62 genes associated with ASD (available commercially from RainDance Technologies™ as the ASDSeq™ Research Screening Panel). These genes were selected to represent the most common single gene etiologies associated with a syndrome that includes autism as a significant clinical feature. The panel was validated at GGC, and diagnostic testing became available in January of this year. Thus far, 66 samples have been submitted for clinical testing. We will discuss the results for these patients and summarize the coverage of this panel along with other run statistics. We have already identified several pathogenic mutations including a mosaic *MECP2* mutation that was previously missed by Sanger sequencing. Our initial findings indicate that this panel provides comprehensive testing for patients with syndromic autism and may also serve as a useful 2nd tier test for patients with features overlapping those of Rett and Angelman syndromes.

2902W

Empowering a clinical genetic setting through the introduction of whole exome sequencing in prenatal and postnatal investigations. *C. Pangalos[1,2], B. Hagnefelt[2], Z. Agioutantis[3], C. Konialis[2].* 1) Clinical Genetics, InterGenetics -Diagnostic Genetic Center, Athens, Greece; 2) Molecular Genetics, InterGenetics -Diagnostic Genetic Center, Athens, Greece; 3) Mineral Resources Engineering, Technical University of Crete, Hania, Greece.

Until now, the `classical' approach to the diagnosis of genetic disorders involves the targeted investigation of specific genes (1–5); however, in everyday clinical practice the geneticist is often confronted with diagnostic riddles, involving rare Mendelian disorders and complex traits, for which phenotype-driven, Sanger sequencing-based testing rarely yields meaningful results. In the last 10 months, we have applied WES in 12 complex clinical cases, involving: an 18yr old male exhibiting multiple renal abnormalities coupled to progressive visual impairment; 3 fetuses, following termination of pregnancy, with skeletal and genital anomalies; an adult with complex poly-neuropathy of unknown etiology; an adult with atypical polycystic kidney disease; an adult with progressive hearing loss; a newborn with possible Treacher-Collins syndrome; a young child with epileptic encephalopathy; a young child with possible Joubert syndrome; a female with degenerative kidney disease and an adult with possible Zellweger syndrome. Exome enrichment was performed using Agilent SureSelect Human All Exon V4 capture, followed by NGS on an Illumina GAIIx at >50x coverage. Sequence alignment and identification of variants was achieved utilizing NextGENe v2.1 sequence analysis software (SoftGenetics), while a custom pipeline (Exome data Management Application - EMA) was developed for variant filtering and mutation identification. In the first of the 2 completed cases to date, known or novel pathological heterozygous mutations were identified in each of 4 genes (NPHP4, RPGRIP1L, CC2D2A, AVIL), consistent with a diagnosis of multi-allelic cilliopathy with retinal degeneration. In the second case of a fetus with skeletal and genital anomalies, a novel pathological hemizygous mutation c.194A>G within the NSDHL gene on Xq28 was identified, and its presence was confirmed in the carrier mother. The presence of this novel mutation in the male fetus is most likely associated with malformation syndromes caused by dysfunction in cholesterol biosynthesis, in agreement with the clinical findings. As WES is proving to be an extremely powerful tool, it may appear as an attractive alternative to often cumbersome clinical evaluation of the patient. Our experience highlights the added value of this new technology when applied in an integrated clinical setting, proving in fact that detailed phenotypic information is a valuable resource for unraveling the genetic causes of complex phenotypes.

2903T

Use of whole exome sequencing to determine the genetic etiology of inherited congenital cataracts. *L.M. Reis[1], R.C. Tyler[1], E.V. Semina[1,2].* 1) Department of Pediatrics and Children's Research Institute, Medical College of Wisconsin, Milwaukee, WI; 2) Department of Cell Biology, Neurobiology and Anatomy, Medical College of Wisconsin, Milwaukee, WI.

Pediatric cataracts are observed in 1–15 per 10,000 births, with 10–25% of cases attributed to genetic causes. Genetic studies have identified mutations in over 34 genes associated with congenital/ juvenile forms of cataracts, but the exact frequencies of mutations in these genes are not known and many cases are still awaiting molecular diagnosis. We screened FOXE3 and PITX3 in 32 probands with congenital cataract and identified two mutations (one in each gene) in probands with cataract and microcornea. The remaining 30 probands were submitted for whole exome sequencing through Perkin Elmer with Geospiza analytical tools; independent gene sequencing was used to confirm identified variants and to screen genes with poor coverage through whole exome sequencing. Initial analysis of known cataract genes in the first 15 set of probands with autosomal dominant congenital cataracts is completed. Probable disease-causing mutations in known genes were identified in nine families including HSF4, CRYBB1, GJA8, BFSP2, CRYGC, CRYGS, CRYGD, and MIP. Ongoing analysis in the remaining 6 families will seek to identify disease-causing mutations in novel genes, utilizing affected and unaffected family members to confirm co-segregation with disease phenotype. Analysis of data from a second set of 15 probands with cataract is ongoing. These data confirm the genetic heterogeneity of congenital cataract, verify the efficacy of utilizing whole exome sequencing to screen a large number of known genes, and suggest that there are additional disease genes yet to be identified.

2904F

A next generation sequencing assay for the diagnosis of the Noonan spectrum of disorders. *A.B. Santani, T. Tischler, A. Sasson, J. Perin, M. Sarmady, E. Frackelton, B. Tweddale, N. Abdel-Magid, S. Panossian, P. Warren, H. Feret, M. Deardorff, A. Wilkens, M. Italia, B. Ruth, J. Miller, E. Zackai, D. Monos, E. Rappaport, H. Hakonarson, P. White, C. Stolle.*Children's Hospital of Philadelphia, 3615 Civic Center Blvd, Philadelphia, PA., USA.

The Noonan spectrum of disorders (NSD) includes Noonan, Costello, LEOPARD and Cardiofaciocutaneous syndromes. This is a clinically heterogeneous group of syndromes with overlapping phenotypic and genetic features. Mutations in multiple genes have been associated with this group of disorders. Testing for NSD using traditional gene sequencing technology can be costly and time consuming. Next generation sequencing (NGS) offers an alternative approach for simultaneous analysis of multiple genes at a reduced cost and improved turnaround time. We have developed and clinically validated a targeted resequencing approach to identify mutations in 11 genes (PTPN11, RAF1, SOS1, BRAF, HRAS, NRAS, KRAS, MEK1, MEK2, SHOC2, CBL) associated with NSD.Amplification of the coding exons and intron/exon boundaries was performed using microdroplet PCR, followed by SOLiD sequencing. After enrichment, up to 10 specimens were barcoded and then pooled for sequencing on a single lane of the SOLiD 5500xl. Sequence alignment, variant detection and annotation were performed using VARify, an in-house bioinformatics workflow. The average depth of coverage was greater than 2000 fold. Approximately 96% of bases had greater than 30X coverage. Amplicons with less than 30X coverage for any base were subjected to Sanger sequencing. On average, less than 5% of amplicons per sample required Sanger sequencing. The ability of this targeted NGS assay to detect mutations in patient samples was validated by resequencing of 25 blinded specimens of known genotypes, of which 6 were from patients with previously characterized disease causing mutations. Analytical performance of the assay was evaluated by comparing genotyping results (obtained by SNP array or Sanger sequencing) and mutation data to NGS results. A total of 168 variants, including 142 single base mutations and 26 indels (up to 6bp), were used for the comparison. All disease causing mutations were correctly identified in all positive control samples; no mutations scored as pathogenic were identified in any negative control samples (no false positives). SNP concordance was 100%. These results demonstrate that both the analytical sensitivity and analytical specificity of the assay is 100%. In summary, we have demonstrated that targeted NGS is a highly sensitive approach for detecting disease causing mutations in genes involved in the Noonan spectrum of disorders and should decrease the time from clinical evaluation to molecular diagnosis.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2905W

French Canadian Founder Mutation in the SBDS Gene causing Shwachman-Diamond Syndrome. *L. Steele[1,2], M. Eliou[1], J.M. Rommens[4], P.N. Ray[1,2], T.L. Stockley[1,3].* 1) Paediatric Lab Medicine, The Hospital for Sick Children, Toronto, Ontario, Canada; 2) Dept of Molecular Genetics; 3) Dept of Laboratory Medicine and Pathobiology; 4) Program in Genetics and Genome Biology, University of Toronto, Toronto, Ontario Canada.

Shwachman-Diamond syndrome (SDS) is an autosomal recessive disorder caused by mutations in the *SBDS* gene, and is characterized by exocrine pancreatic dysfunction, haematological and skeletal abnormalities. The *SBDS* gene (7q11) has an adjacent pseudogene, *SBDS-P* with 97% sequence identity, and the majority of *SBDS* mutations (>80%) arise from gene conversion between exon 2 of the *SBDS* gene and the *SBDS-P* pseudogene. The Molecular Genetics laboratory has offered clinical molecular diagnostic testing for *SBDS* mutations since 2004 (440 probands tested). Molecularly-confirmed SDS probands have either two recurrent mutations in *SBDS* exon 2 from gene conversion, or one recurrent exon 2 mutation and a rare, family-specific mutation in one of the other exons of the *SBDS* gene. Objective: During clinical testing, it was noted that 5 SDS probands had the same rare *SBDS* mutation in exon 1, c.120del, with the second mutation a known recurrent mutation in exon 2 (c.258+2T>C). Recurrent *SBDS* mutations have not been reported outside of exon 2. Further investigation indicated all 5 probands were of French Canadian background, suggestive of a founder mutation. The objective of this study was to determine if the c.120del *SBDS* mutation is a French Canadian SDS founder mutation, using the 5 identified probands and family members. Methods: Sequence analysis of *SBDS* exon 1–5 was performed in all probands, and siblings and parents were tested by targeted mutation analysis as available. Haplotype analysis using STR markers flanking the *SBDS* gene was performed on all samples, and family haplotypes constructed and compared. Results and Conclusions: STR analysis of markers flanking the *SBDS* gene was suggestive of a shared region in four of five families, extending from 3.9 Mb proximal to 0.8 Mb distal of the *SBDS* gene. One of five families had identical shared markers proximal of the gene, but with different markers 0.6 Mb distal of the gene. Overall, results are suggestive of the*SBDS* c.120del mutation being a shared founder in the French Canadian population. Based on our SDS cohort, of whom a total of 48 are of French Canadian background, the incidence of c.120del mutation in SDS probands is may be up to 1 in 20. Therefore, if a child is of French Canadian descent, has failure to thrive and cystic fibrosis testing is negative; a health care provider may want to think of SDS as the frequency of the c.120del mutation appears to be prevalent in this population.

2906T

Confirmation of Cause and Manner of Death Via a Comprehensive Cardiac Autopsy Including Whole Exome Next-Generation Sequencing. *D.J. Tester[1], JJ. Maleszewski[2], TM. Kruisselbrink[2], MJ. Ackerman[1,3,4].* 1) Department of Molecular Pharmacology & Experimental Therapeutics, Mayo Clinic, Rochester, MN; 2) Department of Lab Medicine and Pathology, Mayo Clinic, Rochester, MN; 3) Department of Internal Medicine/Division of Cardiovascular Diseases, Mayo Clinic, Rochester, MN; 4) Department of Pediatric and Adolescent Medicine/Division of Pediatric Cardiology, Mayo Clinic, Rochester, MN.

Annually, the sudden death of thousands of individuals under the age of 35 remains unexplained or poorly characterized despite a medico-legal investigation. Heritable cardiac channelopathies/cardiomyopathies like long QT syndrome (LQTS), catecholaminergic polymorphic ventricular tachycardia (CPVT), Brugada syndrome (BrS), hypertrophic cardiomyopathy (HCM), dilated cardiomyopathy (DCM), and arrythmogenic cardiomyopathy (ARVC) underlie a significant number of sudden unexplained deaths in the young (SUDY). Whole exome sequencing (WES)-based molecular autopsy may provide an efficient manner to identify genetic defects responsible for SUDY. In this pilot study, a WES molecular autopsy was performed on autopsy blood-derived genomic DNA from an otherwise healthy 16-year-old girl who was found dead in her bedroom. Post-mortem examination revealed a moderately enlarged heart with normal ventricular septum to free wall ratio. Histologically, myocyte disarray was noted in the anterolateral left ventricle. Following exome capture, the resulting DNA libraries were sequenced at 96% coverage with a read depth of 35x. Variants were called against the hg19 human reference and analyzed using the TREAT workflow. Following sequencing, gene-specific surveillance of all known LQTS, CPVT, BrS, HCM, DCM, and ARVC genes (68 total) was performed to identify putative disease causing mutation(s). The SUDY victim's exome contained 30,008 variants. Of these, 98 genetic variants were observed among the 68 channelopathy/cardiomyopathy associated genes surveyed. Twenty seven were non-synonymous variants. However, only 1 mutation (R249Q-MYH7) was absent in all publically available exome databases including the 1000 genome project, the NHBLI exome sequencing project, and the 12000 exome chip. Mutations in the MYH7-encoded myosin heavy chain beta (MHC-β) account for ~ 40% of HCM. Specifically, R249Q-MYH7 has been associated previously with familial HCM, sudden death, and impaired MHC-β actin-translocating and actin-activated ATPase activity. Pedigree expansion efforts are currently underway. This is the first report of a post-mortem whole exome sequencing-based cardiac molecular autopsy to confirm the underlying pathogenic substrate most likely responsible for the sudden unexpected death of an otherwise healthy adolescent. WES followed by gene specific surveillance may be a highly efficient and cost-effective approach as part of a comprehensive cardiac autopsy.

2907F

Diagnostic Exome Sequencing Uncovers Mendelian Inheritance in Neurological Disease. *E.C. Chao[1], X. Li[1], K. Gonzalez[1], A. Fatemi[2], J. Cohen[2], w. Zeng[1].* 1) Ambry Genetics, Aliso Viejo, CA; 2) Department of Neurogenetics, Kennedy Krieger Institute and Johns Hopkins School of Medicine, Baltimore, MD.

Whole exome sequencing for clinical diagnostics has been long anticipated, and was first launched by Ambry Genetics one year ago. Here, we describe the two successfulreports of exome sequencing for clinical diagnostic purposes. We sequenced protein-coding regions of the genome (exome) in two families with severe neurologic phenotypes. Both had evaded diagnosis through clinical evaluation and extensive testing over many years. In one family two brothers, with severe mental retardation, autistic features, and self injurious behaviors, were compound heterozygotes for two missense mutations in the *ELP2* gene, which was recently correlated with the phenotype of autosomal recessive intellectual disability (ARID). Each unaffected parent carries one of these mutations; two additional siblings tested by co-segregation analysis carry one and zero of the mutations. In a second case, an adolescent male with spastic diplegia and developmental delay had initially been characterized as suffering from perinatal hypoxic brain injury. Extensive biochemical and genetic testing remained inconclusive. A homozygous nonsense mutation identified in exon 33 of *ALS2*, along with his clinical history is highly consistent with a diagnosis of Infantile-Onset Ascending Hereditary Spastic Paralysis (IAHSP). These findings demonstrate the clinical utility of exome sequencing for diagnostic purposes and highlight initial successes for early adopters of this methodology.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2908W

Whole Exome Sequencing Identifies Potential Homozygous Mutations in a Family with Two Sisters With Brain Malformation. *E. KARACA[1], D. PEHLIVAN[1], C.G. GONZAGA-JAUREGUI[1], M. OZEN[2], A. YUKSEL[3], W. WISZNIEWSKI[1], A. STRAY-PEDERSEN[1], D. MUNZY[4], R.A. GIBBS[1,4], J.R. LUPSKI[1,5,6], Baylor-Hopkins Center for Mendelian Genetics.* 1) Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA; 2) Department of Medical Genetics, Cerrahpasa School of Medicine, Istanbul, Turkey; 3) Department of Pediatrics, Bezmialem University, Istanbul, Turkey; 4) Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX, USA; 5) Department of Pediatrics, Baylor College of Medicine, Houston, TX, USA; 6) Texas Children's Hospital, Houston, TX, USA.

Disorders disrupting normal brain development represent clinically and genetically heterogeneous group which are caused by genomic alterations that involve many different genes. Although recent advances in genomic technologies have led to the discovery of many novel genes that are involved in brain development, many more are likely waiting to be brought to light. To this end, we performed exome sequencing in one family with two affected sisters, born from consanguineous marriage, both showing atrophy of the cerebral cortex, hypoplastic corpus callosum and brain stem, and also having treatment resistant epilepsy, microcephaly, severe growth and developmental delay without speech. Clinical features did not fit with any known syndrome. Exome sequencing revealed probable deleterious mutations in six candidate genes which have not been described before as mutated in brain malformation syndromes. Functional studies of these genes will be performed in order to elucidate their functional impact. Increasing use of exome sequencing approaches, along with other recent genomic technologies, will greatly help to identify causative genes not only in disorders of brain development, but also in other disorders showing Mendelian inheritance.

2909T

Recessive mutations in *MYL2* cause infantile fiber type disproportion and cardiomyopathy. *M. A. J. Weterman[1], P. B. Barth[2], K. Y. van Spaendonck-Zwarts[3], E. Aronica[4], B.-T. Poll-The[2], O. Brouwer[5], J. P. van Tintelen[6], Z. Qahar[1], E. J. Bradley[1], M. B. de Wissel[1], L. Salviati[7], C. Angelini[7,8], L. van den Heuvel[9], Y. E. M. Thomasse[10], A. P. Backx[11], G. Nuernberg[12,13], P. Nuernberg[12,13,14], F. Baas[1].* 1) Dept Genome Analysis/Neurogenetics, K2-213, AMC/UvA, Amsterdam, Netherlands; 2) Div. Pediatric Neurology, Emma Children's Hospital, AMC, Amsterdam, the Netherlands; 3) Dept. Clinical Genetics, AMC, Amsterdam, the Netherlands; 4) Dept. (Neuro)Pathology, AMC, Amsterdam, the Netherlands; 5) Dept. Pediatric Neurology, UMCG, Groningen, the Netherlands; 6) Dept. Genetics, University of Groningen, UMCG, Groningen, the Netherlands; 7) Dept. Pediatrics, University of Padua, Italy; 8) IRCCS S. Camillo, Venice, Italy; 9) Nijmegen Centre for Mitochondrial Disorders, Radboud University Nijmegen Medical Center, Nijmegen, The Netherlands; 10) Dept. of Pediatrics, Pediatric Intensive Care Unit, UMCG, Groningen, the Netherlands; 11) Div. Pediatric Cardiology, Emma Children's Hospital, AMC, Amsterdam, the Netherlands; 12) Cologne Center for Genomics, Cologne, Germany; 13) CMMC, University of Cologne, Cologne, Germany; 14) CECAD, University of Cologne, Cologne, Germany.

IFTDC, a severe disorder with features of fibre type disproportion, myofibrillar disorganization, infantile onset of cardioskeletal myopathy, and early death under 6 months of age was previously reported in three Dutch families. Multipoint parametric linkage analysis of six Dutch patients resulted in the identification of a homozygous region of 2.1 Mb on chromosome 12 that was shared between all Dutch patients with a LOD score of 10.82. Sequence capture and analysis of the entire linkage region identified a homozygous mutation in the last acceptor splice site of the *MYL2* gene as the underlying genetic cause thereby enabling genetic testing for this fatal disorder. *MYL2* encodes a regulatory light chain myosin (MLC-2V). The mutation results in use of a cryptic splice site upstream of the last exon causing a frameshift and replacement of the last 32 codons by 20 different ones. Whole exome sequence analysis of an Italian patient with similar clinical features showed compound heterozygosity for two other mutations affecting the same exon of *MYL2*. The three found mutations disrupt the C-terminal tails of the encoded proteins, affecting the second domain of the two EF-hands which are assumed to function as calcium sensors. Upon binding of calcium, they can undergo a conformational change critical for interactions with downstream targets. Immunohistochemical staining of muscle tissue of the Dutch patients shows a diffuse and mostly weak expression of the mutant protein without the clear fiber specificity seen in normal muscle tissue, while normal protein is absent. Heterozygous missense mutations in *MYL2* have been known to cause dominant hypertrophic cardiomyopathy, however, none of the parents showed signs of cardiomyopathy. In conclusion, the identification of mutations in the last exon of *MYL2* has revealed the cause of this disorder which is a novel autosomal recessive myosinopathy due to defects changing the C-terminal tail of the ventricular form of the myosin regulatory light chain.

2910F

Molecular diagnosis in Stargardts disease through a tiered sequencing approach. *J. Zaneveld[1], H. Wang[1], F. Wang[1], L. Mao[2], K. Zhang[2], Y. Duan[2], J. Zhu[2], J. Chiang[3], R. Chen[1].*1) Molecular and Human Genetics, Baylor College of Medicine, Houston, TX; 2) Institute for Genomic Medicine, UCSD, San Diego, CA; 3) Casey Eye Institute, Oregon Health and Science University, Portland, OR.

Stargardts disease is a form of early onset macular degeneration known to be caused by mutations in ABCA4. In order to find disease causing alleles in patients lacking known Stargardts mutations, we have proceeded with a tiered sequencing approach. First, patient DNA from genes known to cause retinal disease were sequenced using capture sequencing technology. If we were unable to identify a strong candidate for the disease causing allele(s) within the capture region, we proceeded with whole exome sequencing and, when possible, linkage analysis. Using this approach, we have analyzed three large families affected with Stargardts disease, as well as 24 sporadic cases of the disease. Analysis is ongoing, but so far we have identified several novel disease causing allele candidates, as well as a candidate for a novel Stargardts disease causing gene. This research raises the strong possibility that there are novel Stargardts disease genes, and demonstrates the utility of capture sequencing in molecular diagnosis.

2911W

Identification of mutations in the prostaglandin transporter gene SLCO2A1 in Japanese patients with pachydermoperiostosis. *J. Kudoh[1], T. Sasaki[2,3], A. Shimizu[4], A. Shiohama[1], A. Hirakiyama[5,6], T. Okuyama[6], A. Seki[7], K. Kabashima[8], A. Otsuka[8], A. Ishiko[9], K. Tanese[10], S. Miyakawa[10], J. Sakabe[11], M. Kuwahara[12], M. Amagai[3], H. Okano[13], M. Suematsu[14], H. Niizeki[5].* 1) Lab Gene Med, Keio Univ Sch Med, Tokyo, JapanLaboratory of Gene Medicine, Keio University School of Medicine, Shinjuku-ku, Tokyo, Japan; 2) Center for Integrated Medical Research, Keio University School of Medicine, Shinjuku-ku, Tokyo, Japan; 3) Department of Dermatology, Keio University School of Medicine, Shinjuku-ku, Tokyo, Japan; 4) Department of Molecular Biology, Keio University School of Medicine, Shinjuku-ku, Tokyo, Japan; 5) Department of Dermatology, National Center for Child Health and Development, Setagaya-ku, Tokyo, Japan; 6) Department of Laboratory Medicine, National Center for Child Health and Development, Setagaya-ku, Tokyo, Japan; 7) Department of Orthopedics, National Center for Child Health and Development, Setagaya-ku, Tokyo, Japan; 8) Department of Dermatology, Kyoto University Graduate School of Medicine, Kyoto, Japan; 9) First Department of Dermatology, School of Medicine, Toho University, Ota-ku, Tokyo, Japan; 10) Division of Dermatology, Kawasaki Municipal Hospital, Kawasaki, Kanagawa, Japan; 11) Department of Dermatology, Hamamatsu University School of Medicine, Hamamatsu, Shizuoka, Japan; 12) Department of Plastic Surgery, Nara Medical University, Kashihara, Nara, Japan; 13) Department of Physiology, Keio University School of Medicine, Shinjuku-ku, Tokyo, Japan; 14) Department of Biochemistry and Integrative Medical Biology, Keio University School of Medicine, Shinjuku-ku, Tokyo, Japan.

Pachydermoperiostosis (PDP), also known as primary hypertrophic osteoarthropathy, is a rare genetic disorder characterized by 3 major symptoms: pachydermia including cutis verticis gyrata (CVG), periostosis, and finger clubbing. In addition, several other symptoms, including sebaceous hyperplasia, hyperhidrosis, and arthropathy have also been reported. Recently, a homozygous mutation in the gene HPGD, which encodes 15-hydroxyprostaglandin dehydrogenase (15-PGDH), was found to be associated with PDP. We have also attempted to find HPGD mutations in Japanese PDP patients; however, no HPGD mutations have been identified so far, suggesting the existence of other causative gene(s) responsible for PDP in the Japanese population. We therefore aimed to identify a novel responsible gene for PDP using exome sequencing by next-generation DNA sequencer (NGS). Four PDP patients, including 2 patient-parent trios were enrolled in this study. Entire coding regions were captured using the SureSelect Human All Exon v4 (Agilent Technologies, Inc.) and sequenced with the Genome Analyzer IIx (Illumina, Inc.) to identify candidate mutations associated with PDP. The candidate mutations were subsequently sequenced using the Sanger method. To determine clinical characteristics, we analyzed histological samples, as well as serum and urinary prostaglandin E2 (PGE2) levels for each of the 4 PDP patients, and one additional patient with idiopathic CVG. From initial analyses of exome sequencing data we identified mutations in the solute carrier organic anion transporter family member 2A1 (SLCO2A1) gene, encoding prostaglandin transporter (PGT), in all of the three PDP patients examined. Follow-up Sanger sequencing identified 5 different SLCO2A1 mutations (c.940+1G>A, p.E427_P430del, p.G104Ter, p.T347I, p.Q556H) in 4 unrelated PDP patients. However, no significant mutation in the SLCO2A1 gene was identified in the sole patient with CVG. Among the 8 alleles of SLCO2A1 mutant genes identified in this study, 4 (50%) included the same splice-site mutation c.940+1G>A. Haplotype analysis indicated that the c.940+1G>A mutation is a founder mutation in the Japanese population. Although the SLCO2A1 gene is only the second gene discovered to be associated with PDP, it is likely to be a major cause of PDP in the Japanese population.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2912T

CIAS1 mutations associate with familial Mediterranean Fever in Turkish patients and a novel V734G mutation in MEFV. *E. Yosunkaya[1], N. Buyru[2].* 1) Dept. Medical Genetics, Istanbul Univ. Cerrahpasa Medical Sch., Istanbul, Turkey; 2) Dept. Medical Biology, Istanbul Univ. Cerrahpasa Medical Sch., Istanbul, Turkey.

Familial Mediterranean Fever (FMF), an autosomal recessive disease, affecting people originating from Mediterranean basin, is characterized by short lasting acute serositic attacks. The frequency and duration of attacks and accompanying clinical features not only vary among patients, but also between the attacks of a patient. It is suggested that the phenotype is affected by genotype, environmental factors, epigenetic mechanisms and other modifying genes. Symptoms mainly consisting of fever, abdominal and articular pain are due to inflammation cascade problems. It is understood that abnormal pyrin, resulting from MEFV gene mutations, would be inadequate to inhibit the effect of cryopyrin, so inflammation is triggered. In this study with 192 patients referred for MEFV mutation testing, we examined the genotype-phenotype correlations in FMF and investigated CIAS1 gene mutations, which are known to be responsible from cryopyrinopathies. The relationship between MEFV genotypes and the phenotypes with regard to the demographical, clinical and laboratory findings in patients showed statistically significant correlations in approximately 30 parameters. DNA sequencing of the 5' end of the 10th exon of the MEFV gene, which involves most FMF mutations was also performed. A novel V734G mutation was identified in an FMF patient. We detected V200M and L307P mutations in the CIAS1 gene in 6 people with various MEFV mutations. In conclusion, our data, which correlated nearly 30 phenotypic parameters with MEFV genotypes also revealed that V734G, a novel mutation in MEFV gene is responsible from FMF in a patient, and V200M and L307P mutations in CIAS1 gene might have modifying effects on the FMF phenotype.

2913F

Exome sequencing of a dog trio affected with a canine progressive retinal degeneration. *SJ. Ahonen[1,2,3], H. Lohi[1,2,3].* 1) Dept. of Veterinary Biosciences, University of Helsinki, Finland; 2) Dept. of Molecular Genetics, the Folkhälsan Institute of Genetics, Finland; 3) Dept. of Medical Genetics, Research Programs Unit, Molecular Medicine, University of Helsinki, Finland.

Retinitis pigmentosa (RP) is a group of inherited diseases in humans causing retinal degeneration and gradual decline of vision due to a loss of photoreceptor cells function. Primarily RPs are rod degenerations followed by loss of cone function. Many causing genes have been found in humans, but still there are RP patients with unknown genetic cause. Canine progressive retinal atrophy (PRA) is a reminiscent of human RP. PRAs are characterized as well by progressive loss of retinal photoreceptor cells, usually leading to complete loss of vision. Dogs have been successfully used as animal models for human RPs and several mutations have been found in both species. Continental Toy Spaniels include both Phaléne and Papillon breeds, which are separated only based on the posture of the ears. Genetically the breeds are the same and littermates are registered either as Phalénes or Papillons based on their ear positions. Continental Toy Spaniels are affected by a bilateral progressive retinal degeneration with similarities to progressive rod-cone degeneration caused by prcd-gene. We studied a trio of Phalénes with parents and a PRA affected dog to identify the potential coding PRA mutation using exome capture followed by Illumina sequencing. The affected dog was diagnosed with bilateral retinal degeneration at the age of 3 years. Both parents were examined healthy over the age of 6 years. Based on the pedigree analysis, PRA was assumed to be recessively inherited. Data analysis revealed over 150000 single nucleotide variants (SNV) and 36000 indels in each individual. About 26% of SNVs and 0.3 % of indels were known SNPs. Filtering with the assumption of the autosomal recessive mode and using seven other dogs from two other breeds as controls decreased the number of variants to 63 SNVs and 21 indels for validation study in a larger sample cohort. Candidate variants were prioritized according to the relevant function, retinal expression or possible association with retinal degeneration in other species. Ongoing validation is likely to identify a causative gene for PRA in the Continental Toy Spaniels and benefit human RP research as the causative gene can be tested in human RP patients.

2914W

Mutations in NESP4, a KASH domain protein linking the nucleus to the cytoskeleton, lead to hearing impairment in humans and mice. *K.B. Avraham[1], Z. Brownstein[1], H.F. Horn[2], D.R. Lenz[1], S. Shivatzki[1], A. Dror[1], K.J. Roux[3], S. Kozlov[4], B. Burke[2], C.L. Stewart[2].* 1) Department of Human Molecular Genetics and Biochemistry, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel; 2) Institute of Medical Biology, A*STAR, Singapore; 3) Sanford Research/USD Children's Health Research Center, Sioux Falls, SD, USA; 4) National Cancer Institute, Frederick, MD, USA.

Hereditary hearing loss in humans and mice is shown for the first time to be caused by mutations in Spectrin-repeat containing, nuclear envelope 4 protein (NESP4/SYNE4/C19orf46) that is essential for the maintenance of sensory outer hair cells. Two families of Jewish Israeli-Iraqi origin presented with autosomal recessive non-syndromic hearing loss, designated DFNB76. Linkage analysis and exome capture followed by paired-end massively parallel sequencing led to the detection of a 2 nucleotide deletion in the coding region of *NESP4*, with exon skipping leading to a predicted truncation. *Nesp4*$^{-/-}$ mice were found to have a similar pattern of hearing loss, as determined by auditory brainstem response (ABR). First identified in secretory epithelial cells, NESP4 is a KASH-domain protein that resides in the outer nuclear membrane (ONM) as part of the nuclear envelope-associated LINC (Linker of Nucleoskeleton and Cytoskeleton) complexes. These complexes consist of KASH domain proteins in the ONM and SUN domain proteins in the inner nuclear membrane (INM), which together link the nucleoplasm to the cytoskeleton. We found NESP4 to be expressed in the mouse cochlea, in the inner and outer hair cell ONM. Significantly, in the *Nesp4*$^{-/-}$ mice, outer hair cells are formed, but gradually to die as the cochlea matures, leading to progressive hearing loss that is complete by 2 months of age. *Sun1*-null mice exhibited similar cochlear abnormalities. Our findings mark NESP4 and SUN1 as essential proteins for cochlear outer hair cell viability and hearing, potentially due to their role in maintaining correct nuclear position in the outer hair cells of the sensory epithelia.

2915T

Genetic characterization of a large keratinization disorder cohort, utilizing whole exome sequencing to discover mutations in known and novel genes. *L.M. Boyden[1], C. Tian[2], C. Saraceni[2], E. Loring[1], L. Milstone[2], R.P. Lifton[1], H. Vernon[3], D. Zand[4], A.S. Paller[5], K.A. Choate[2].* 1) Department of Genetics, Yale University School of Medicine, New Haven, CT; 2) Department of Dermatology, Yale University School of Medicine, New Haven, CT; 3) Commonwealth Dermatology, Richmond, VA; 4) Children's National Medical Center, Washington, DC; 5) Department of Dermatology, Northwestern University Feinberg School of Medicine, Chicago, IL.

Disorders of keratinization (DOK) are severe skin diseases featuring generalized or local scaling with additional epidermal and systemic sequelae. There is substantial heterogeneity, with dominant, recessive, X-linked, and sporadic forms. Over 50 genes have been identified for DOK, yet mutations in these genes explain only a portion of observed phenotypes. Currently there is no systematic method facilitating both comprehensive identification of mutations in known genes and discovery of novel genes. We have recruited a cohort of 180 well-phenotyped DOK kindreds. We are employing a tiered strategy whereby cases are first screened for mutations in commonly causative genes, and then whole exome sequencing is used either to identify mutations in other known DOK genes or for novel gene discovery. To date, pre-screening for mutations in 8 genes has been performed in 147 kindreds, with mutations identified in 82. In cases of autosomal recessive congenital ichthyosis (ARCI), homo- or compound heterozygous mutations were seen in TGM1, ALOX12B, NIPAL4, CYP4F22, SPINK5, and ALOXE3. In cases of dominant epidermolytic hyperkeratosis (EHK) or ichthyosis with confetti (IWC), heterozygous mutations in KRT1 and KRT10 were observed. Whole exome sequencing of 45 remaining cases has revealed 18 with mutations in known genes. ARCI cases were found to have mutations in ABCA12, PNPLA1, ALOXE3, CYP4F22, and TGM1. Dominant disease included a palmoplantar keratoderma patient with a LOR mutation, and a patient diagnosed with EHK harboring a mutation in KRT2 (known to cause phenotypically similar ichthyosis bullosa of Siemens). Exome sequencing allowed discoveries that would have been overlooked by traditional approaches. In one example, a patient born of consanguineous union was found to be compound heterozygous for ABCA12 mutations, a genetic diagnosis that would have been missed by homozygosity mapping. In another, a patient with severe syndromic ichthyosis had been seen by at least 6 experts and presented at a national dermatology meeting, with no definitive diagnosis. Exome sequencing revealed a hemizygous mutation (R429H) in MBTPS2, an X-linked gene known to cause the typically milder ichthyosis syndromes IFAP and KFSD, but severe disease in the case of R429H. Discovery efforts within the remaining cohort are underway, including sequencing of parent-child trios to identify de novo mutations. Several genes are currently under investigation as potential novel causes of DOK.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2916F

X-linked Familial Focal Segmental Glomerulosclerosis with First-Degree Heart Block caused by a mutation in the NXF5 gene. *T. Esposito[1], R.A. Lea[2], B.H. Maher[2], D. Moses[2], H.C. Cox[2], S. Magliocca[1], V. D'Alessio[1], A. Angius[3], T. Titus[4], T. Kay[4], A. Parnham[4], F. Gianfrancesco[1], L.R. Griffiths[2].* 1) Inst Gen & Biophysics, Italian Natl Res Council, Naples, Naples, Italy; 2) Genomics Research Centre, Griffith Health Institute, Griffith University, Queensland, 4222, Australia; 3) Institute of Genetic and Biomedical Research, National Research Council of Italy, Cagliari, Italy; 4) Gold Hospital Gold Coast, Queensland, 4222, Australia.

Focal segmental glomerulosclerosis (FSGS) is a disease that attacks the kidney's filtering system (glomeruli), causing serious scarring. FSGS is a cause of nephrotic syndrome in children and adolescents and is an important cause of kidney failure in adults. There are currently several genes known to cause the hereditary forms of FSGS that have contributed to the understanding of disease pathology. We recruited a large, multigenerational Australian pedigree of 89 individuals in which FSGS co-segregates with first-degree heart block. Only males exhibit the full form of both disorders, which is suggestive of X-linked recessive inheritance. Of further interest is the occurrence of renal tumours in this family, which has added another element of complexity to understanding the genetic basis of the disease. This family appears to be unique. There have been no reports of any family co-inheriting these disorders, nor has there been a previous report of FSGS transmitted in an X-linked fashion. Through a combined approach of linkage and haplotype analysis, we found that the disease gene(s) reside in a 21.19 cM interval between the markers DXS8077 and DXS8064. We then used a whole exome sequencing approach to identify a mutated gene. Nuclear RNA Export Factor 5 (NXF5) is located in the linkage interval and segregated with the disease phenotype in the pedigree. Moreover, the R113W mutation, in NXF5 gene, was predicted to be deleterious by bioinformatics tools implying a causative role and was not found in a healthy population. The NXF5 gene belongs to the NXF gene family which, in humans, also includes the NXF1, NXF2 and NXF3 genes, all clustered on Xq22.1, except NXF1, which is located on chromosome 11. Evolutionary conservation analysis shows high similarity for these proteins. Interestingly, the arginine at position 113 is highly conserved in all genes of this family and all species. In eukaryotes, the nuclear export of mRNA is mediated by nuclear export factor 1 (NXF1) receptors. As demonstrated in mice, Nxf7 which is analogous to the human NXF5 gene, has cytoplasmic RNA transport factor properties. It has further been shown that Nxf7 interacts with cytoplasmic microtubules and actin filaments and is involved in cellular motility. Functional studies are now ongoing to determine the functional consequence of this novel mutation to development of FSGS and heart block phenotypes.

2917W

Mutations in *SLC52A2* impaire riboflavin transport and cause Brown-Vialetto-Van Laere syndrome. *T.B. Haack[1,2], C. Makowski[3], Y. Yao[4], E. Graf[1], M. Hempel[2], T. Wieland[2], U. Tauer[3], U. Athing[5], J.A. Mayr[6], P. Freisinger[7], H. Yoshimatsu[4], K. Inui[4,8], T.M. Strom[1,2], T. Meitinger[1,2], A. Yonezawa[4], H. Prokisch[1,2].* 1) Institute of Human Genetics, Helmholtz Zentrum München, Neuherberg 85764, Germany; 2) Institute of Human Genetics, Technische Universität München, Munich 81675, Germany; 3) Department of Pediatrics, Technische Universität München, Munich 80804, Germany; 4) Department of Pharmacy, Kyoto University Hospital, Kyoto 606-8507, Japan; 5) Department of Clinical Chemistry, Städtisches Klinikum München, Munich 80804, Germany; 6) Department of Paediatrics, Paracelsus Medical University Salzburg, Salzburg 5020, Austria; 7) Department of Pediatrics, Klinikum Reutlingen, Reutlingen 72764, Germany; 8) Kyoto Pharmaceutical University, Yamashina-ku, Kyoto 607-8414, Japan.

Brown-Vialetto-Van Laere syndrome (BVVLS [MIM 211530]) is a rare neurological disorder presenting with infancy onset sensorineural deafness and ponto-bulbar palsy. Mutations in*SLC52A3* coding for riboflavin transporter 2 (hRFT2) have been recently identified as the molecular genetic correlate in a number of BVVLS individuals. Exome sequencing of a BVVLS individual lead to the identification of compound heterozygous mutations in *SLC52A2*, a new member of the riboflavin transporter family coding for riboflavin transporter 3 (hRFT3). Functional studies confirmed that the gene products of both mutant alleles have reduced riboflavin transport activities. While mutations in *SLC52A3* are associated with decreased plasma riboflavin levels, concordant with a role of *SLC52A3* in riboflavin uptake from food, the *SLC52A2*-mutant individual showed normal plasma riboflavin concentrations, a finding in line with a postulated function of *SLC52A2* in riboflavin uptake from blood into target cells. We initiated an oral high-dose riboflavin supplementation which lead to moderate clinical improvement. However, saturation of intestinal uptake capacity and high rates of renal excretion are likely to limit maximum riboflavin blood concentration. We therefore now started an intravenous riboflavin regimen. Evaluation of the clinical efficacy is ongoing.

2918T

Identification of a causative variant underlying nonsyndromic autosomal-dominant intellectual disability in an extended family. *N. Jinawath[1], B. Pupacdi[2], E. Wohler[3], V. Charoensawan[4], D. Wattanasirichaigoon[5].* 1) Research Center, Faculty of Medicine Ramathibodi Hospital, Mahidol University, Bangkok, Thailand; 2) Chulabhorn Research Institute, Bangkok, Thailand; 3) Kennedy Krieger Institute, Baltimore, MD, USA; 4) Department of Biochemistry, Faculty of Science, Mahidol University, Bangkok, Thailand; 5) Department of Pediatrics, Faculty of Medicine Ramathibodi Hospital, Mahidol University, Bangkok, Thailand.

Intellectual disability (ID) is a common neurodevelopmental disorder with global prevalence between 1 and 3%. Despite its high prevalence, little is known about the genetics of nonsyndromic intellectual disability (NSID). In this study, we set out to identify a causative variant in an extended Thai family with NSID. This family exhibits a four-generation pedigree with extensive history of NSID in every generation, including six affected males and two affected females. The proband is a 5-year-old female with moderate to severe intellectual disability and mildly dysmorphic facies. Her father has mild intellectual disability, and her mother and siblings are phenotypically normal. Karyotype, FISH for DiGeorge syndrome, subtelomeric FISH, and CGG expansion analysis of *FMR1* were all normal. We performed copy number variation (CNV) analysis using the Illumina HumanCytoSNP-12 Beadchip in five affected members of this family, including the proband, and identified no pathogenic or unknown significance CNVs. Assuming that the inheritance pattern is autosomal dominant, we analyzed allele sharing between the five affected individuals using SNPduo and narrowed down the genomic regions which share one allele (Identity by State (IBS) = 1) to 460 Mb. To identify causative variants, we performed whole genome sequencing (WGS) of the proband. We first focused our analysis on exonic regions, which included 16,426 single nucleotide variants (SNV) using ANNOVAR. After filtering out synonymous SNVs and those nonsynonymous SNVs that were in the dbSNP135 and 1000 Genomes databases, considered benign by SIFT and Polyphen, and outside of IBS1 shared regions, we were left with 32 genes with heterozygous missense SNVs. Among these candidates was *SPTAN1*, a gene on chromosome 9q34.11 that belongs to a family of widely-distributed filamentous cytoskeletal proteins known to cause intellectual disability. Dominant-negative mutations in this gene are the cause of early infantile epileptic encephalopathy-5. Additionally, we also identified two known autism and ID-related genes, *ABCA2* and *FASN*, in the 14 small deletion CNVs analyzed from WGS data that overlapped with shared IBS1 regions. The validation of these findings in other affected family members is underway. We propose that genome-wide genotyping analysis of multiple affected members in combination with whole genome sequencing of a proband may help simplify the identification process of causative variants in large family studies.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2919F

Targeted Exome Capture and Paired-End Massively Parallel Sequencing Reveals New Mutations for Human Hereditary Deafness in the Middle East. *M. Kanaan[1], Z. Brownstein[2], A. Abu Rayyan[1], D. Karfunkel[2], D. Dweik[1], Y. Bhonker[2], A. Yeheskel[3], L. Friedman[2], N. Kol[4], O. Yaron[4], V. Oron-Karni[4], M. Frydman[2,5], N. Shomron[4,6], K.B. Avraham[2,4].* 1) Department of Biological Sciences, Bethlehem University, Bethlehem, Palestinian Authority; 2) Department of Human Molecular Genetics and Biochemistry, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel; 3) Bioinformatics Unit, George S. Wise Faculty of Life Sciences, Tel Aviv University, Tel Aviv, Israel; 4) Genome High-Throughput Sequencing Laboratory, Tel Aviv University, Tel Aviv, Israel; 5) Danek Gartner Institute of Human Genetics, Sheba Medical Center, Tel Hashomer, Israel; 6) Department of Cell and Developmental Biology, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel.

Identification of genes responsible for medically important traits is a major challenge in human genetics. While classic techniques such as linkage analysis and Sanger sequencing have led to the discovery of over 100 genes for hearing loss, many more genes remain to be discovered. This is particularly true of the Middle Eastern population, with a large number of different ethnic groups and high rates of consanguinity. Exome capture and massively parallel sequencing can be exploited to address this challenge for hereditary deafness, a highly genetically heterogeneous disease. A targeted capture pool was used for identifying mutations in all known human genes and human orthologs of mouse genes responsible for hearing loss. The Agilent SureSelect Target Enrichment system was used to capture the genomic regions harboring 284 genes including a total of 118 human protein-coding genes, three human microRNAs and the human orthologues of 163 genes associated with inner ear expression or deafness in the mouse. The final capture design targeting 4,475 exons from the 284 genes was 1.86 Mb. The multiplexed libraries representing 96 Palestinian Arab and Israeli Jewish patients (48 libraries per lane) were analyzed with paired-end sequencing at a read length of 2×101 bp, using the Illumina HiSeq 2000. The median base coverage was 113-641X, with minimal coverage of 84% at 10 reads per base and 73% at 30 reads. SNP, indel and CNV analyses were performed. Coordination with homozygosity mapping in consanguineous families optimized bioinformatics analysis. Novel mutations in previously known human deafness genes were discovered. Most compelling, a number of mutations were found in genes previously known only to be involved in mouse deafness. Protein structure predictions were made to provide insight into how the mutations lead to hearing loss. An increase in the number of known genes will have implications not only for the Middle East, but worldwide, as many mutations first found in this region have turned out to be present in other populations. This strategy allows for improved diagnostics, facilitating discovery of the causative mutation in an economically and temporally feasible manner, and establishing etiologically-based genetic counseling and hearing loss management. Further gene discovery will allow for a better understanding of the mechanisms of hearing loss, facilitating therapeutic development. Funded by the National Institutes of Health (NIDCD) R01DC011835.

2920W

Defective ion transport in sweat glands cause generalized isolated anhidrosis. *J. Klar[1], M. Sobol[1], K. Mäbert[1], M. Tariq[2], M. Rasool[2], M. Jameel[2], T. Naeem[2], N.A. Malik[2], A. Johansson[1], L. Feuk[1], S.M. Baig[2], N. Dahl[1].* 1) Department of Immunology, Genetics and Pathology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden; 2) Human Molecular Genetics Laboratory, National Institute for Biotechnology and Genetic Engineering (NIBGE), Faisalabad, Pakistan.

Anhidrosis is defined as reduced or absent perspiration in response to an appropriate stimulus. Reports on congenital isolated generalized anhidrosis with morphologically normal sweat glands are extremely rare. We have identified a consanguineous Pakistani family comprising in total 13 individuals affected by generalized and isolated anhidrosis segregating as an autosomal recessive trait. Affected individuals show no other associated ectodermal abnormalities and they present with normal hair, teeth and skin. The patients suffer from heat intolerance and exposure to heat results in severe uncomfort with increased body temperature accompanied by a compensatory increase of heart rate. Starch-iodine test on skin surface of affected individuals excluded normal hydrolysis of starch supporting complete anhidrosis.

To identify the genetic basis for this inherited form of anhidrosis we performed autozygosity mapping using SNP array (Affymetrix 250k Nsp array). We identified a single homozygous region on chromosome 13 and subsequent fine mapping using microsatellite markers confirmed a shared homozygous region in all affected individuals. A custom array of the entire candidate region (NimbleGen Sequence Capture Microarrays, Roche) was used to enrich for DNA in one affected individual. High throughput sequencing (Illumina HiSeq) of enriched DNA identified a unique homozygous missense variant after filtering against common polymorphisms in available databases. The missense variant alters an evolutionary conserved codon and it segregates in a homozygous state with the disease in all affected family members. Software predictions indicate that the mutation has a severe functional impact on the protein which is involved in intercellular ion signaling. Histological analysis of paraffin sections from human skin using an antibody specific to the gene product confirms expression in eccrine sweat glands.

We present herein and to our knowledge, the first examples of a gene behind isolated anhidrosis. The findings emphasize the critical role of ion binding and transport in normal sweating.

2921T

Exome sequencing identifies a Novel Mutation in GJA1 as the Cause of Alopecia, Leukonychia, Skin Hyperkeratosis and Craniofacial Anomalies. *Z. Lin, H. Wang, J. Zhang, Q. Chen, M. Lee, X. Wang, R. Li, Y. Yang.* Dermatology, Peking University First Hospital, Beijing, Beijing, China.

Background Exome sequencing, the next-generation sequencing technology, is a powerful and cost-effective approach for dissecting the genetic basis of monogenic disorders, even clinically undiagnosed diseases. Objectives To identify the causative gene of a familial case presenting with alopecia, leukonychia, skin hyperkeratosis and craniofacial anomalies using exome sequencing. Methods Mutation analyses of KRT1, KRT10, GJB2 and GJB6 were performed in the proband. Exome sequencing was applied subsequently to identify the causative gene, followed by Sanger sequencing for verification and to rule out the possibility of single nucleotide polymorphism (SNP). Results No pathogenic mutations were detected in KRT1, KRT10, GJB2 or GJB6. Among the filtered variants obtained by exome sequencing, a missense mutation (p.Gly8Val) in GJA1 was noted and verified by Sanger sequencing. Cosegregation of the mutation with the phenotype in the family was confirmed and the possibility of SNP was excluded. Conclusions We identified a novel missense mutation in GJA1 underlying the phenotype of alopecia, leukonychia, skin hyperkeratosis and craniofacial anomalies in a familial case by exome sequencing.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2922F

Mutation discovery in ENU-derived mouse models for metabolic bone disease by whole chromosome and whole exome sequencing. *B. Lorenz-Depiereux[1], S. Sabrautzki[2], S. Diener[1], E. Graf[1], T. Wieland[1], A. Benet-Pagès[1], S. Eck[1], B. Rathkolb[2,3], M. Klaften[4], J.A. Aguilar-Pimentel[2,5], J. Calzada-Wack[6], E. Janas[6], E. Wolf[3], M. Ollert[5], F. Neff[6], M. Hrabé de Angelis[2,7], T.M. Strom[1,8].* 1) Institute of Human Genetics, Helmholtz Center Munich, German Research Center for Environmental Health, Neuherberg, Germany; 2) Institute of Experimental Genetics, Helmholtz Center Munich, German Research Center for Environmental Health, Neuherberg, Germany; 3) Chair for Molecular Animal Breeding and Biotechnology, Gene Center of the Ludwig-Maximilians-University Munich, Munich, Germany; 4) Karlsruher Institute of Technology DE Innovation Management (IMA), Eggenstein-Leopoldshafen, Germany; 5) Klinikum Rechts der Isar der TU Munich, Clinic and Policlinic of Dermatology and Allergology, Munich, Germany; 6) Institute of Pathology, Helmholtz Center Munich, German Research Center for Environmental Health, Neuherberg, Germany; 7) Chair of Experimental Genetics, TU Munich, Freising-Weihenstephan, Germany; 8) Klinikum Rechts der Isar der TU Munich, Institute of Human Genetics, Munich, Germany.

Animal models are required to understand the molecular mechanisms of metabolic bone disorders in which imbalances of bone metabolism and mineralization lead to a variety of different phenotypes. Spontaneous mutations, targeted mutations and conventional knockout models are routinely used to study the molecular mechanism of a putative mutation. Random N-ethyl-N-nitrosourea (ENU) mutagenesis is a complementary approach to obtain mouse models for human diseases. Within a large-scale genome-wide Munich ENU Mutagenesis Project mutant mouse models for metabolic bone disease were identified using three blood plasma parameters (total alkaline phosphatase activity (ALP), total calcium (Ca), and inorganic phosphate (Pi) levels) which were commonly used as biochemical markers in patients with metabolic bone disease. All mutants were generated on a C3HeB/FeJ genetic background. The availability of massively parallel sequencing provides an opportunity to sequence mouse mutants with manageable costs at an appropriate time. Here, we describe two different sequencing strategies based on next generation sequencing technology, (1) chromosome sequencing, and (2) exome sequencing, to identify novel genes involved in phosphate homeostasis and bone metabolism. In two mutant mouse lines with high ALP activity (BAP004, BAP005), the disease causing mutations were identified by linkage analysis, chromosome sorting and whole chromosome sequencing (BAP004: Jak1 c.1933T>C, p.Ser645-Pro and BAP005: Asgr1 c.815A>G, p.Tyr272Cys). We performed exome sequencing using in-solution exome capture (Agilent) on HiSeq2000 instruments (Illumina) in 19 mutant mouse lines (3 recessive and 16 dominant lines) and one C3HeB/FeJ wt line as control. We identified between 3 and 42 SNVs (single nucleotide variants) in each mutant mouse line. Mapping data (chromosomal linkage) were available for 8 mouse lines. Each putative SNV was evaluated using segregation analysis of a cohort of mutant and control mice by capillary sequencing. To date, we identified one candidate gene within the 3 recessive mutant mouse lines. Within the 16 dominant mouse lines 7 candidate SNVs (4 novels and 3 common) were identified, 7 mouse lines are currently under investigation, and in 2 mouse lines no mutation has been identified yet. For the specific evaluation of the potential causative SNVs further studies as functional analysis or association to human diseases are necessary.

2923W

Exome sequencing identified a novel mutation in CACNA1S in a Japanese family with Malignant Hyperthermia. *N. Matoba[1], H. Iwasa[2], M. Kohda[1], Y. Yatsuka-Kanesaki[2], Y. Ichihara[3], H. Kikuchi[3], Y. Suzuki[4], S. Sugano[4], Y.K. Hayashi[5], I. Nishino[5], Y. Okazaki[1,2].* 1) Division of Translational Research, Research Center for Genomic Medicine, Saitama Medical University, Hidaka, Saitama, Japan; 2) Division of Functional Genomics & Systems Medicine, Research Center for Genomic Medicine, Saitama Medical University, Hidaka, Saitama, Japan; 3) Department of Anesthesiology, Saitama Medical University Hospital, Moroyama, Saitama, Japan; 4) Laboratory of Functional Genomics, Department of Medical Genome Sciences, Graduate School of Frontier Sciences, the University of Tokyo, Minato-ku, Tokyo, Japan; 5) Department of Neuromuscular Research, National Institute of Neuroscience, National Center of Neurology and Psychiatry (NCNP), Kodaira, Tokyo, Japan.

Background Malignant Hyperthermia (MH) is a disorder of skeletal muscle triggered by some anesthetics and muscle relaxants. MH is frequently inherited in an autosomal dominant manner. To date, 6 genetic regions associated with MH have been reported, but only 2 genes have been identified, which are *RYR1* (encoding ryanodine receptor 1) and *CACNA1S*(encoding L-type Ca^{2+} channel subunit α). Approximately, 80% of patients with MH have mutated *RYR1*. As previously reported (Carlos, et al.,2006), we experienced a MH case without*RYR1* mutations by mutation screening of entire *RYR1* coding region.

Methods We performed whole-exome sequencing of 5 individuals of a Japanese family with MH using Hiseq2000(Illumina). Three of them were healthy and the rest were the carriers. (diagnosed by the rate of Ca^{2+} induced Ca^{2+} release from SR)

About 23,000 genetic variants in the exon/splicing sites were identified per individual. To reduce the numbers of variations and identify potentially pathogenic mutations, we used network analysis and expression analysis following a prioritization scheme like previously reported (Hoistchen et al., 2010).

Results Three carriers had 243 variants in common and those were not identified in healthy individuals. 158 candidate genetic variants remained after excluding the known variations (MAF>0.01 in dbSNP132). After comparison to the genes predicted by either network analysis or expression analysis, we identified p.I639V mutations in *CACNA1S*. The mutation was confirmed by Sanger sequencing of the DNA from the family including proband. The position of the mutation is likely on transmembrane (S6 of repeat II) according to Uniprot and the corresponding amino acid was highly conserved.

Conclusion We identified a novel heterozygous mutation in *CACNA1S* in the proband and two carriers that was unidentified in the healthy subjects in the family. Further work includes electrophysiological analysis for confirmation. Also, we are currently working on the exome sequencing analysis of additional families with MH.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2924T

Identification of a novel missense mutation in RAD51 in a large family with congenital mirror movements. *O.E. Onat[1], S. Gulsuner[1], R. Bilgen[2], G.M. Dal[1], K. Bilguvar[3], H. Boyaci[4,5], K. Doerschner[4,5], H. Uysal[2], M. Gunel[3], T. Ozcelik[1,6].* 1) Department of Molecular Biology and Genetics, Faculty of Science, Bilkent, Ankara 06800, Turkey; 2) Department of Neurology, Akdeniz University Hospital, Antalya 07058, Turkey; 3) Department of Neurosurgery, Department of Neurobiology and Department of Genetics, Center for Human Genetics and Genomics and Program on Neurogenetics, Yale University School of Medicine, New Haven, CT 06510, USA; 4) Department of Psychology, Faculty of Economics, Administrative and Social Sciences, Bilkent University, Ankara 06800, Turkey; 5) National Research Center for Magnetic Resonance, Bilkent University, Ankara 06800, Turkey; 6) Institute of Materials Science and Nanotechnology, Bilkent University, Ankara 06800, Turkey.

Congenital mirror movements (CMM) are a rare and heterogeneous group of disorders characterized by involuntary contralateral movements of mainly the upper extremities during intentional movements on the opposite side. Isolated cases are usually familial and suggest autosomal dominant inheritance with incomplete penetrance. In two chromosome 18 linked families, causative mutations were identified in DCC (Science 328:592, 2010; MRMV1; MIM:157600). Here, we describe a three-generation consanguineous Turkish family with six members affected by CMM. Linkage analysis with a dominant model and 90 percent penetrance parameters resulted in peaks on 15q13.3-q21.1, 15q26.2, and 19q12 with maximum multipoint LOD scores of 3.6, 2.6, and 2.6, respectively. However, a region of homozygosity segregating with the phenotype was not observed, and thus excluded the possibility of recessive inheritance of the disease allele in this consanguineous family. Whole-exome sequencing of an affected individual uncovered 7 coding, 33 intronic and 3 intergenic novel variants located within the three linkage intervals, which were filtered against the dbSNP132 dataset. Segregation analysis, population filtering using 1000 genomes and EVS data sets, and conservation considerations using prediction tools revealed a novel missense mutation (c.404C>T [p.T134N], RefSeq accession number NM_002875) in exon 5 of RAD51 (MIM:179617), consistent with the dominant inheritance of the disease allele in the family. The mutation resides in the highly conserved AAA (ATPases associated with diverse cellular activities) domain of the protein, and it was not observed in 436 chromosomes from healthy individuals coming from a geographical matched region. Recently, truncating mutations in RAD51 were identified in two families with CMM (Am J Hum Genet 90:301,2012; MRMV2; MIM:614508). Our findings support the totally unexpected role of RAD51 in neurodevelopment and further suggest that alterations of this gene may lead to neurological phenotypes.

2925F

***De Novo* Mutations of the Gene Encoding the Histone Acetyltransferase KAT6B Cause Genitopatellar Syndrome.** *M.A. Simpson[1], C. Deshpande[2], D. Dafou[1], L.E.L.M. Vissers[3], W.J. Woollard[1], S.E. Holder[4], G. Gillessen-Kaesbach[5], R. Derks[3], S.M. White[6], P. Cohen-Snuijf[7], S.G. Kant[8], L.H. Hoefsloot[3], W. Reardon[9], H.G. Brunner[3], E.M.F.H. Bongers[3], R.C. Trembath[1,10].* 1) Division of Genetics and Molecular Medicine, King's College London School of Medicine, Guy's Hospital, London SE1 9RT, UK; 2) Clinical Genetics, Guy's and St Thomas' National Health Service (NHS) Foundation Trust, Guy's Hospital, London SE1 9RT, UK; 3) Department of Human Genetics, Radboud University Nijmegen Medical Centre, Nijmegen 6500 HB, The Netherlands; 4) North West Thames Regional Genetics Service, North West London Hospitals NHS Trust, Harrow HA1 3UJ, UK; 5) Institut für Humangenetik, Universität zu Lübeck, Lübeck 23538, Germany; 6) Genetic Health Services Victoria, Murdoch Childrens Research Institute, Melbourne 3052 Australia; 7) Ipse de Bruggen, Centre for People with Intellectual Disability, Nieuwveen 2441 CP, The Netherlands; 8) Centre for Human and Clinical Genetics, Department of Clinical Genetics, Leiden University Medical Center, Leiden 2300 RC, The Netherlands; 9) National Centre for Medical Genetics, Our Lady's Hospital for Sick Children, Crumlin, Dublin 12, Ireland; 10) Queen Mary University of London, Barts and The London School of Medicine and Dentistry, London E1 2AD, UK.

Genitopatellar syndrome (GPS [MIM 606170]) is a rare disorder in which patellar aplasia or hypoplasia is associated with external genital and renal anomalies, congenital flexion deformities of the limbs, distinctive facial features, corpus callosum agenesis, and severe intellectual disability. To search for disease alleles in GPS we undertook whole exome sequencing of six unrelated affected individuals. All had characteristic clinical features of GPS. As sibling recurrence of GPS has previously been observed, we initially compared the six exome profiles under a model of a rare autosomal-recessive inheritance, but did not identify any gene harboring novel homozygous or compound heterozygous protein-altering variants in more than one individual. Subsequently, we interrogated these data under the assumption that each of the causative alleles is dominant and had arisen *de novo*; requiring at least one previously unobserved heterozygous nonsynonymous or splice-site substitution or a coding insertion or deletion in the same gene in all six individuals. This process did not reveal a single gene matching the criteria in all six individuals. However, evaluation of these data with a prior expectation of genetic heterogeneity highlighted *KAT6B* as the only candidate gene harboring previously unobserved heterozygous variants in five of the six individuals; a single nonsense variant (c.3877A>T p.Lys1293X) and three frameshift indels, including a 4 bp deletion observed in two unrelated affected individuals (c.3768_3771del p.Lys1258GlyfsX13, c.3680_3695del p.Asp1227GlufsX11, and c.3773dup p.Trp1259ValfsX12). All identified mutations are located within the terminal exon of the gene and are predicted to generate a truncated protein product lacking evolutionarily conserved domains. Using an RT-PCR approach we demonstrate that total abundance of *KAT6B* transcript in individuals with GPS is equivalent to wild-type controls and were able to detect a protein of the size of the predicted truncated KAT6B in primary skin fibroblasts from an affected subject. *KAT6B* encodes a member of the MYST family of histone acetyltranferases. We demonstrate a reduced level of both histone H3 and H4 acetylation in patient-derived cells suggesting that dysregulation of histone acetylation is a direct functional consequence of GPS alleles. These findings define the genetic basis of GPS and illustrate the critical and complex role of the regulation of histone acetylation during development.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2926W

Focal Facial Dermal Dysplasia, Type IV is associated with mutations in *CYP26C1*. *A. Slavotinek[1], P. Mehrotra[1], B. Li[1], I. Nazarenko[2], P. Ling-Fung[3], R.Z. Lao[3], C. Chu[3], M. Yahyavi[1], C. Chou[4], A.L. Marqueling[5], K. Cordoro[5], I. Frieden[5], M-A. Morren[6], K. Devriendt[7], T. Prescott[8], T. Glaser[4], P-Y. Kwok[3], M. Petkovich[9], R. Desnick[2].* 1) Dept. Pediatrics, U585P, University California, San Francisco, San Francisco, CA; 2) Dept. of Genetics and Genomic Sciences, Mount Sinai School of Medicine, New York, NY 10029; 3) Cardiovascular Research Institute, University of California, San Francisco San Francisco, CA. 94143-3118; 4) Dept. Internal Medicine and Human Genetics, University of Michigan, Ann Arbor; 5) Dept. of Dermatology, University of California, San Francisco, San Francisco, CA94143-0316; 6) Dermatology Unit, University of Leuven, Leuven, Belgium; 7) Center for Human Genetics, University of Leuven, Leuven, Belgium; 8) Dept. Medical Genetics, Oslo University Hospital, Oslo, Norway; 9) Dept. of Biomolecular and Medical Sciences, Cancer Research Institute, Queen's University, ON K7L 3N6.

We report a sibling pair with focal facial dermal dysplasia (FFDD) Type IV manifesting as lesions resembling aplasia cutis located in a preauricular distribution on both cheeks along the line of fusion of the maxillary and mandibular facial prominences. Both children had small, intraoral polyps on their buccal mucosa, a finding that has not previously been reported. Mutations in the*TWIST2* gene, causative for Setleis syndrome, were excluded in both offspring. We assumed autosomal recessive inheritance and used exome sequencing to identify novel sequence variants that were shared by both siblings. We found two mutations in *CYP26C1* - c.844_851dupCCATGCA, predicting p.Glu284fsX128, a tandem duplication of 7 basepairs that was inherited from the children's mother and c.1433G>A, predicting p.Arg478His, that was inherited from their father and that was predicted to be deleterious. Sequencing of additional type IV FFDD patients revealed that three previously reported individuals (Prescott et al., 2006) were homozygous for the same duplication as found in this family. The duplication was present in 1/250 (0.3%) Caucasian control chromosomes, 0/200 Hispanic control chromosomes and 1/236 (0.4%) chromosomes from individuals with eye or diaphragmatic birth defects. *Cyp26c1* is involved in retinoic acid (RA) degradation, together with *Cyp26a1* and *Cyp26b1*, and the gene is expressed in the first pharyngeal arch at the location of embryological fusion of the maxillary and mandibular processes. There is no known phenotype in animal models of loss of gene function on a wildtype background. The effects of disordered retinoid metabolism on skin development have previously been studied in *Cyp26b1* homozygous null mice, which have reduced thickness of the cornified layer of skin, reduced filaggrin and arrest of hair follicle growth at E17.5 and E18.5. Experiments using RA-soaked beads at specific developmental stages in embryonic chicken skin explant cultures also showed that the beads can create a radial zone of inhibition of skin appendage formation (Chuong et al., 1992), demonstrating that excess RA may induce skin defects. The low frequency of the duplication in control chromosomes poses the question as to whether variable in-utero exposure to retinoids can influence penetrance of the phenotype. We conclude that mutations in *CYP26C1* are associated with FFDD type IV in humans. Chuong et al. Development 1992;115:839; Prescott T et al. Eur J Med Genet 2006;49:135.

2927T

Identity-by-Descent Mapping Reveals a New Locus for Primary Congenital Glaucoma, GLC3E, on Chromosome 19p13.2. *H. Verdin[1], B.P. Leroy[1, 2], B. D'haene[1], F. Coppieters[1], S. Lefever[1], P.G. Kestelyn[2], E. De Baere[1].* 1) Center for Medical Genetics, Ghent University, Ghent, Belgium; 2) Department of Ophthalmology, Ghent University Hospital, Ghent, Belgium.

Primary congenital glaucoma (PCG) is caused by developmental anomalies of the trabecular meshwork and the anterior chamber angle, resulting in an increased ocular pressure (IOP) and optic nerve damage from birth or early infancy. The prevalence of PCG is estimated to be 1:10.000 in Western populations with higher prevalences in inbred populations. In general PCG displays an autosomal recessive inheritance and is genetically heterogeneous. To date, four PCG loci are known (GLC3A-D), in which two genes have been identified, *CYP1B1* and *LTBP2*. Here, we aimed to map the disease gene in a large, four-generation consanguineous family with PCG, originating from Jordan. Mutations in known PCG genes (*CYP1B1* and *LTBP2*) as well as anterior segment dysgenesis genes (*FOXC1* and *PITX2*) were excluded respectively. Identity-by-descent (IBD) mapping was performed in six affected members using genomewide SNP genotyping with 250K arrays (Affymetrix). Several gene prioritization tools were used to make a selection of candidate genes for Sanger sequencing. Two affected individuals underwent exome enrichment (TruSeq Exome Enrichment Kit, Illumina) and sequencing (2×100 cycles, HiSeq, Illumina). The CLC Genomics Workbench (CLC bio) was employed for read mapping and variant calling. The common IBD regions did not overlap with any known PCG loci. Filtering on both size of the region and number of consecutive homozygous SNPs revealed a new candidate region on 19p13.2, named GLC3E. This region measures 2.67 Mb and contains 93 genes. Using prioritization tools, *BEST2* was selected as the best candidate gene. Indeed, *Best2* is expressed in non-pigmented epithelial (NPE) cells of ciliary body, which is responsible for formation of aqueous humour. Also, *Best2*$^{-/-}$ mice have significantly lower IOP than wild type littermates. Sanger sequencing of the *BEST2* coding region in affected individuals revealed no mutations however. Exome sequencing was performed in two affected individuals. Data analysis is ongoing. We identified a potential new PCG locus, named GLC3E, confirming the genetic heterogeneity of PCG, and representing a unique opportunity to identify the third PCG gene.

2928F

Mutations in *HOXB1* cause autosomal recessive congenital facial palsy with sensorineural hearing loss and strabismus. *B.D. Webb[1], S. Shaaban[2, 3, 4], H. Gaspar[1, 5], L.F. Cunha[1], C.R. Schubert[3, 6], K. Hao[1], C.D. Robson[7], W. Chan[2, 8], C. Andrews[2, 3, 8], S. MacKinnon[9], D.T. Oystreck[10, 11], D.G. Hunter[9], A.J. Iacovelli[1], X. Ye[1], A. Camminady[5], E.C. Engle[2, 3, 8, 9], E.W. Jabs[1].* 1) Department of Genetics and Genomic Sciences, Mount Sinai School of Medicine, New York, NY; 2) Department of Neurology, Children's Hospital Boston, Boston, MA; 3) Department of Neurology, Harvard Medical School, Boston, MA; 4) Department of Ophthalmology, Mansoura University, Daqahlia, Egypt; 5) Institute of Human Genetics, University Hospital Freiburg, Freiburg, Germany; 6) Department of Pediatrics, Harvard Medical School, Boston, MA; 7) Department of Radiology, Children's Hospital Boston, Boston, MA; 8) Howard Hughes Medical Institute, Chevy Chase, MD; 9) Department of Ophthalmology, Children's Hospital Boston, Boston, MA; 10) Department of Ophthalmology, College of Medicine, King Saud University, Riyadh, Saudi Arabia; 11) Division of Ophthalmology, University of Stellenbosch, Tygerberg, South Africa.

The HOX genes are a family of highly conserved, homeodomain-containing transcription factors which specify regional differences along the anterior-posterior axis during early development. In this study, we utilized whole exome sequencing to identify a causative gene mutation in *HOXB1* in a family with autosomal recessive congenital facial palsy with hearing loss and strabismus. 175 additional probands with similar clinical features were screened for mutations in *HOXB1* (102 with a diagnosis of Moebius syndrome) and a second family was identified with the same homozygous *HOXB1* c.619C>T mutation. The clinical features of these two families were examined and correlate extensively with the phenotype of the *Hoxb1*$^{-/-}$ knock-out mouse reported previously. This Arg207Cys HOXB1 missense variant identified results in a substitution at the highly conserved arginine 5 residue of the homeodomain, which when mutated in other proteins is known to cause human disease. Modeling of the mutant HOXB1/PBX1 complex shows reduction in affinity for DNA. Additionally, transfection experiments show differential activation of the mutant versus wild-type HOXB1 for its auto-regulatory element. Because both affected families were of conservative German American background, haplotype analysis was performed and both families shared a common haplotype, suggesting a founder mutation. Here is the first demonstration of human developmental disorders that can be classified as HOXB1opathies.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2929W

Genetic heterogeneity underlying female infertility and sensorineural hearing loss and identification of *LARS2* as a third gene for Perrault syndrome. *S.B. Pierce[1], T. Walsh[1], R. Michaelson-Cohen[2], R.E. Klevit[3], M.-C. King[1], K. Gersak[4], E. Levy-Lahad[2].* 1) Department of Medicine (Medical Genetics), University of Washington, Seattle, WA; 2) Medical Genetics Institute, Shaare Zedek Medical Center, Jerusalem, ISRAEL; 3) Department of Biochemistry, University of Washington, Seattle, WA; 4) Department of Obstetrics and Gynecology, University Medical Center, Ljubljana, SLOVENIA.

Perrault syndrome is a heterogeneous recessive disorder characterized by ovarian dysgenesis in females and sensorineural hearing loss in males and females, with additional neurological symptoms present in some cases. We previously identified mutations in two genes, *HARS2* and *HSD17B4*, as the cause of Perrault syndrome in two families, reflecting this genetic heterogeneity. *HARS2* encodes the mitochondrial histidyl-tRNA synthetase and thus is required for translation of proteins encoded by the mitochondrial genome. *HSD17B4* encodes a 17beta-hydroxysteroid dehydrogenase, which is a peroxisomal enzyme involved in fatty acid beta-oxidation and steroid metabolism. Through exome sequencing, we have now identified mutations in a third gene, *LARS2*, encoding the mitochondrial leucyl-tRNA synthetase, as the cause of Perrault syndrome in two other families. The proband of a consanguineous Palestinian family, and her two brothers with hearing loss, are all homozygous for the private variant *LARS2 c.1565C>A* (p.T522N). The proband of a non-consanguineous Slovenian family is compound heterozygous for private variants *LARS2 c.1076delT* and *LARS2 c.1886C>T* (p.T629M). LARS2 p.T522N affects a conserved residue adjacent to residues involved in leucyl-adenylate substrate binding, suggesting that this mutation may alter enzyme activity. *LARS2 c.1076delT* leads to a stop at codon 374 of the 903-amino acid protein. LARS2 p.T629M affects a residue in the leucyl-specific insertion domain, whose role in catalysis is poorly understood. We are evaluating the functional consequences of this mutation at the transcript and protein levels. A role for leucyl-tRNA synthetase in maintenance of fertility is supported by a previously identified mutation in the *Caenorhabditis elegans* homolog *lars-2*, which results in a loss of germ cell differentiation and sterility. The fact that *LARS2*, like *HARS2*, is involved in mitochondrial translation also supports a causal role for the *LARS2* mutations in Perrault syndrome. Other families with Perrault syndrome under study in our laboratory do not carry mutations in *HARS2*, *HSD17B4*, or *LARS2*. Causative mutations in still other genes will expand our understanding of the biology underlying Perrault syndrome.

2930T

Exome sequencing in apparently recessive and apparently X-linked HSP families identifies an ATL1 mutation hotspot which is associated with low penetrance dominant inheritance. *C. Beetz[1], R.E. Varga[1], H. Fadel[2], R. Schüle[3, 4], I. Valenzuela[5], F. Speziani[4], G. Rudenskaja[6], G. Nürnberg[7], H. Thiele[7], J. Altmüller[7], V. Alvarez[8], J. Gamez[9], P. Nürnberg[7], S. Zuchner[4].* 1) Institute of Clinical Chemistry, University Hospital Jena, Jena, Germany; 2) Department of Neurology, Al Kortobi Hospital, Tangier, Morocco; 3) Department of Neurodegenerative Disease, Hertie-Institute for Clinical Brain Research and Center for Neurology, Tübingen, Germany; 4) Hussman Institute for Human Genetics, University of Miami Miller School of Medicine, Miami, FL; 5) Clinical Genetic Unit, Hospital Vall d'Hebrón, Barcelona, Spain; 6) Genetic Counselling Department, Medical Genetics Research Centre, Russian Academy of Medical Sciences, Moscow, Russia; 7) Cologne Centre for Genomics, University of Cologne, Cologne, Germany; 8) Laboratory of Molecular Genetics -Genetic Unit, Hospital Universitario Central de Asturias, Oviedo, Spain; 9) Department of Neurology, Hospital Universitari Vall d'Hebron, VHIR Universidad Autonoma de Barcelona, Barcelona, Spain.

The hereditary spastic paraplegias (HSPs) are a diverse group of movement disorders. They are classified according to age at onset (early vs. late), clinical presentation (pure vs. complex), and mode of inheritance, or according to the genetic defect. Despite >45 loci and >20 HSP genes identified, cumulative frequencies derived from studies targeting individual genetic forms seem to not even explain 80% of cases. This may partially be due to inappropriate patient selection for such gene-centred investigations. We performed exome sequencing in index cases from two large families in which several members each suffer from clinically pure, early onset forms of HSP. Consanguinity of the parents in family 1 and only males being affected in family 2 suggested autosomal recessive and X-linked inheritance, respectively. A homozygous variant in ATL1 (c.1244G>A, p.R415Q), i.e. a well-established gene for the autosomal dominant form SPG3A, was the only plausible variant in family 1. It was also homozygous in two other patients but, surprisingly, heterozygous in the fourth patient and in seven unaffected family members. In family 2, none of the variants found on the X-chromosome segregated. However, all patients carried one ATL1 c.1243C>T (p.R415W) allele and this variant was shared by at least five unaffected females and one unaffected male. Interestingly, p.R415W has been reported by one previous study where it was also associated with reduced penetrance. We subsequently identified one out of 87 apparently sporadic patients and two out of 23 dominant families to also harbour the p.R415W mutation. Direct or indirect evidence for non-manifesting carriers was again obtained. As geographic origin and haplotype analysis argued against a founder effect for p.R415W, we considered a mutation hotspot. Indeed, bisulfate sequencing revealed c.1243 and c.1244 to represent a completely methylated CpG dinucleotide, spontaneous deaminations of which would explain the above transition mutations. A substantial fraction of the patients diagnosed in the present study would not routinely have entered targeted ATL1 sequencing (apparently X-linked, apparently recessive, apparently sporadic). SPG3A due to alterations at the mutation hotspot we identified may therefore be a rather frequent but currently underestimated cause of HSP. Consequently, we suggest targeted screening for c.1243 and c.1244 alterations in all patients with pure and/or early onset HSP regardless of family history.

2931F

Autosomal Recessive Lethal Congenital Contractural Syndrome Type 4 (LCCS4) Caused by a Mutation in MYBPC1. *O.S. Birk[1,2], B. Markus[1], G. Narkis[1], R.Z. Birk[3], I. Cohen[1], D. Landau[2].* 1) Morris Kahn Center for Human Genetics, National Institute for Biotechnology in the Negev, Ben-Gurion University, Beer-Sheva, Israel; 2) Genetics Institute, Soroka Medical Center, Beer-Sheva, Israel; 3) Ariel University Center, Ariel, Israel.

Autosomal recessive lethal congenital contractural syndrome (LCCS) is a severe form of neuromuscular arthrogryposis. We previously showed that this phenotype is caused in two unrelated inbred Bedouin tribes by homozygous mutations in either ERBB3 or in PIP5K1C of the phophatidyl inositol pathway. However, the molecular basis of the same phenotype in other tribes remained elusive. Whole exome sequencing identified a novel LCCS founder mutation within a minimal shared homozygosity locus of ~1Mb in two affected individuals of different tribes: a homozygous premature stop producing mutation in MYBPC1, encoding myosin binding protein C, slow type. A dominant miss-sense mutation in MYBPC1 was previously shown to cause mild distal arthrogryposis. We now show that a recessive mutation abrogating all functional domains in the same gene leads to LCCS. This study highlights the power of whole exome sequencing in elucidating the molecular basis of phenotypes common to remotely related individuals in inbred communities.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2932W

Exome sequencing of a pedigree with Caudal Regression Syndrome (CRS). *G. CHENG[1], E.HM. WONG[2], PC. SHAM[2], S.S. CHERNY[2], S. MAAS[3], S.W. SCHERER[4], C.R. MARSHALL[4], S.L. PEREIRA[4], P.KH. TAM[1], M.M. GARCIA-BARCELÓ[1].* 1) Department of Surgery, Li Ka Shing Faculty of Medicine, University of Hong Kong, Hong Kong; 2) Department of Psychiatry, Li Ka Shing Faculty of Medicine, University of Hong Kong, Hong Kong; 3) Department of Pediatrics and Clinical Genetics, Academic Medical Center Amsterdam, Netherlands; 4) Department of Molecular and Medical Genetics, The Centre for Applied Genomics and Program in Genetics and Genomic Biology, The Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada.

Caudal Regression Syndrome (CRS) is a rare, congenital disorder characterized by varying degrees of agenesis of the caudal spine, and is usually associated with general anomalies of other systems. Although maternal diabetes is a well-known risk factor, the exact cause of CRS is still unknown. CRS mainly occurs sporadically, except for Currarino syndrome (CS) or CRS sub-phenotype IV, which follows an autosomal-dominant pattern of inheritance. We present a three generation Caucasian family of 11 individuals of whom 6 are affected with CRS subtype III or V based on the sacral/coccygeal agenesis pattern. A presacral mass was present in 5 of the 6 affected individuals. The medical history of family was re-assessed and the phenotype/individual ID was confirmed. G-banding karyotype of the affected members was normal. To identify the genetic defect underlying the disorder, DNA from the 11 family members were submitted to whole genome single nucleotide polymorphism genotyping (Omni2.5) and whole exome sequencing. Further, using 31,520 pruned tag SNPs, parametric linkage analysis using MERLIN revealed 4 linkage peaks with LOD score≈1.8 under the autosomal dominant inheritance model. The linkage peak regions lied in chr1, chr11 and chr12 and spanned 19.57Mb (ranges from 1.3Mb to 8.2 Mb; ≈ 33.18cM). Whole exome sequencing for the identification of causal variants was performed using Illumina HiSeq platform with TruSeq Exome enrichment kit, and SOLID5500 platform with Agilent SureSelect Exome enrichment kit. The sequencing data retrieved ≈65K recalibrated single nucleotide variants/insertions/deletions for each individual on average and the output showed a concordance rate of 96.2% with Omni2.5 (reads mapping by bwa/bfast+bwa/mrsFast; post-processing and variants calling by GATK pipeline/mpileup/Varscan/freebayes). Besides, we performed structural variants (SVs) calling from Omini2.5 density file using PennCNV, iPattern, QuantiSNP, or CNVPartition as well as from the exome sequencing data using BreakDancer-Max, VariationHunter, Pemer, SpitRead or ExomeDepth. Filtering for causal variants was performed according to 1) novel small variants within the linkage peak regions which were shared only by patients; 2) large structural variants (>0.1Mb, based on the tag SNP resolution for linkage analysis) each of which were examined for the inheritance pattern observed. Several plausible genes with functional mutations have been identified.

2933T

Loss of function mutations in OMIM genes reveal a burden of disease susceptibility in a consanguineous population. *K.A. Fakhro[1], J.L. Rodriguez-Flores[2], N.R. Hackett[2], J. Salit[2], J. Fuller[2], J.A. Malek[1], L. Chouchane[1], R. Badii[3], A. Al-Marri[3], J.G. Mezey[2,4], R.G. Crystal[2].* 1) Dept of Genetic Medicine, Weill Cornell Medical College, Doha, Qatar; 2) Dept of Genetic Medicine, Weill Cornell Medical College, New York, NY; 3) Hamad Medical Corporation, Doha, Qatar; 4) Dept of Biological Statistics and Computational Biology, Cornell University, Ithaca, NY.

Exome sequencing was used to study genetic variation in the Qatari population. We sequenced 100 healthy Qataris (50X average depth, with >80% of sites at >10x depth) representing the 3 major Qatari genetic subpopulations. Reads were aligned to the GRCH37 human reference genome using BWA 0.59, variants were called using GATK 1.2 and annotated using SNPEFF 2.1. A total of 174×10^3 variants were identified with quality scores >50, of which 585 were predicted to cause loss of function (frameshift/stop-gain/splice) in 435 unique OMIM disease genes. Of these, 184 (35%) were high-quality, novel Qatari variants (absent from public databases) of which 115 were observed in 2 or more individuals, suggesting a population allele frequency ≥1%. In order to enhance specificity, we discarded all mutations that occur in highly repetitive regions and poly-N tracks or that occurred at a position that was within 2 bp of at least 1 common SNP or indel and any variants in alternatively spliced exons. After filtering, there were a total of 20 high-confidence loss of function mutations in 20 unique OMIM genes. Ten of these were single-base substitutions that led to stop-gains, 3 were single-base substitutions of a conserved splice-site residue within 2 bp of the intron/exon junction, 6 were insertions (1–2 bp), and 1 a deletion (1 bp). All exonic substitutions resulted in the introduction of stop-codons within the transcript in silico, with a median resultant truncated protein size of 48% of full-length (range 8–95%). None of these truncations occurred in an alternatively spliced exon, suggesting that the majority of these mutations could affect gene expression levels. Half of these mutations (n=10) truncated genes responsible for autosomal recessive severe childhood disease. For example, 22 individuals (AF 11%) had a single-base substitution resulting in a stop-gain in TBX15 (Cousin syndrome; truncated protein size: 66%) and 13 individuals had a 2-base insertion resulting in a downstream stop-codon in NPHP3 (nephronophthisis; truncated protein size 42%). Further, truncations in the thyroid receptor TSHR (hypothyroidism) and insulin receptor INSR (diabetes mellitus) were observed at frequencies of 11.5% and 2% respectively. Overall, the carrier frequency for these 20 mutations was rare-to-moderate (mean 4.9%, range 1–12%), which bears significant implications for a region of the world with a high rate of consanguinity.

2934F

Causal Gene Discovery in Mendelian Disorders Using Whole Exome Sequencing. *S. Jhangiani[1], M. Bainbridge[1], J. Lu[1], M. Wang[1], H. Dinh[1], Y. Han[1], J. Santibanez[1], M. Caramins[3], P. Campeau[2], B. Lee[2], J. Reid[1], J. Lupski[2], E. Boerwinkle[1], D. Muzny[1], R. Gibbs[1].* 1) Human Genome Sequencing Center, Baylor College of Medicine, Houston, Texas; 2) Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas; 3) South Eastern Sydney Area Health Service, Sydney, Australia.

Using next-generation sequencing technologies with predominantly whole exome applications, the Human Genome Sequencing Center has generated 14Tb of high quality data for greater than 1,500 exomes in efforts towards discovering causative alleles for Mendelian disorders. More than 100 studies, extending locally from the Texas Medical Center to worldwide collaborations, have ranged in disorders including neurological, skeletal, cardiac, tumor formation, immunological and other phenotypes. Our high-throughput multiplexed capture pipeline, producing 2K capture libraries per month, has enabled cost efficient and high quality data production. A new implementation to the pipeline enables automated annotated variant file generation using Atlas SNP 2 for SNP identification. The utility of whole exome sequencing and an automated pipeline has proven to be highly effective in associating genes with diseases. Exome sequencing can be used to elucidate specific pathways responsible for diseases. For three individuals with genitopatellar syndrome, a skeletal dysplasia expressing cerebral and genital anomalies, exomes were sequenced and showed *de novo* heterozygous truncating mutations in *KAT6B*. Mutations in this gene resulted in a loss of the highly conserved transcription activation domain affecting skeletogenesis. In a study of spinocerebellar ataxia two related individuals were sequenced using two orthogonal chemistries and shared a rare deleterious hemizygous mutation on *GJB1*. This gene has been identified in patients with CMT and Dejerine-Sottas syndrome. This work supports the use of using whole exome sequencing in the discovery of disease causing alleles.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2935W

Whole genome sequencing and copy number analysis of exome sequencing in two families with split-hand/split-foot malformation identifies chromosomal rearrangements affecting putative exonic enhancers. *H. Lango Allen[1], R. Caswell[1], P. Turnpenny[1,2], C. Turner[1,2], C. Wragg[3], W. Xie[1], M. Weedon[1], X. Xu[1], S. Ellard[1].* 1) Peninsula Medical School, University of Exeter, Exeter, UK; 2) Department of Clinical Genetics, Royal Devon and Exeter Hospital, Exeter, UK; 3) Regional Genetics Unit, Southmead Hospital, Westbury on Trym, Bristol, UK.

We report three individuals from two families with split-hand/split-foot malformation (SHFM) in whom next generation sequencing was performed to investigate the cause of their phenotype.

The first proband has a de novo balanced translocation t(2;7)(p25.1;q22) identified by karyotyping following a normal array CGH result. To characterise the breakpoints, we undertook paired-end whole genome sequencing on a single Illumina HiSeq2000 lane, and obtained mean coverage of 8x. This was sufficient to characterise the breakpoints: there were seven read pairs where one of the pairs mapped to chromosome 2 and the other to chromosome 7, and there were four reads partially mapping to both chromosomes, i.e. across the breakpoints. The exact breakpoints were confirmed by Sanger sequencing of a single PCR product. The chromosome 7 breakpoint is situated within the SHFM1 locus on chromosome 7q21.3 and results in the physical separation of the recently identified DYNC1I1 exonic enhancers (eExons) from their target genes, *DLX5* and *DLX6*, implicated as candidates for SHFM in mouse knockout models.

The mode of inheritance within the second family was thought to be X-linked recessive, and exome sequencing was performed to search for a mutation within the SHFM2 locus at Xq26 in the affected proband and his uncle. No coding mutation was found but copy number analysis of exome data identified a ~100kb deletion within the chromosome 7 SHFM1 locus. The *DLX5*and *DLX6* genes are disomic, but the deletion includes the putative *DYNC1I1* exon 15 and 17 enhancers.

Split hand foot malformation type 1 is an autosomal dominant trait with reduced penetrance and variable expression. We have narrowed the critical region from 0.9Mb to 0.1Mb and this deletion provides further evidence for a key role of DYNC1I1 exonic enhancers in normal limb formation.

2936T

Searching novel genes for hereditary hearing loss in multiplex families using next generation sequencing. *Y.H. Lin[1], C.C. Wu[1,2], Y.C. Lu[1], C.J. Hsu[1], P.L. Chen[2].* 1) Department of Otolaryngology, National Taiwan University Hospital, Taipei, Taiwan; 2) Department of Medical Genetics, National Taiwan University Hospital, Taipei, Taiwan.

Despite the clinical utility of genetic diagnosis in addressing idiopathic sensorineural hearing impairment (SNHI), current strategy of mutation screening using Sanger sequencing suffers from the limitation that only a limited number of DNA fragments associated with common deafness mutations can be genotyped. Consequently, definite genetic diagnosis cannot be achieved in many families with obvious family history. To investigate the diagnostic utility of next generation sequencing (NGS) and to delineate the genetic epidemiology of uncommon deafness genes in an East Asian population, in the present study, we applied the NGS technique to 12 Han Chinese multiplex families (i.e. families with multiple affected members) with idiopathic SNHI. NimbleGen sequence capture array was designed to target all exons and 100 bp of flanking sequence from 79 common deafness-associated genes. NGS by HiSeq2000 using paired-end (2×100) protocol was performed to get a greater than 30-fold coverage across the targeted regions. We applied bowtie, samtools, GATK, Picard and IGV for the bioinformatics analyses. After variant calling, variant allele frequency > 5% in 5400 NHLBI exomes and in 1000 Genomes Project 2012 were filtered out. PolyPhen2 and SIFT were used to assess the effect of non-synonymous changes on the protein. Five indel (insertion/deletion) variants and 36 non-synonymous variants were identified in the 12 multipex families. After validation by Sanger sequencing and mutation effect assessment by PolyPhen2 and SIFT, twelve variants with high probability of changing the function of gene were observed in 5 autosomal dominant families, and 1 autosomal recessive family, respectively. Those variants include one known pathogenic variant (*GJB2* p.R75Q) and 11 novel variants (*GRHL2* p.K9R, *KCNQ4* p.S680F, *CCDC50* p.D295Y,*MYH9* p.E1256K, *MYO1A* p.E456G, *GJB4* p.C169W, *DIAPH1* p.R1146W, *GJB6* p.L76fs, *WFS1* p.V412A, *CDH23* p.D428N, and *CDH23* p.V858I). The assignment of causality for certain variants needs to be achieved by checking the segregation pattern in the pedigrees and possible functional study in the future. In conclusion, NGS enables a precise genetic diagnosis in certain multiplex families with idiopathic SNHI by identifying mutations in relatively uncommon deafness genes.

2937F

Combination of genomic technologies and consanguinity in order to identify pathogenic variants in recessive disorders. *P. Makrythanasis[1], M. Nelis[1], F.A. Santoni[1], M. Guipponi[2], F. Béna[2], A. Vannier[1], G. Duriaux-Sail[1], S. Gimelli[2], E. Stathaki[2], E. Falconnet[1], S. Temtamy[3], A. Megarbane[4], M. Aglan[3], M.S. Zaki[3], S. Fokstuen[2], A. Bottani[2], A. Masri[5], S. Psoni[6], S. Kitsiou[6], H. Fryssira[6], N. All-Allawi[7], A. Sefiani[8], S. Al-Hait[9], S. Elalaoui[8], N. Jalkh[4], L. Al-Gazali[10,11], F. Al-Jasmi[10,11], H. Chaabouni Bouhamed[12], H. Hamamy[1], S.E. Antonarakis[1,2].*1) Dept Medical Genetics & Dev, University of Geneva, Geneva, Geneva, Switzerland; 2) Service of Genetic Medicine, University Hospitals of Geneva, Geneva, Switzerland; 3) Department of Clinical Genetics, National Research Centre, Cairo, Egypt; 4) Medical Genetics Unit, Saint Joseph University, Beirut, Lebanon; 5) Pediatric Department, Jordan University Hospital, Amman, Jordan; 6) Department of Medical Genetics, University of Athens, Athens, Greece; 7) Department of Pathology, College of Medicine, University of Dohuk, Dohuk, Iraq; 8) Département de génétique médicale, Institut National d'Hygiène, Rabat, Morocco; 9) Genetic Clinic, Al Amal Maternity Hospital, Amman, Jordan; 10) Department of Paediatrics, Faculty of Medicine and Health Sciences, United Arab Emirates University, Al-Ain, United Arab Emirates; 11) Department of Pediatrics, Tawam Hospital, United Arab Emirates University, Al-Ain, United Arab Emirates; 12) Department of Human Genetics, University Tunis El Manar, Faculty of Medicine, Tunis, Tunisia.

Consanguinity and inbreeding increase the sharing of alleles among individuals; thus a considerable number of autosomal recessive phenotypes occur in offspring(s) of consanguineous couples. We have collected samples from consanguineous families with phenotypes of unknown etiology that are compatible with autosomal recessive transmission, in order to identify the functional genomic variation responsible for the disease. Any phenotype and family history compatible with autosomal recessive inheritance (and unknown molecular defect) was a candidate for participation in the study. 42 families of different ethnic background have been collected so far. From each family, blood DNA from the patient(s), all unaffected siblings and the parents is extracted. Samples are i/ analyzed by array-CGH for the detection of homozygous deletions; ii/ genotyped with a 720K SNP array in order to identify Runs of Homozygosity (ROH) and define the areas of the genome that could include the causative variant; iii/ exome sequenced (one affected individual per family). The variants within the targeted areas are selected for the downstream analysis. The mean coverage is 130x and 98.2% of the total coding region of RefSeq is covered at least 8x. By comparing the genotyping with the sequencing data, we found that the high throughput sequencing for Single Nucleotide Variants (SNV) that passed the quality threshold had a specificity of 99.95%, sensitivity of 97.7%, positive predictive value (PPV) 99.2% and negative predictive value (NPV) 98.6%. On average we identified 21901 variants per exome. So far we completely analysed 26 families and identified the causative variation in known genes in 3 of them: in *VLDLR*, causing disequilibrium syndrome, in *FKTN* causing Fukuyama muscular dystrophy and in *DMP1* causing hypophosphatemic rickets. In 12 further families 23 candidate genes/variants have been identified (more than 1 candidate genes were found per family). We present the genomic variants and the corresponding phenotypes of these families. In the remaining 11 families the likely molecular defect has not been identified using our experimental protocol. Consanguineous families provide an opportunity to identify pathogenic variants in known as well as candidate genes responsible for recessive phenotypes and rapidly fill in the gap between genotype and phenotype.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2938W

A single exome variant is the only expected variant by likelihood ratio for a rare heritable de novo dominant disorder in a three generation family with two affected. *S.M. Marchegiani[1,5], T.C. Markello[1,2,3], L.A. Wolfe[1], K. Fuentes-Fajardo[1], D.R. Adams[1,2,3], W.A. Gahl[1,2,3], J.C. Mullikin for NISC[4], T. Davis[1], J.P. Accardi[1], C.J. Tifft[1,3], C.F. Boerkoel[1,2,3].* 1) NIH Undiagnosed Diseases Program, Bethesda, MD; 2) Medical Genetics Branch, National Human Genome Research Institute; 3) Office of the Clinical Director, National Human Genome Research Institute; 4) NIH Intramural Sequencing Center, National Human Genome Research Institute; 5) Department of Pediatrics, National Capital Consortium, Bethesda, MD.

High throughput sequencing (HTS) refined the predicted mutation rate to 1.3 × 10-8 per base pair per generation, which yields approximately 1.3 (n= 0-4, P>0.95) mutations per diploid exome per generation. Given the known mutation rate and size of the exome, there is a P<0.05 probability that exome sequencing (ES) will identify more than one variant that is both deleterious and consistent with inheritance in the case where a de novo mutation occurred in a parent and correctly segregates into two offspring in a third generation. We illustrate this concept by identifying one novel variant in a de novo affected father and his affected daughter within a three generation family. The father and daughter carry the diagnosis ablepharon macrostomia (AMS), a rare presumed dominant disorder characterized by absent eyelids, dysmorphic facies, lax redundant skin and ambiguous genitalia. Exome sequencing identified a T>A substitution in the 5'UTR of TRIM23, a well-conserved ADP-ribosylation factor (ARF) gene on Chromosome 5 that has a role in vesicular trafficking. The variant from exome sequencing is only present in the affected father and daughter, and is absent in all unaffected family members, including both paternal grandparents. This variant is not reported in current databases. Consanguinity was excluded by high density single nucleotide polymorphism (SNP) array analysis. Pseudodominant inheritance is unlikely for a rare disease in an outbred family. Evidence of (non)penetrance was critically examined during clinical phenotyping, including comprehensive eye and skin exams with light and EM histopathology. We confirmed complete absence of features consistent with ablepharon macrostomia in the unaffected paternal grandparents, unaffected mother and unaffected brother, supporting a de novo dominant mutation in the father with subsequent inheritance by his daughter. We additionally established elastic fiber electron microscopy as another modality to clinically phenotype affected versus unaffected individuals in this family. The presence of a rare, heritable de novo dominant disorder in a three generation family provides an ideal model to propose that an identified and Sanger validated variant, like the one described for TRIM23 in this pedigree, is the only expected variant and thus likely causal for the condition in that family.

2939T

Whole-genome sequencing identifies mutations in known and novel genes for early infantile epileptic encephalopathy. *H.C. Martin[1], A.T. Pagnamenta[1], K. Hudspith[1], A. Rimmer[1], R. Copley[1], E. Sadighi Akha[1], J. Broxholme[1], A. Kanapin[1], J.-B. Cazier[1], D. Shears[2], H. Stewart[2], D. Bentley[3], J. Taylor[1], E. Blair[2], P. Donnelly[1].* 1) Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom; 2) Department of Clinical Genetics, Oxford Radcliffe Hospitals National Health Service Trust, Churchill Hospital, Oxford, United Kingdom; 3) Illumina Cambridge, Chesterford Research Park, Little Chesterford, Cambridge, UK.

Early Infantile Epileptic Encephalopathy (EIEE) is a severe neurological disorder causing intractable seizures, intellectual disability and developmental delay. One subtype of EIEE, Ohtahara syndrome (OS), produces a characteristic “burst-suppression” pattern on the electroencephalogram. OS patients are routinely screened for mutations in several genes, including *ARX*,*CDKL5*, *SPTAN1* and *STXBP1*, as well as for large copy number variants, but most cases remain unexplained. As part of the WGS500 project at the Wellcome Trust Centre for Human Genetics, we used the Illumina HiSeq system to do high-coverage whole-genome sequencing (WGS) of six EIEE trios in which these standard tests had been negative: four OS cases and two girls with severe early-onset seizures. In each trio, we screened for rare or novel coding variants that would fit a *de novo*, simple recessive, compound heterozygous or X-linked model. Two OS cases had novel *de novo* nonsynonymous mutations in ion channel genes, one in *KCNQ2* and the other in *SCN2A*, both of which had previously been implicated in EIEE. Another OS case proved to have paternal isodisomy for chromosome 9, and, consistent with a recessive model on this chromosome, we identified a novel homozygous nonsynonymous mutation in an ion channel gene that has recently been linked to a similar epilepsy phenotype. In the fourth OS patient there were two strong candidates: compound heterozygous nonsynonymous variants in an ion channel gene that has been associated with muscular disorders, and a homozygous variant that affected splicing of a gene in the glycophosphatidyl inositol biosynthesis pathway. The two other epilepsy patients had *de novo* mutations in genes involved in Ras/MAPK and Wnt signalling. These three pathways have been previously linked to epileptogenesis (Bassuk et al. AJHG, 2008; Adachi et al. Seizure, 2012; Johnston et al. AJHG, 2012) and functional studies are underway to determine whether these variants are pathogenic. This work highlights the need to expand the panel of genes routinely screened in EIEE patients, emphasises the utility of WGS in elucidating unanticipated disease mechanisms and sheds new light on pathways important for neural development.

2940F

Novel intellectual disability genes identified by exome sequencing. *R. Rabionet[1], O. Drechsel[1], A. Puig[1], J. Gonzalez[1], I. Madrigal[2], M.I. Alvarez[2], N. Baena[3], M. Viñas[3], S. Ossowski[1], M. Guitart[3], M. Mila[2], X. Estivill[1].* 1) Genes & Disease programme, Center for Genomic Regulation, Barcelona, Spain; 2) Servei de Genètica, Hospital Clínic i Provincial de Barcelona, Barcelona, Spain; 3) Laboratori de Genètica, Corporació Sanitària Parc Taulí, Sabadell, Barcelona, Spain.

Intellectual disability (ID) is a genetically heterogeneous disorder affecting 1–3% of the population. In many cases, it presents as a sporadic disorder, although it can also show X-linked, AR and AD inheritance. About 30% of cases of ID can be explained by structural variants, and more than hundred genes have been implicated in its pathophysiology. Nevertheless, a large proportion of cases remains unexplained. A recent study has shown that *de novo* point mutations are a frequent cause of ID, which would make next generation sequencing technologies extremely powerful tools to identify rare *de novo* genetic causes of ID.

We recruited 35 trio cases of sporadic idiopathic ID, all of them negative for the fragile X expansion and without cytogenetically visible abnormalities. The exome was captured with Agilent sure select technology, multiplexed, and sequenced in an Illumina HISeq2000 lane. Potential damaging variants were detected by an in-house developed pipeline for variant annotation. Briefly, non-synonymous variants present in the case are selected and filtered based on their predicted functionality (*condel* score) and their frequency in known databases (EVS, 1000genomes, dbSNP). Then, *de novo* variants as well as potential X-linked or recessive genes (carrying homozygous or compound heterozygous variants in the case) were identified and annotated against a list of known ID-related genes. As a final step, the presence and *de novo* state of the identified variants were confirmed by Sanger sequencing. In the first 14 trios analysed, we detected an average of 8520 non-synonymous variants in the cases, of which an average of 27 were *de novo*. Frequency and functionality based filtering reduced the number of potential candidate ID genes harboring *de novo* variants to 0–10 per case. Two male cases carried X-linked variant in a known ID gene, two cases were compound heterozygous for two rare/novel potentially damaging non-synonymous variants in ID genes, and in one case we detected a *de novo* variant in a previously described ID gene. In two independent trios we identified a different *de novo* splicing variant in the same kinase gene. In four other cases we identified novel *de novo* variants in genes related with neurite outgrowth, gene silencing or epigenetic control of gene expression, which would be good candidates for novel ID genes.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2941W

Targeted deep resequencing identifies a mutation in MID2, as causal for X-linked intellectual disability with varied disease severity. *B. K. Thelma[1], S. G. Thenral[1], A. Michealraj[1], M. Kabra[2], G. Kaur[3], R. C. Juyal[4].* 1) Dept Genetics, Univ Delhi, South Campus, New Delhi, India; 2) Department of Pediatrics, All India Institute of Medical Sciences (AIIMS), New Delhi; 3) Department of Physiology, Government Medical College and Hospital, Chandigarh, India; 4) National Institute of Immunology, New Delhi.

Approximately 10% of X-linked intellectual disability (XLID) in males is contributed by copy number variants and monogenic causes account for an additional ~10–12% with >90 genes implicated in both syndromic and non-syndromic forms. Rapid technological advancement has aided the identification of several ID genes, but each of these account only for a small number of families with the exception of FMR1. However, many of these ID genes with diverse molecular functions cluster in specific cellular and molecular pathways. Ubiquitin- proteosome is one such emerging pathway involved in the regulation of proteins, notably during neuronal development and remodeling of the synapse. In the present study, by targeted resequencing of the linked region we report a novel missense mutation (NM012216.3 c.1040G>A; Q9UJV3 p.Arg347Gln) in MID2, which encodes ubiquitin ligase E3, as the likely cause of X-linked intellectual disability in a large Indian kindred. The mutation was observed in all affected and obligate carriers but not in any unaffected males of the family or in the population controls (n= 200). Screening of additional idiopathic individuals with intellectual disability subjects (n=250) did not reveal this or any additional mutations in the coding exons of this gene. When transiently expressed in HEK293T cell line, the mutation was found to abolish the function of the COS domain in the protein. The GFP tagged mutant protein accumulated in the cytoplasm instead of binding to the cytoskeleton resulting in altered the sub-cellular localization. This study highlights the growing role of the ubiquitin pathway in intellectual disability. The MID2 mutation, conferring a phenotype different from that of Opitz syndrome, is of considerable interest in ID etiology. It not only contributes to the literature on functional significance of TRIM proteins but also demonstrates the wide range of phenotypic variability caused by mutations in similar domains of paralogous proteins.

2942T

Exome sequencing and functional biology reveal novel genes causing Infantile Mitochondrial Encephalopathy. *P. Bonnen, A. Besse, T. Donti, S. Lalani, F. Scaglia, W. Craigen, B. Graham.* Molecular & Human Genetics, Baylor College of Medicine, Houston, TX.

Mitochondrial disease is a diverse group of disorders that are estimated to occur with a combined incidence of 1/5,000. Approximately ninety percent of pediatric onset cases are caused by high penetrance recessive mutations in the nuclear genome, however after exhausting all available diagnostic tests most patients remain without a molecular diagnosis. Using a combination of sequencing, bioinformatics, and cellular and molecular biology we have discovered high-confidence pathogenic mutations. We identified pediatric patients who have been pre-screened and shown to be negative for mitochondrial and known nuclear gene candidates but have overwhelming evidence supporting a mitochondrial disease diagnosis. A battery of mitochondrial functional assays have been performed on all samples with available tissues including electron transport chain analysis, mitochondrial DNA copy number, FACS based analysis of relative mitochondrial membrane potential and reactive oxygen species. Exome sequencing has been completed on 20 patients. We have discovered and validated the pathogenic mutation in five patients. A causal relationship has been established using cDNA complementation and mitochondrial functional assays. This work has resulted in the discovery of novel disease genes, improved molecular diagnosis, and new insights into the pathogenetic mechanisms underlying mitochondrial encephalopathy.

2943F

***TIMM44* mutations identified by family-based WES cause severe mitochondrial respiratory chain disease due to defective mitochondrial protein import.** *M.J. Falk[1], S. Srinivasan[2], S. Dingley[1], J. Ostrovsky[1], M. Tsukikawa[1], E. Polyak[1], E. Place[1,3], M. Consugar[3], J.C. Perin[4], N. Avadhani[2], E.A. Pierce[3], X. Gai[5].* 1) Divisions of Human Genetics and Metabolic Disease, Department of Pediatrics, The Children's Hospital of Philadelphia and University of Pennslyvania Perelman School of Medicine, Philadelphia, PA; 2) Department of Animal Biology, School of Veterinary Medicine, University of Pennsylvania, PA; 3) Ocular Genomics Institute and Berman-Gund Laboratory for the Study of Retinal Degenerations, Department of Ophthalmology, Massachusetts Eye and Ear Infirmary, Harvard Medical School, Boston, MA; 4) Center for Biomedical Informatics, The Children's Hospital of Philadelphia, Philadelphia, PA; 5) Department of Molecular Pharmacology and Therapeutics, and Center for Biomedical InformaticsLoyola University Chicago Health Sciences Division, Maywood, IL.

Primary mitochondrial respiratory chain (RC) disease is a highly heterogeneous group of energy deficiency disorders, where mutations in more than 100 nuclear and all 37 mtDNA genes explain the genetic etiology for less than half of cases. To improve the diagnosis and understanding of pathogenic mechanisms in individual RC diseases, we have adopted a family-based whole exome sequencing (WES) approach. **METHODS:** We have completed WES analyses on 18 individuals from 6 kindreds, with multiple additional analyses currently ongoing. We employed a family-based WES approach using the Agilent 50 Mb “MitoExome” capture and Illumina HiSeq paired end sequencing with Sanger validation of variant segregation. **RESULTS:**Of the 6 kindreds in whom sequencing is complete, novel disease-causing recessive mutations were identified in one known gene (*RRM2B*). 1 to 2 lead candidate mutation(s) were identified for each of the other kindreds that involve novel genes not previously linked to human disease. One simplex case is a 7-year-old girl with severe multi-systemic manifestations of primary RC disease characterized by marked complex I-III deficiency, severe lactic acidosis, Leigh syndrome, GI dysmotility, and profound global developmental delay. WES analysis of the proband and her healthy parents revealed only a single gene, *TIMM44*, with biparental inheritance of rare, compound heterozygous, non-synonymous sequence variants (p.N251K and p.P308Q) predicted to be damaging by both Polyphen2 and SIFT. *TIMM44* encodes a key component of the TIM23 inner mitochondrial membrane (IMM) protein import complex. Functional analyses in fibroblast cell lines (FCLs) from the proband showed significantly impaired expression of a complex V matrix subunit that requires TIM23 complex import, as well as increased PINK1 expression, which typically occurs only by impairing its membrane potential-dependent TIM23 transport. Targeted IMM import assays in the proband's FCLs and isolated mitochondria confirmed defective mitochondrial import. Wild-type *TIMM44* cDNA rescue to ameliorate these functional abnormalities is underway. **CONCLUSIONS:** *TIMM44* is a novel cause of primary RC disease due to impaired mitochondrial protein import. A family-based WES analysis approach permits efficient determination of the precise genetic etiology for sporadic cases of mitochondrial disease, overcoming the substantial challenge of locus heterogeneity and rare individual gene contributions to disease burden.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2944W

A novel form of Limb Girdle Muscular Dystrophy caused by impairment of an ER-to-Golgi trafficking component. *N. Bögershausen[1], Y. Li[1], J.C. von Kleist-Retzow[2], R. Wirth[1], G. Nürnberg[3], H. Thiele[3], J. Altmüller[3], B. Schoser[4], P. Nürnberg[3], R. Heller[1], B. Wollnik[1].* 1) Institute of Human Genetics, University Hospital Cologne, Cologne, Germany; 2) Pediatrics Department, University Hospital Cologne, Cologne, Germany; 3) Cologne Center for Genomics, University of Cologne, Cologne, Germany; 4) Friedrich-Bauer-Institute, Ludwig-Maximilian-University Munich, Munich, Germany.

The Limb Girdle Muscular Dystrophies (LGMD) are a heterogeneous group of degenerative myopathies caused by functional alterations in the dystrophin-glycoprotein complex. Progressive proximal muscle weakness is characteristic of the disease and it may be associated with involvement of the heart and bulbar muscles, and/ or intellectual disability in a minority of cases. We examined a consanguineous Syrian family with three affected family members showing a hitherto undescribed form of LGMD. The clinical phenotype involved progressive muscle weakness of the proximal limbs, highly elevated serum creatine kinase (9–16 fold), hip dysplasia, scoliosis, and hypertrichosis in all three patients, along with eye abnormalities in two patients. There was no obvious cardiac or bulbar muscle involvement. We performed whole-exome sequencing in the index patient. There were no novel variants in any gene known to cause LGMD. Bioinformatic analysis of all known and novel exome variants for stretches of homozygosity drastically reduced the number of relevant variants. To refine the homozygous regions we performed genome wide homozygosity mapping for the whole family and found a significantly linked homozygous 1 Mb region on chromosome 4. Combination of mapping and exome data pointed to a single novel variant, a homozygous missense mutation in a gene encoding a component of a multi-subunit tethering complex that is involved in early ER-to-Golgi trafficking. Co-segregation of the mutation within the family was confirmed by Sanger sequencing. It was not found in 100 healthy Turkish controls or in any current database of human variation (including 10.000 alleles of the Exome Variant Server). We screened 32 patients with unexplained LGMD, but could not yet identify any further mutations in this gene. The mutation affects a highly conserved amino acid residue in a functional protein domain of this cargo protein and it is predicted to be probably damaging. Post translational modification of skeletal muscle proteins is important for their specific function during myogenesis and muscle regeneration, as exemplified by mutations in the Golgi resident glycosyltransferase fukutin that cause LGMD2I. We propose that the identified mutation might lead to alterations in post translational modification of muscle proteins and thereby cause a novel form of LGMD. Functional studies aiming to prove this mutational effect are in process.

2945T

Tackling giants with next-generation sequencing: homozygous or compound heterozygous truncating mutations of *TTN* from exome analysis define novel forms of cardiomyopathy with skeletal myopathy. *C. Chauveau[1, 2], C. Julien[2], H. Marks[3], R. Foley[4], A.L. Kho[5], B. Talim[6], MC. Arne-Bes[7], E. Uro-Coste[7], P. Maury[7], A. Vihola[8], B. Udd[8], H. Topaloglu[6], S. Moore[9], M. Gautel[5], C. Bonnemann[4], ME. Samuels[2], A. Ferreiro[7, 10].* 1) Groupe Myologie, U787 INSERM-UPMC paris VI, france, paris, France; 2) Centre de Recherche de l'Hôpital Ste-Justine, Université de Montréal, Montréal, Canada; 3) Section of Neurology, Department of Pediatrics, St. Christopher's Hospital for Children, Drexel University College of Medicine, Philadelphia USA; 4) National Institutes of Health, Bethesda, MD, US; 5) Cardiovascular Division and Randall Division for Cell and Molecular Biophysics, King's College London, London, United Kingdom; 6) Pediatric Pathology Unit, Child Health and Diseases Department, Faculty of Medicine, Hacettepe University, Turkey; 7) CHU Rangueil, Toulouse, France; 8) Folkhälsan Institute of Genetics and Department of Medical Genetics, Haartman Institute, University of Helsinki, Helsinki, Finland; 9) University of Iowa, Iowa City, USA; 10) Centre de Référence des maladies neuromusculaires, GH Pitié-Salpêtrière, Paris, France.

Cardiomyopathies are potentially lethal conditions in which the muscular function of the myocardium is altered. They are frequent and can be associated with skeletal muscle disease. We describe the genetic analysis focusing on the titin gene (*TTN*, 363 exons) of 24 families with primary cardiomyopathy associated with congenital core myopathy. Titin plays a key structural and functional role in striated muscles. However, *TTN*'s size and complexity preclude full screening by Sanger sequencing, and until recently only a small number of mutations had been found. Since mutations in the titin M-line domain have been associated with cardiac and skeletal muscle phenotypes, we started with Sanger sequencing of the M line-encoding region encompassed by the last 6 exons of *TTN* (6.5% of the gene). We identified 4 novel nonsense (1 homozygous, 2 compound heterozygous, 1 heterozygous) and 1 heterozygous missense mutation in 4 families (16.6% of our cohort). The 2 heterozygous changes were inherited from healthy fathers. Subsequently we performed whole exome sequencing (WES) on one sample from those last two families using the current version of the Agilent 50Mb SureSelect targeting library and successfully captured and generated sequence for essentially all *TTN* exons (except a few alternative exons not targeted by the library). In each family, WES identified one new heterozygous mutation (missense and truncating); these were confirmed as compound heterozygotes by Sanger resequencing of the patients and parents. All 8 parents were heterozygous and healthy, suggesting that not all truncating *TTN* mutations, particularly if in the M-line, manifest unless associated with a second mutation. This can potentially explain the reported high prevalence of *TTN* truncating mutations in controls (3%) from a recent study which identified truncating *TTN* mutations in 27% cases of a series of dilated cardiomyopathy and which proposed a dominant negative effect of these mutations. The range of cardiological manifestations in our cohort was large and included previously unreported phenotypes. Our results expand the spectrum of TTN-related conditions and confirm that WES opens a new era for the study of giant and complex genes, of which *TTN* is the paradigm. We also propose that *TTN* is involved in an unexpectedly high proportion and large range of striated muscle phenotypes, and suggest that it should be considered and tested in any uncharacterized cardiac and/or skeletal muscle condition.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2946F

Exome sequencing for the molecular diagnosis of muscle disorders: successes and challenges encountered. *K.K. McDonald[1], J. Stajich[1], C.P. Blach[1], A.E. Ashley-Koch[1,2], M.A. Hauser[1,2].* 1) Center for Human Genetics, Duke University, Durham, NC; 2) Department of Medicine, Duke University Medical Center, Durham, NC.

The molecular diagnosis of muscle disorders is challenging: genetic heterogeneity (>100 genes) precludes exhaustive clinical testing, and prioritizing sequencing of specific genes is difficult due to the similarity of clinical presentation. We have evaluated the usefulness of whole exome sequencing as a diagnostic approach for autosomal dominant muscular dystrophy by examining the extent of high quality sequence coverage of known muscle disease genes while searching for disease-causing variants. Exome sequencing was performed for 11 individuals from 4 families, using the Agilent SureSelect Human All Exon 50Mb kit and the Illumina HiSeq 2000 (2×100bp). Alignment was performed with BWA, and the Genome Analysis Toolkit was used for local realignment, quality score recalibration, and variant calling. After duplicate removal, the mean coverage per exome ranged from 42X to 84X, and 85.7%-91.8% of the expected capture region was covered at ≥10X. We created target intervals spanning the coding and splice regions of 102 skeletal and/or cardiac muscle disease genes, 499,520 bases total. For each exome, “failed” positions were identified that had <10 reads with QUAL scores ≥50. 4.9% of bases within the gene set failed in every exome. 11 genes (SMN1, FKRP, KBTBD13, DES, PABPN1, BIN1, COL6A1, PLEC, SEPN1, ISCU and POMT2) failed across ≥20% of their coding/splicing regions. Analysis of the exome data in one family resulted in the identification of a known nonsense mutation in filamin C, p.W2710X. If a minimum depth cut-off was applied as part of a quality filter, the mutation would not have been identified in one of the cases because only 5 reads were present at the position. In another family, we identified a known, functionally-tested, splice mutation in *desmin*, IVS3+3A>G. Interestingly, because it is not in a canonical splice site, the functional annotation fields in SeattleSeq 131 do not indicate that it is a splicing variant. Use of the “distancetosplice” field for filtering, however, did result in easy identification of the variant. In conclusion, while exome sequencing provides reliable, high-quality data for most exons in known muscle disease genes, gaps and low quality regions remain that must be evaluated on an individual basis. Additionally, filtering using distance to splice sites rather than disruption of canonical splice sites avoids potentially missing functional variants. To date, we have identified causative mutations for 2 of 4 families.

2947W

Usher-exome: an efficient diagnostic approach when used in combination with LSDB USHBases, Variant Manager USHVAM and USMA. *A.F. Roux[1, 2], T. Besnard[2, 3], G. Garcia Garcia[2, 4], D. Baux[1], C. Vaché[1], L. Larrieu[1], V. Faugère[1], J. Millan[4], M. Claustres[1, 2, 3].* 1) Molecular Genetics, CHU Montpellier, Montpellier, France; 2) INSERM U827, Montpellier, France; 3) Univ, Montpellier I, Montpellier, France; 4) Instituto de Investigación Sanitaria, Valencia, Spain.

Usher syndrome (USH) is a rare autosomal recessive disorder that associates sensorineural hearing loss (HL) and retinitis pigmentosa. Some patients will also exhibit vestibular areflexia (VA). Clinical and genetic heterogeneity is recognized as the 3 clinical subgroups, defined mainly on the degree of HL and VA, can be caused by mutations in one of the 9 known genes. If, in most cases clinical signs are in line with the defined causative gene, some patients will carry mutations in a gene responsible for a different subgroup (defined as atypical Usher). We have used so far Sanger sequencing, coupled with CGH-array to identify mutations in USH patients and the strategy was efficiently based on haplotyping followed by cascade sequencing of the candidate gene(s) according to their implication. We could show that the mutation detection rate was more than 90% but this approach remained laborious in terms of time and costs and inefficient for atypic USH patients. We have recently designed a targeted exome consisting of 330 kb of captured sequences (SeqCap EZ Choice, Nimblegen), sequenced using the GS Junior system (Roche 454). This design has been tested on 43 patients originated from France (22) and Spain (21), previously screened by Sanger in at least one USH gene (20 with 1 mutation, 23 with no mutation). This strategy was then applied to 8 patients for whom no analysis in any of the Usher genes had been performed. An average of 400 variants/patient has been identified. Successive filtering has been applied. An average of 20 variants/patient needed further analysis. In addition to central databases, studies were performed using tools dedicated to Usher syndrome (USHbases, USHVAM and USMA) in order to further assess the pathogenicity of each variant. By comparison with the variants previously identified by Sanger, we observed 3.5% of false negative among all the variants: 1% was due to poor coverage, the remaining was due to misreading in the homopolymeric regions. Ten additional mutations could be identified. The pathogenic mutations could be identified in 86% of the new USH cases, but that was possible thanks to the availability of dedicated tools and excellent knowledge on the Usher genes. The coverage was estimated to be > 40 reads for 92% of the covered regions, (minimum required for proper molecular diagnosis). Although an optimization of the design is necessary, these results are very promising for a future transfer to diagnostic facilities.

2948T

Loss of function mutations in *HINT1* are a major cause of autosomal recessive axonal neuropathy with neuromyotonia. *M. Zimon[1,2], J. Baets[2,3,4], L. Almeida-Souza[2,5], J. Nikodinovic[6], Y. Parman[7], E. Battaloglu[8], V. Guergueltcheva[9], I. Tournev[9], M. Auer-Grumbach[10], P. De Rijk[11], T. Müller[12], E. Fransen[13], P. Van Damme[14,15,16], W. Löscher[17], N. Baris{{ak}}ić[18], Z. Mitrovic[19], S. Previtali[20,21], H. Topaloglu[22], G. Bernert[23], A. Beleza-Meireles[24,25], B. Ishpekova[9], K. Peeters[1,2], A. Hahn[26], S. Züchner[27], V. Timmerman[2,5], P. Van Dijck[28], V. Milic Rasic[6,29], A. Janecke[12], P. De Jonghe[2,3,4], A. Jordanova[1,2,30].* 1) VIB Department of Molecular Genetics, VIB-University of Antwerp, Antwerpen, Antwerp, Belgium; 2) Neurogenetics laboratory, Institute Born-Bunge, University of Antwerp, Antwerpen, Belgium; 3) Neurogenetics Group, VIB Department of Molecular Genetics, University of Antwerp, Belgium; 4) Department of Neurology, Antwerp University Hospital, Antwerp, Belgium; 5) Peripheral Neuropathies Group, VIB Department of Molecular Genetics, University of Antwerp, Belgium; 6) Clinic for Neurology and Psychiatry for Children and Youth, University of Belgrade, Belgrade, Serbia; 7) Department of Neurology, Istanbul Medical Faculty, Istanbul University, Istanbul, Turkey; 8) Department of Molecular Biology and Genetics, Bogazici University, Istanbul, Turkey; 9) Department of Cognitive Science and Psychology, New Bulgarian University, Sofia, Bulgaria; 10) Department of Internal Medicine, Division of Endocrinology and Metabolism, Medical University of Graz, Graz, Austria; 11) Applied Molecular Genomics Unit, VIB Department of Molecular Genetics, University of Antwerp, Belgium; 12) Department of Pediatrics I, Innsbruck Medical University, Innsbruck, Austria; 13) Department of Medical Genetics and StatUA Center for Statistics, University of Antwerp, Belgium; 14) Experimental Neurology and Leuven Institute for Neurodegenerative disorders (LIND), University of Leuven, Leuven, Belgium; 15) Vesalius Research Center, VIB, Leuven, Belgium; 16) Department of Neurology, University Hospital Leuven, University of Leuven, Leuven, Belgium; 17) Department of Neurology, Innsbruck Medical University, Innsbruck, Austria; 18) Department of Paediatrics, University of Zagreb, Medical School, University Hospital Centre Zagreb, Zagreb,Croatia; 19) National Center for Neuromuscular Diseases, Department of Neurology, University Hospital Center Zagreb, Zagreb, Croatia; 20) Institute for Experimental Neurology, Division of Neuroscience, San Raffaele Scientific Institute, Milano, Italy; 21) Department of Neurology, San Raffaele Scientific Institute, Milano, Italy; 22) Department of Paediatric Neurology, Faculty of Medicine, Hacettepe University, Ankara, Turkey; 23) Department of Paediatrics, Gottfried von Preyer'sches Kinderspital, Vienna, Austria; 24) Autonomous Section of Health Sciences, University of Aveiro, Portugal; 25) Department of Genetics, Coimbra Paediatric Hospital, CHUC-EPE, Portugal; 26) Department of Clinical Neurological Sciences, London Health Sciences Centre, University of Western Ontario, London, Ontario, Canada; 27) Hussman Institute for Human Genetics, University of Miami Miller School of Medicine, Miami, Florida, USA; 28) VIB Department of Molecular Microbiology, Laboratory of Molecular Cell biology, University of Leuven, Leuven, Belgium; 29) Faculty of Medicine, University of Belgrade, Belgrade, Serbia; 30) Department of Medical Chemistry and Biochemistry, Molecular Medicine Center, Medical University-Sofia, Sofia, Bulgaria.

Inherited Peripheral Neuropathies are a group of genetically and clinically heterogeneous disorders affecting 1 in 2500 individuals. Although more than 50 genes were associated with IPN, number of families still remains without genetic diagnosis pointing to novel genetic entities. Until recently, small family sizes hampered gene discovery however current technological advances present new opportunities for unraveling CMT molecular architecture.

In this study, we show that combined SNP genotyping, linkage analysis and whole genome sequencing were successful in discovering genetic cause in a small Belgian family diagnosed with autosomal recessive (AR) axonal neuropathy and neuromyotonia. After data filtering within the three linked regions, we identified two heterozygous mutations in *HINT1* (*histidine triad nucleotide binding protein 1*) - gene located on chromosome 5. Subsequent patient cohort screening allowed finding of 8 different mutations in a total of 32 families. Interestingly, the majority of patients carrying *HINT1* mutations were diagnosed with a motor greater than sensory axonal neuropathy associated with action myotonia in the hands and/or neuromyotonic and myokymic discharges on needle electromyography. Studies in yeast cells and patients lymphoblasts showed that mutations lead to either enzymatically inactive protein or to its post-translational degradation.

In conclusion, we present here combined genetic and functional evidences that the maintenance of functional HINT1 is crucial for the peripheral nervous system (PNS) physiology. Frequency of mutations in a cohort of patients diagnosed with autosomal recessive neuropathy was 11 % however in a specific group of patients with AR axonal neuropathy and neuromyotonia it raised to 76 %. Our findings are also of great importance for future molecular diagnostics and patient counseling.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2949F

Identification of novel genes in human primary immunodeficiency diseases using exome sequencing. *S. Khan[1], B. Wakeland[1], C. Liang[1], M. De la Morena[2], N. Van Oers[1], E. Wakeland[1].* 1) University of Texas Southwwestern Medical Center, Dallas, TX; 2) Children's Medical center, Dallas, TX.

Severe immunodeficiencies are rare in the human population (>1/70000) and their genetic etiology is very difficult to decipher. More than 150 genes have been linked to various forms of primary immunodeficiencies (PID) and most of these are the consequence of a monogenic defect and hence follow a simple mendelian inheritance. In this study we utilized whole exome sequencing strategy to identify the causative genetic lesions in patients with severe immunodeficiencies of unknown etiology. This strategy has proven to be a powerful, efficient strategy for identifying genes underlying rare mendelian disorder. We subjected five families with one or more children who have severe PID with an unknown genetic basis to whole exome sequencing using SureSelect Human All Exon Kit from Agilent technologies and sequenced in a single lane of an Illumina GAIIx. On average 4.5 Gbp of sequence data was generated per sample. Raw sequence data in fastq format were converted to BAM file format and aligned using the Genome Analysis Toolkit (GATK) which is a structured programming framework developed by the Broad Institute. GATK consists of robust analysis tools including depth of coverage analyzers, a quality score recalibrator, a SNP/indel caller and a local realigner. The data was further analyzed using Variant Call Format (VCF) tools package. On average individual genomes differed from the reference sequence by 67089 total SNVs. To distinguish potentially deleterious mutations from other variants, only nonsynonymous (NS) variants and short coding region insertions or deletions (indels; I) were considered. We will present the analysis for all the five families at ASHG 2012. The causative variants for the PID exhibited by the proband was sought using several filtering strategies that led to a list of candidate variations that were novel, deleterious mutations in affected child within a PID family. These variants are further processed through the PolyPhen 2 Gateway and SIFT to predict their potential effects on gene function and to generate a list of variants with potentially damaging SNVs. These mutations are now being validated using standard Sanger sequencing. The future studies will characterize the impact of these variants on their expression by RNA-SEQ analysis of PBMC from the parents and/or siblings that are heterozygous for the candidate variant. We hope that these data will identify novel genes and their functional role within the human immune system.

2950W

The ADAMTS18 gene is responsible for autosomal recessive syndromic retinal dystrophy. *S. Banfi[1, 5], I. Peluso[1], F. Testa[2], M. Pizzo[1], R. Collin[3], N. Meola[1], M. Mutarelli[1], G. Dharmalingam[1], M. Melone[4], I. Conte[1], F. Simonelli[2].* 1) TIGEM, Fondazione Telethon, Naples, Italy; 2) Ophthalmology, Second University of Naples, Naples, Italy; 3) Nijmegen Centre for Molecular Life Sciences, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands; 4) Clinical and Experimental Medicine and Surgery, Second University of Naples, Naples, Italy; 5) 5Medical Genetics, Department of General Pathology, Second University of Naples, Naples, Italy.

Inherited retinal dystrophies are a group of genetically heterogeneous disorders that lead to severe visual deficits. They can be caused by mutations in over 200 genes and there is evidence for the presence of as yet unidentified genes in a significant proportion of patients. By using homozygosity mapping and whole exome sequencing, we identified, in an Italian ARRD patient, a homozygous missense mutation in the ADAMTS18 gene, which was recently proposed to be linked to Knobloch syndrome, a rare developmental disorder that affects the eye and the occipital skull. In vivo gene knockdown and rescue assays performed in the medaka fish (Oryzias latipes) model organism confirmed the pathogenic role of the mutation. This study reveals that mutations in the ADAMTS18 gene can cause a broad phenotypic spectrum of eye disorders and contributes to shed further light on the complexity of ARRD.

2951T

Novel *PLP1* gene mutation discovered by whole genome sequencing in brothers with infantile onset dopa-responsive dystonia and delayed central nervous system demyelination. *R.L. Margraf[1], J. Durtschi[1], K. Mallempati[1], J. Bonkowsky[3], R. Lutz[4], K.V. Voelkerding[1,2], K.J. Swoboda[5].* 1) ARUP Institute for Clinical and Experimental Pathology, Salt Lake City, UT; 2) Department of Pathology, University of Utah School of Medicine, Salt Lake City, Utah; 3) Department of Pediatrics, University of Utah School of Medicine, Salt Lake City, Utah; 4) Munroe-Meyer Institute for Genetics & Rehabilitation, University of Nebraska Medical Center, Omaha, Nebraska; 5) Departments of Pediatrics and Neurology, University of Utah School of Medicine, Salt Lake City, Utah.

Next-generation sequencing has provided a paradigm shift in our ability to identify causal variants in undiagnosed neurogenetic diseases and facilitated expansion of phenotypes associated with previously well-described neurologic disorders in which an atypical phenotype results in an expensive diagnostic odyssey. The male proband in this family presented in early infancy with generalized dystonia and normal brain MRI imaging in the first two years of life, and he was diagnosed with cerebral palsy. When his brother presented with a similar phenotype, further diagnostic workup revealed CSF neurotransmitter and biopterin deficiency thought to be consistent with a diagnosis of GTPCH1 deficiency, and dystonic symptoms were partially dopa-responsive. Several years later the diagnosis was revisited following sudden death of the proband in the setting of progressive spasticity and opisthotonic posturing. Repeat MRI neuroimaging in the younger brother revealed a previously undiagnosed leukodystrophy. To elucidate the genetic basis of these neurological symptoms, whole genome sequencing of family members was performed, including both parents, the two affected boys and an unaffected sister. Candidate gene variant discovery analysis involved heuristic filtering and probabilistic approaches facilitated by Golden Helix and VAAST software analyses, respectively, revealing an X-linked mutation in the *PLP1* gene (proteolipid protein 1 gene, c.617T>A, p.M206K) in the two affected male siblings and the mutation was heterozygous in their carrier mother, which was Sanger confirmed. This mutation was not found in the NCBI dbSNP135, 1000 Genomes, or HGMD databases and appears novel. The PLP1 protein is the predominant myelin protein present in the CNS, playing a role in the stabilization and maintenance of myelin sheaths and mutations in *PLP1* have been reported in the X-linked dysmyelinating disorders, Pelizaeus-Merzbacher disease (PMD) and spastic paraplegia type 2. PMD mutations have been reported that flank either side of the p.M206K variant found in our patients. While mutations in *PLP1* have been reported to cause both dystonia and leukodystrophy, the lack of early CNS dysmyelinating/demyelinating features and CSF studies indicating a possible diagnosis of dopa-responsive dystonia contributed to a substantial delay in diagnosis. This report emphasizes the value of whole genome sequencing for causal variant discovery in children with undiagnosed neurologic disorders.

2952F

Digenic inheritance in Autosomal Recessive Non-syndromic Hearing Loss cases carrying GJB2 heterozygote mutations: assessment of GJB4 and GJA1. *D. Kooshavar[1,2], M.R. Noori Daloii[1], M. Hashemzadeh Chaleshtori[2].* 1) Department of Medical Genetics, Faculty of medicine, Tehran University of Medical Science, Tehran, Iran; 2) Cellular and Molecular Research Center, Shahrekord University of Medical Science, Shahrekord, Iran.

Autosomal Recessive Non-syndromic Hearing Loss (ARNSHL), in up to 50 percent of the cases, is caused by mutations in the GJB2 gene which encodes the gap-junction (GJ) protein connexin (Cx) 26. Nevertheless 10 to 42 percent of patients with recessive GJB2 mutations are heterozygous carriers. Mutations of GJB4 and GJA1 encoding Cx30.3 and Cx43 respectively, can lead to hearing loss. Different connexins can integrate in heteromeric and heterotypic GJ assemblies. This study aims to determine whether variations in any of the genes GJB4 or GJA1 can be the second mutant allele in digenic mode of inheritance in the GJB2 heterozygous cases studied. We examined 44 unrelated GJB2 heterozygous ARNSHL subjects and 400 normal hearing individuals from different geographic and ethnic areas in Iran, for sequence variations in the two genes, using polymerase chain reaction followed by direct sequencing also Restriction Fragment Length Polymorphism (RFLP). Sequence analysis of GJB4 showed five heterozygous variations (c.451C>A, c.219C>T, c.507C>G, c.155_158delTCTG, c.542C>T) that the last one was not detected in any of control samples. The altered amino acid was highly conserved. There were three heterozygous substitutions (c.758C>T, c.717G>A, c.3*dupA) in GJA1 in five cases. We propose that GJB4 c.542C>T variation and with lower possibility other variations of GJB4 and GJA1, can be assigned to ARNSHL in a digenic pattern in GJB2 heterozygote mutation carriers. Further functional studies should be done to definitely enlighten the role of these variations in digenic form of ARNSHL.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2953W

Exocrine and endocrine pancreatic damage in cystic fibrosis are associated with SLC26A9. *D. Soave[1], T. Chiang[2], M. Miller[2], D. Su[2], K. Keenan[2], W. Li[1,2], W. Ip[2], F. Wright[3], S. Blackman[4], H. Corvol[5,6], M. Knowles[7], G. Cutting[4,8], M. Drumm[9], L. Sun[1], J. Rommens[2,10], P. Durie[2,11], L. Strug[1,2].* 1) Biostatistics, University of Toronto, Toronto, Canada; 2) Research Institute, The Hospital for Sick Children, Toronto, On, Canada; 3) Department of Biostatistics, University of North Carolina, Chapel Hill, North Carolina, USA; 4) Department of Pediatrics, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA; 5) AP-HP, Hopital Trousseau, Pediatric Pulmonary Department, Inserm U938, Paris, France; 6) Pierre et Marie Curie University-Paris 6, Paris, France; 7) Cystic Fibrosis-Pulmonary Research and Treatment Center, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA; 8) McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA; 9) Departments of Pediatrics and Genetics, Case Western Reserve University, Cleveland, Ohio, USA; 10) Department of Molecular Genetics, Univ. of Toronto, Toronto, On, Canada; 11) Department of Pediatrics, Univ. of Toronto, Toronto, On, Canada.

Degree of lung and exocrine pancreatic damage (PD), and the occurrence of co-morbidities such as intestinal obstruction at birth (MI), varies in Cystic Fibrosis (CF) patients with the same CFTR gene mutations. This suggests involvement of other genes. In determining modifying genes related to early CF morbidity, we identified SLC26A9, SLC6A14 and SLC9A3 as associated with MI; SLC26A9 has also been reported as associated with CF-related diabetes (CFRD). Serum immunoreactive cationic trypsinogen concentration (IRT) acts as a marker of exocrine PD in CF patients. Using longitudinal IRT measurements we asked whether early PD is associated with the three MI modifiers, and whether early PD can predict CFRD later in life. CF patients enrolled in the Canadian CF Gene Modifier Study and in earlier studies of PD (Durie et al. 1986) were included. In 129 patients, there were 1 to 16 measurements per person, up to 23 years of age, with one measurement within 24 months following birth. CFRD was diagnosed in 37 of 129 patients. Genetic associations with IRT at birth and rate of decline were determined for 8 SNPs in the three MI genes. We used linear mixed effects (LME) modeling accounting for limits of detection of the IRT assay (less than 3 ng/mL). IRT and age were natural log transformed, each SNP and its interaction with age were then added to the model, adjusting for MI. To determine whether risk of CFRD was associated with IRT, similar LME modeling was used to estimate subject-specific IRT measurements at birth and rate of decline and these estimates were assessed as predictors of CFRD in a Cox proportional hazards model. Three of the four SNPs in the SLC26A9 gene were associated with IRT at birth (min p=0.003 at rs7512462; each C allele results in a 0.96 increase in log(IRT)) and its rate of decline (p=0.008). SNPs in SLC6A14 or SLC9A3 had p>0.05. Estimated IRT at birth, but not rate of decline, was associated with CFRD (p=0.004). A 46.7% decrease in the risk of CFRD is estimated for every one unit increase in log(IRT) at birth (log(IRT) interquartile range: 2.28 to 4.38). SLC26A9 is associated with MI, CFRD and IRT at birth and its rate of decline, suggesting potential for a common therapeutic intervention. Estimated PD at birth is a strong predictor of later-onset CFRD, suggesting the possibility of predicting CFRD risk at birth. These findings require confirmation in a second cohort and estimation of CFRD predictive models with a newborn-screened sample.

2954T

Clinical findings of a three-generation family with GLA nonsense mutation (W162X). *E. Severin[1], A. Stan[2], C. Dragomir[2], G.J. Sarca[3].* 1) Carol Davila University of Medicine and Pharmacy, Bucharest, Romania; 2) Genetic Lab, Bucharest, romania; 3) Prof.Dr. C. Angelescu Hospital, Bucharest, Romania.

Background: The deficiency of α-galactosidase A, a lysosomal enzyme, is caused by different mutations in the GLA gene located on Xq22. Complete or partial deficiency of this enzyme results in the accumulation of globotriaosylceramide in cells throughout the body, leading to Fabry disease phenotype. Objectives: To describe and compare clinical findings, to confirm the diagnosis in male and female family members and to identify the mutation causing Fabry disease. Materials and Methods: 7 family members ranging in age from 8 to 56 years, 2 males and 5 females, were clinically assessed and also evaluated for enzyme activity. Genomic DNA was isolated from blood samples of all family members and analyzed for GLA gene mutation. Results: The proband, a 31-year-old male, first experienced symptoms onset at 7 years. He had a less severe phenotype with no cardiac or renal involvement. His maternal uncle had a severe classic form of Fabry disease (left ventricular hypertrophy, cerebrovascular manifestations, hearing loss, renal insufficiency). Both were found to have one copy of GLA mutation (c.485G>A). All five females were carrier with milder symptoms (acroparaesthesiae, pain or angiokeratomas). Conclusions: Fabry disease showed a wide range of clinical manifestations in this family. According to the results genetic testing should be considered especially for women and families in whom the diagnosis of Fabry disease is suspected based on early clinical signs and symptoms.

2955F

Targeted array-CGH analysis to identify copy-number changes underlying ciliopathies. *A. Lindstrand[1], C. Carvalho[2], D. Pehlivan[2], R.D. Clark[3], C.A. Johnson[4], H. Omran[5], B. Franco[6], H. Kremer[7], P.L. Beales[8], P.J. Scambler[8], E.R. Maher[9], J.R. Lupski[2,10], N. Katsanis[1].* 1) Center for Human Disease Mod., Duke University Medical Center, Durham, NC; 2) Department of Molecular and Human Genetics, Baylor College of Medicine, Houston TX, USA; 3) Division of Medical Genetics, Department of Pediatrics, Loma Linda University School of Medicine, Loma Linda, California; 4) Section of Ophthalmology and Neurosciences, Leeds Institute of Molecular Medicine, St James's University Hospital, Leeds, UK; 5) Department of Pediatrics and Adolescent Medicine, University Hospital Muenster, Muenster, Germany; 6) Telethon Institute of Genetics and Medicine (TIGEM), Napoli, Italy; 7) Departments of Otorhinolaryngology and Human Genetics, Radboud University Medical Centre, Nijmegen, The Netherlands; 8) Molecular Medicine Unit, University College London Institute of Child Health, London, UK; 9) Department of Medical and Molecular Genetics, Institute of Biomedical Research, University of Birmingham, Birmingham, UK; 10) Texas Children's Hospital, Houston TX, USA.

Disruption of ciliary and basal body function has been associated with a growing number of human genetic disorders, collectively termed ciliopathies. Although clinically distinct, ciliopathies manifest similar hallmark features including retinal degeneration, renal and pancreatic cysts, liver fibrosis, polydactyly and other skeletal abnormalities, situs inversus, and central and peripheral nervous system defects. Despite extensive sequencing efforts, 20–50% of the mutational basis for ciliary disorders remains unknown. In addition, genetic studies have indicated that ciliopathy genes can contribute both causal and modifying alleles across the ciliopathy spectrum, supporting the hypothesis that the mutational load in the ciliary proteome plays a critical role in defining disease penetrance and expressivity. To establish the genetic contribution of copy number variations (CNVs) throughout a spectrum ciliopathies we have developed a custom designed oligonucleotide array-CGH (aCGH) targeting the coding exons of 772 genes of interest. A total of 401 DNA samples from ciliopathy patients were evaluated using this targeted aCGH: Bardet-Biedl syndrome (BBS) = 92; Jeune syndrome (JS) = 52; Joubert syndrome (JBTS) = 4; Meckel-Gruber syndrome (MKS) = 87; Primary Ciliary Dyskinesia (PCD) = 83; Oral-facial-digital type I syndrome (OFD1) = 48; Usher syndrome (USH) =27; unknown ciliopathy = 8. We discovered one or more distinct CNVs in 43 (10.7%) unrelated individuals, including 33 deletions and 3 duplications in known ciliopathy associated genes. Interestingly, we detected an accumulation of the recurrent NPHP1 deletion in our BBS, OFD1 and PCD cohorts and 5 (18.5%) unrelated families with Usher syndrome harbored an identical 182 kb deletion of USH2A. In addition, eight contiguous gene deletions and duplications were detected which are likely to be deleterious. Deletion breakpoints have been confirmed in 16 cases by PCR and sequencing. None of the aberrations were detected in 229 control individuals. Our data suggest a major contribution of CNVs in ciliopathy disease burden. In addition, we found new candidate ciliopathy loci that are presently undergoing in vivo and in vitro validations.

2956W

Is the R202Q mutation clinically relevant in familial mediterranean fever, though? *S. Duzenli, A.O. Arslan.* Medical Genetics, Abant Izzet Baysal University, BOLU, BOLU, Turkey.

Familial Mediterranean Fever (FMF) is an autosomal recessive disorder characterized by recurring attacks of fever and serositis, affecting Arabs, Armenians, Jews and Turks. Up to date, determined disease causing mutations on the marenostrin/pyrin gene are 222, of which about 100 are said to cause clinical symptoms including amyloidosis. Most seen disease causing mutations in the Turkish population are M694V, V726A, A744S, E148Q, M694I, M680I, F497L and R202Q. We aimed to evaluate the phenotype-genotype correlation of MEFV gene mutation R202Q of exon 2, of which clinical effects are still debatable. Five hundred and eleven patients of Turkish origin with FMF symptoms, aged between one month and 65 years were screened for the eight mutations mentioned above. The molecular genetic testing was undertaken by using appropriate primers and probes on a Real-Time PCR system, after extraction of DNA from whole blood. Of the 511 clinically significant patients, 332 carried at least one of the mutations. Ninety nine patients were only heterozygote R202Q carriers. Compound heterozygosity with R202Q and another mutation was detected in 138 patients. In total, of the 511 patients, 332 had mutations on the MEFV gene, 32 were homozygous for R202Q and 138 had at least one R202Q mutation. According to Naimushin et al. the impact on the pyrin-protein structure of R202Q mutation is slightly higher than the E148Q mutation. It is further as good as accepted, that the E148Q mutation is associated with FMF symptoms. Accordingly, one could expect that both have a close-rangely effect on clinical outcome. When taken into account that MEFV mutation distribution and effect vary considerabely high among populations, we can conclude that the effect of the R202Q mutation on the phenotype of Turkish FMF patients can not be ignored. This rather should be discussed and studied further, to determine whether it is clinically relevant for the populations affected.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2957T

Molecular basis of microphthalmia/anophthalmia. *N. Chassaing[1], A. Causse[2], A. Delahaye[3], A. Vigouroux[1], P. Calvas[1].* 1) Service de Génétique Médicale, Hôpital Purpan, Université Toulouse III, Toulouse, France; 2) Service d'Ophtalmologie, Hôpital Purpan, Université Toulouse III, Toulouse, France; 3) Service d'Histologie, Embryologie, et Cytogénétique, AP-HP, Hôpital Jean Verdier, Bondy, France.

PURPOSE: Microphthalmia/Anophthalmia (M/A) are the most severe ocular developmental defects corresponding to a reduce size or an absence of the ocular globe. Incidence of M/A is estimated to be 1/10.000. Numerous genes have been implicated either in non syndromic and syndromic forms of M/A. The aim of this study was to evaluate respective responsibility of participation of the 7 major genes involved in M/A in human. METHODS: 135 patients suffering from M/A were included in this study. Among these patients, 35 displayed anophthalmia, 47 microphthalmia, and 53 colobomatous microphthalmia. The ocular involvement was isolated in 76 patients, and associated with other malformations in 59 patients. Most cases (110/135) were sporadic. Mutations screening was conducted in *SOX2*, *PAX6*, *OTX2*, *RAX*, *FOXE3*, *VSX2*, and *GDF6* entire coding regions and intron-exon borders by direct sequencing. Furthermore, we developed a semi-quantitative approach (QMPSF) in order to detect deletion or duplication events for *SOX2, PAX6*, *OTX2*, *RAX*, and *VSX2*. Mutated patients' phenotypes were examined in order to obtain a better delineation of each gene phenotypic spectrum. RESULTS: Molecular analysis allows identification of the causative mutation in 32/135 patients (24%). Mutations in*SOX2* represent the most common cause of M/A (18/135, 13%). *SOX2* mutations represent more than a half of the total of identified mutations (18/32, 56%) in this cohort. Interestingly, whole gene deletions account for 27% (5/18) of the identified *SOX2* mutations. *OTX2* and *RAX* are mutated in 4% (5/135), and 3% (4/135) of patients respectively. *FOXE3*, *PAX6*, *VSX2* and *GDF6*are only rarely involved. Most mutations were identified among anophthalmic patients (21/32, 65%). The mutation detection rate is of 24% in this cohort, but increases in the subgroup of anophthalmic patients (21/35, 60%). CONCLUSION: In this the large, molecular screening allows identification of the molecular defects in one quarter (32/135) of patients. We provide a rough evaluation of the frequency of these genes involvement in M/A and recognized some unusual phenotypes. Finally, these results underline the genetic heterogeneity of M/A and as the majority of patients do not harbour mutations in known M/A genes other genes remain to be.

2958F

Identification and functional investigation of non-coding NIPBL regulatory elements. *D. Braunholz[1], J. Wilde[4], L.D. Michelis[4], K.S. Wendt[2], E. Watrin[3], J. Eckhold[1], I.D. Krantz[4], G. Gillessen-Kaesbach[1], M.A. Deardorff[4], F.J. Kaiser[1].* 1) Institut für Humangenetik, Lübeck, Germany; 2) Dept. of Cell Biology Erasmus MC Dr. Molewaterplein 50 3015 GE Rotterdam, Netherlands; 3) Institut de Génétique et Développement de Rennes, CNRS Faculté de Médecine 2 avenue du Pr Bernard, 35043, Rennes, France; 4) Children`s Hospital of Philadelphia, Division of Human Genetics, Philadelphia, United States.

Cornelia de Lange Syndrome (CdLS) is a rare genetic developmental disorder with a variable severity of phenotype. Patients demonstrate morphological anomalies that include characteristic facial features, a range of malformations of the upper extremities, growth retardation and cognitive delays. Nearly 60% of classical CdLS patients carry mutations in NIPBL, compared with only 10% of mutations found in four other identified genes. Current molecular diagnostics, including whole exome sequencing, are restricted to the coding region of NIPBL. It is known that even small changes in NIPBL transcript levels have pathogenic effects in animal model systems, suggesting that mutations in NIPBL regulatory elements could also cause CdLS. To investigate this, we used in silico promoter prediction tools followed by reporter assays to identify a 2 kb-region that includes the first exon and 5' region of the first intron as the NIPBL promoter. Furthermore, we narrowed the core promoter to a 143 bp region upstream of exon 1. Sequencing analysis of the 2 kb NIPBL promoter in 110 mutation-negative CdLS patients identified two novel base substitutions. In vitro reporter gene assays and quantitative RT-PCR of blood samples or cell lines demonstrated that these mutations cause a significant reduction of the NIPBL transcript levels. In silico analysis also suggested a putative regulatory element within intron 9 that shows strong enrichment of the histone 3 lysine 4 monomethylation in various cell types, signifying a regulatory element. Reporter gene analysis demonstrates strong repression activity of this element toward the NIPBL promoter. As next step we will perform sequencing analysis of this regulatory element in a large cohort of patients with CdLS and related phenotypes as well as further molecular investigations, such as chromatin conformation capturing (3C), to examine the functional interaction of this cis-regulatory element with the NIPBL promoter. In summary, we identified and functionally characterized two regulatory elements within non-coding regions of NIPBL that influence NIPBL gene expression. Therefore, mutations within these elements may alter NIPBL-protein level to cause CdLS or related phenotypes.

2959W

Molecular genetic studies of nonsyndromic oculocutaneous albinism in the Pakistani population. *T. Jaworek[1], T. Kausar[1,2], N. Tariq[2], S. Sadia[2], M.I. Maqsood[2], A. Sohail[2], M.A. Bhatti[2], M. Ali[2], S. Riazuddin[1], R.S. Shaikh[2], Z.M. Ahmed[1,2].* 1) Cincinnati Children's Hospital Research Foundation, Cincinnati, OH. USA; 2) Institute of Biotechnology, Bahauddin Zakariya University, Multan, Pakistan.

Oculocutaneous albinism (OCA) is caused by a group of genetically heterogeneous inherited defects that result in the loss of pigmentation in the eyes, skin and hair. Mutations in the TYR, OCA2, TYRP1 and SLC45A2 genes have been shown to cause isolated OCA. No comprehensive analysis has been conducted to study the spectrum of OCA alleles prevailing in the Pakistani albino populations. We enrolled 75 large Pakistani families and screened them for OCA genes and a candidate gene, SLC24A5. Screening of the TYR gene revealed four known and three novel mutations in ten families. Ex vivo studies revealed the retention of an EGFP-tagged mutant tyrosinase in the endoplasmic reticulum (ER) at 37°C, but a significant fraction of two mutant proteins left the ER in cells grown at a permissive temperature (31°C). Three novel and two known mutations of OCA2 were found in fourteen families. Sequence analysis of the SLC45A2 gene revealed three novel probable pathogenic variants in three families. No mutation in TYRP1 and SLC24A5 was found in the remaining families, suggesting there may be an additional gene in which mutant alleles cause nonsyndromic OCA. Therefore, we performed a genome wide linkage analysis and mapped a novel locus for nonsyndromic OCA. Based on our results, it is tempting to speculate that OCA2 mutations are the most common cause of nonsyndromic OCA in the Pakistani population. We also report the clinical phenotypes of multiple affected individuals from each family. Overall, our study contributes to the development of genetic testing protocols and genetic counseling for OCA in Pakistani families. Details of these novel mutant alleles and a new OCA locus will be presented at the meeting.

2960T

A Case of Lymphedema with Microcephaly and Chorioretinopathy with a Mutation in the KIF11 Gene. *D. Finegold[1,2], E. Lawrence[1], K. Levine[1], R. Ferrell[1].* 1) Dept of Human Genetics University of Pittsburgh; 2) Childrens Hospital of Pittsburgh.

Ostergaard et al 2012 found KIF11 mutations in families with autosomal-dominant microcephaly associated with congenital lymphedema and chorioretinopathy. Subsequently, we carried out Sanger sequencing for all 22 coding exons in KIF11 in a set of 187 probands ascertained on primary lymphedema, identifying a heterozygous deletion in exon 20 in one family (Arg953fsX1012). This 17 bp deletion leads to a nonsense protein of 60 amino acids prior to a stop codon altering and truncating the normal protein. The family has been previously described as having Microcephaly-lymphedema-chorioretinal dysplasia syndrome, MLCRD OMIM #152950 (Crowe and Dickerman 1986 and Limwongse et al 1999). The proband also was diagnosed with ASD (atrial septal defect) at the age of 14. ASD was also noted in two of the patients in the Ostergaard et al paper. Potentially ASD could be part of the overall syndrome but with a lower penatrance. The mother also carried the heterozygous deletion and was affected with microcephaly and chorioretinopathy and mild lyphedema. This observation further confirms the findings of Ostergaard et al, expanding the molecular spectrum of mutation in KIF11.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2961F

miRNA-411 negatively regulates YAF2 and myogenic factors. *Y. Chen[1,2], N. Harafuji[1], P. Schneiderat[3], M.C. Walter[4].* 1) Center for Genetic Medicine Research, Children's National Medical Center, Washington DC; 2) Department of Integrative Systems Biology and Department of Pediatrics, George Washington University, Washington DC; 3) Department of Neurology, Ludwig-Maximilians-University of Munich, Germany; 4) Friedrich-Baur-Institute, Department of Neurology, Ludwig-Maximilians-University of Munich, Germany.

Facioscapulohumeral muscular dystrophy (FSHD) is an autosomal dominant muscular dystrophy caused by de-repression of double homeobox protein 4 due to a contraction of the D4Z4 array at chromosome 4q35. Several studies suggest that defects in myogenesis are involved in the disease mechanisms. In this study, we hypothesized that misregulation of microRNAs (miRNAs) in FSHD is involved in the myogenic defects and disease progression. To identify misregulated miRNAs in FSHD myoblasts, miRNA expression profiling was performed using TaqMan microRNA arrays. Twenty-six and 11 miRNAs were differentially expressed in FSHD myoblasts during proliferating and 48 hours post-differentiation, respectively. The differential expression of miR-411 was validated by individual quantitative RT-PCR assay and additional arrays. In situ hybridization showed cytoplasmic localization of miR-411 in the myoblasts. The mRNA expression changes of the myoblasts were determined by mRNA expression profiling, and 4 potential miRNA-411 targets including YY1 associated factor 2 (YAF2) were identified. Over-expression of miR-411 in murine myoblasts, C2C12 cells, showed a reduction of YAF2 mRNA expression and myogenic factors myoD and myogenin. The findings suggest that the higher expression of miR-411 in FSHD primary myoblasts might be involved in the defects of myogenesis reported in FSHD.

2962W

Transcriptional profiling in facioscapulohumeral muscular dystrophy to identify candidate biomarkers. *F. Rahimov[1,2], O.D. King[3], D.G. Leung[4,5], G.M. Bibat[4,5], C.P. Emerson Jr.[3], K.R. Wagner[4,5], L.M. Kunkel[1,2].* 1) Genetics, Children's Hospital Boston, Boston, MA; 2) Harvard Medical School, Boston, MA; 3) Boston Biomedical Research Institute, Watertown, MA; 4) The Kennedy Krieger Institute, Baltimore, MD; 5) Neurology and Neuroscience, The Johns Hopkins School of Medicine, Baltimore, MD.

FSHD is a progressive neuromuscular disorder caused by contractions of repetitive elements within the macrosatellite D4Z4 on chromosome 4q35. The pathophysiology of FSHD is unknown and as a result, there is currently no effective treatment available for this disease. In order to better understand the pathophysiology of FSHD and develop mRNA-based biomarkers of affected muscles, we compared global analysis of gene expression in two distinct muscles obtained from a large number of FSHD subjects and their unaffected first-degree relatives. Gene expression in two muscle types was analyzed using GeneChip Gene 1.0 ST arrays: biceps, which typically shows an early and severe disease involvement, and deltoid, which is relatively uninvolved. For both muscle types, the expression differences were mild: using relaxed cutoffs for differential expression (fold-change $\geq$1.2 and nominal *P*-value <0.01), we identified 191 and 110 genes differentially expressed between affected and control samples of biceps and deltoid muscle tissues, respectively, with 29 genes in common. Controlling for a false-discovery rate <0.25 reduced the number of differentially expressed genes in biceps to 188 and in deltoid to 7. Expression levels of 15 genes altered in this study were used as a “molecular signature” in a validation study of an additional 26 subjects, and predicted them as FSHD or control with 90%; accuracy based on biceps and 80% accuracy based on deltoids. Genes and molecular pathways shown to be disrupted in FSHD should elucidate the underlying mechanisms of FSHD pathogenesis and facilitate development of novel therapies.

2963T

The second most prevalent mutation of *fukutin* in Japanese Fukuyama muscular dystrophy. *K. Kobayashi[1], R. Kato[1], E. Kondo[2], M. Osawa[3], K. Saito[2], T. Toda[1].* 1) Div Neurology/Molecular Brain Science, Kobe Univ Grad Sch Medicine, Kobe, Japan; 2) Inst Medical Genetics, Tokyo Women's Medical Univ, Tokyo, Japan; 3) Dept Pediatrics, Tokyo Women's Medical Univ, Tokyo, Japan.

Fukuyama muscular dystrophy (FCMD), one of the most common autosomal recessive disorders in Japan, is characterized by congenital muscular dystrophy combined with neuronal migration disorder and ocular abnormalities. Most of the Japanese FCMD patients are homozygous for a founder haplotype that carries the SVA-type retrotransposon insertion mutation in the *fukutin* gene causing abnormal splicing, whereas some other patients are compound heterozygous for the founder insertion and various kinds of less common mutations. However, no mutation that corresponds to the second most prevalent haplotype has been found so far. Here we show that this haplotype carries a deep intronic mutation (c.647+2084G>T), which is the same as the mutation found in Korean patients. It activates a pseudo-exon between exon 5 and 6, leads to frameshift and premature termination of fukutin synthesis, and causes hypoglycosylation of α-dystroglycan in the patients. This mutation seen in Japanese and Korean patients might have been derived from the same ancestor.

2964F

Molecular study of KRT14 mutations for making a screening panel test for Iranian patient affected with Epidermolysis Bullosa. *M. Mahdavi[1], M. Moghaddam[1], A. Kakavand[1], N. Hatamnejadian[1], B. Sedaghati khayyat[1], H. Dehghanpour[1], A. Ebrahimi[1,2].* 1) Parseh Medical Genetics Counsell, Tehran, Iran; 2) Shiraz medical science university.

Epidermolysis bullosa (EB) is a rare, heterogeneous, trauma-induced blistering disease. The identified genes include those that encode keratins 5 and 14 in epidermolysis bullosa simplex, collagen VII in dystrophic epidermolysis bullosa, and laminin 5 in Herlitz junctional epidermolysis bullosa.The subtypes have been classified according to skin morphology.The Simplex form is more common variation which has four subtypes.Although the Skin Biopsy is the first step in EB Testing strategy, direct molecular methods is needed to confirm and localize the mutations.We aimed to evaluate the KRT14 hot spot exons as a noninvasive, rapid and cost benefit panel test for first step screening of common mutations in EB affected patients.Methods: About 90 patients clinically diagnosed as EB, were collected and classified clinically. All DNA samples sequenced directly for selected exons of KRT14 .Result: about 30%of affected probands have mutation in KRT14 selected exons that most of the mutation were C to G convertion that changes Ser to Thr and CAG(Glutamin) deletion in first exon.These mutations were novel.KRT14 selected exons beside KRT5 and COL7A1 hot spots could show about 50% of EB types as a screening panel.Key words: EB, KRT14, Sequencing, KRT5, COL7A1,Hot spot.

2965W

Mutations in ANO5 in the Norwegian Limb Girdle Muscle Dystrophy population. *M. Van Ghelue[1,2], M. Ingebrigtsen[1], A. Skogstad[1], B. Hestholm[1], C. Wahl[3], S.I. Mellgren[4,5], F. Rasmussen[6], M. Rasmussen[7], B. Karimé[6], C. Jonsrud[1].* 1) Department of Medical Genetics, Division of Child, University Hospital of North Norway, Tromso, Norway; 2) Institute of Clinical Medicine, University of Tromsoe, Norway; 3) National Neuromuscular Centre, University Hospital of North Norway, Tromsoe, Norway; 4) Department of Clinical Medicine, University of Tromsoe, Tromsoe, Norway; 5) Department of Neurology, University Hospital of North Norway, Tromsø, Norway; 6) Department of Neurology, University Hospital, Oslo, Norway; 7) Department of Paediatrics, University Hospital Oslo, Norway.

Recently, recessive mutations in the ANO5 gene have been identified in two distinct phenotypes, non-dysferlin LGMD2B and distal non-dysferlin Miyoshi-like dystrophy families. Although heterogeneity in clinical features associated with ANO5 sequence variants, has been recognized, most patients presented with adult onset slowly progressive myopathy and high CK levels. In the current study, 26 Norwegian patients with high CK levels, presenting with myopathy of unknown etiology and with no mutations in the FKRP gene, were screened for ANO5 mutations. The most commonly found truncating c.191dupA mutation in exon 5 was identified in 4 patients. However, only a single patient was homozygous for this founder mutation. The other 3 patients were compound heterozygous with less common ANO5 variants. Although half of the investigated patients were females, mutations in ANO5 were exclusively found in male patients. In these patients, the clinical phenotype was heterogeneous and the severity of symptoms varied extensively between patients. Both, molecular results and the clinical findings retrospectively analyzed, of the 4 unrelated patients will be presented.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2966T

Mutation of the *EYS* gene implicated in autosomal recessive retinitis pigmentosa leads to abnornmal cellular aggregates and cell death. *W. Mustill[1], P. Kruczek[1], A. C. Zelhof[2], S. S. Bhattacharya[1].* 1) Institute of Ophthalmology, University College London, London, United Kingdom; 2) Department of Biology, Indiana University, Bloomington, USA.

Retinitis pigmentosa (RP) is a genetically heterogeneous disease leading to blindness due to photoreceptor degeneration. The condition can be inherited in a dominant, recessive, X-linked, or oligogenic fashion. Recessive and isolated cases account for 50–60 % of all forms of RP. In the early stages, rod photoreceptor function is predominantly impaired, leading to defective dark adaptation, night blindness, and constriction of the visual field, followed by impairment of visual acuity from the loss of cone photoreceptor cells and eventually complete blindness. Recent studies have demonstrated a significant prevalence of arRP patients with mutations in the EYS gene. However, little is known of the mechanism by which EYS mutations lead to RP. In this study we demonstrate that abnormal cellular aggregates are formed in cells transfected with mutant EYS DNA compared to WT DNA, and this leads to increased cell death. The WT construct was generated by cloning cDNA of the EYS gene into a GFP-tagged vector and the mutant constructs were generated from this using site directed mutagenesis. The mutant EYS DNA constructs were generated from previously identified sequence variants from arRP patients from France and Spain, and consisted of mis-sense and non-sense (truncating) mutations. This study is the first to investigate the effect of EYS mutations in vitro and is likely to play an important role in understanding the progression of RP.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2967T

Mitochondrial DNA variations in human ova and blastocyst and its clinical implications and IVF outcome. *M. Shamsi[1], G. Periyasamy[2], L. Chawla[3], N. Malhotra[3], N. Singh[3], S. Mittal[3], P. Talwar[4], K. Thangaraj[2], R. Dada[1].* 1) Anatomy, All India Institute Medical Sciences, New Delhi, India; 2) Evolutionary and Medical Genetics Laboratory, Centre for Cellular and Molecular Biology, Hyderabad, India; 3) Department of Obstetrics and Gynaecology, All India Institute of Medical Science, New Delhi, India; 4) ART centre, Army Research and Referral Hospital, New Delhi, India.

Background: Mitochondrial DNA (mtDNA) is maternally inherited and mutations in mtDNA of female germline determine oocyte function and may deteriorate indices of embryo quality and development. Hence, females having pathogenic mtDNA mutations are at high risk for conception difficulties, pre or post implantation loss, fetal anomalies or increased risk of transmission of mtDNA mutations in offspring. Therefore, we investigated the effect of mtDNA variations on the reproductive capacity of oocyte and early embryo (blastocyst) in idiopathic infertile females opting for assisted conception. Methods: Complete mitochondrial genome from single cell 49 oocyte and 18 blastocyst from 67 females was sequenced using long range PCR and analyzed. Findings: Samples (group A) with confirmed disease associated mtDNA mutations- T3396C (Non syndromic hearing loss), T10454C (hearing loss), G11696A (Lebers hereditary optic neuropathy), T12311C (chronic progressive external ophthalmoplegia) and T15908C (hearing loss) had lower fertilization rate and reduced embryo quality as compared to samples which did not have disease associated mtDNA mutations (group B). No difference in the implantation rate and clinical pregnancy rate observed was observed between the groups. Our analysis revealed 437 nucleotide variations most of which were in NADH Dehydrogenase (ND) genes that encode mitochondrial enzyme Complex I. 40.29 % samples had either a disease associated or non-synonymous novel or pathogenic mutation in the evolutionarily conserved region of mitochondrial genome. No association of any haplogroup with mtDNA-associated infertility was found in our study. Interpretations: Dependence of oocyte and embryo on mtDNA differ in various stages of development, therefore pathogenic affects of mtDNA variations manifest differentially during fertilization, implantation and embryonic growth. Screening mtDNA from oocyte or blastocyst offers a better diagnostic approach to understand the etiology of assisted reproduction technique failure or idiopathic infertility. It will help to reduce the risk of transmission of mtDNA diseases to the offspring and also to understand the dynamics of such transmission. MtDNA sequence analysis of oocyte is a simple, efficient and clinically reliable method, which can be used in diagnostic workup of infertile women.

2968T

Cumulus cell microRNA profiles are associated with oocyte aneuploidy. *B. R. McCallie[1], J. C. Parks[1], N. Treff[2], R. T. Scott[2], W. B. Schoolcraft[3], M. G. Katz-Jaffe[1,3].* 1) National Foundation for Fertility Research, Lone Tree, CO; 2) Reproductive Medicine Associates of New Jersey, Morristown, NJ; 3) Colorado Center for Reproductive Medicine, Lone Tree, CO.

Oocyte aneuploidy is a leading cause of miscarriage and age related infertility in human reproduction. MicroRNAs (miRNA) play critical roles in gene regulation and have been shown to be aberrantly expressed in human tumors with chromosome abnormalities. The aim of this study was to investigate a relationship between human oocyte aneuploidy and cumulus cell (CC) miRNA profiles. CCs were donated by couples undergoing infertility treatment with IRB approval and patient consent. Screening for all 23 oocyte chromosomes was performed on biopsied polar bodies using SNP microarray. Groups were classified according to oocyte chromosome constitution: (A) CCs from cumulus oocyte complexes with a euploid oocyte (n=6) and (B) CCs from cumulus oocyte complexes with an aneuploid oocyte (n=6). Expression profiling of 377 miRNAs was performed on individual human CCs using quantitative real-time PCR with the Taqman® Array Human MicroRNA Card (Life Technologies). Predicted target genes were identified with Pathway Studio (Ariadne Genomics) and validated by quantitative real-time PCR. MicroRNA profiling of human CCs identified the expression of 99 miRNAs, 10 (10%) which were differentially expressed in association with oocyte chromosome constitution ($P<0.05$). Six miRNAs showed increased expression with oocyte aneuploidy ($P<0.05$), including miR-29a, whose function when over expressed results in upregulated TP53 levels and the induction of apoptosis. This was confirmed by quantitative real-time PCR with increased TP53 expression observed in CCs of aneuploid oocytes compared to euploid oocytes ($P<0.05$). In addition, four miRNAs displayed decreased expression in association with aneuploid oocytes ($P<0.05$), including miR-16, whose function when decreased results in increased VEGF levels, critical for folliculogenesis. Elevated VEGF concentrations have been correlated with decreased pregnancy rates following infertility treatment. Quantitative real-time PCR validated higher VEGF gene expression in CCs of aneuploid oocytes compared to euploid oocytes ($P<0.05$). This study identified altered miRNA expression of human CCs in association with oocyte aneuploidy. Ongoing analysis may reveal critical roles in gene modulation and pathway regulation that reflect compromised follicular development and oocyte aneuploidy. Evaluation of miRNA expression and their target genes could allow for a non-invasive method for selection of euploid, developmentally competent oocytes.

2969T

Prenatal diagnosis of sirenomelia by combining two-dimensional ultrasound, new fetal skeletal rendering, three-dimensional helical computer tomography and magnetic resonance imaging. *X-l. CHEN[1], R. Liu[1], X-H. Yang[1], N. Zhong[2].* 1) Department of Ultrasound, HuBei Maternal and Children Health Care Hospital, Wuhan, Hubei, China; 2) New York State Institute for Basic Research in Developmental Disabilities, Staten Island, NY.

To evaluate the value of prenatal diagnosis of skeletal abnormalities in sirenomelia, a prospective study was conducted for the period of September 2010 to February 2012. Fetal skeletal rendering, three-dimensional helical computer tomography (3D-HCT) and magnetic resonance imaging (MRI) were performed after two-dimensional ultrasound (2D-US) in seven cases of sirenomelia. Diagnostic skeletal detailed findings with each of the three techniques were compared with postnatal radiological findings and 3D-HCT. All cases performed postnatal autopsy, five cases performed chorionic villus and/or cord blood sampling. Six cases are singleton and one is conjoined twins. The abnormalities associated with seven siremomelia cases included different degrees of fusion of the lower extremities, bilateral renal agenesis, absent bladder, absent external genitalia and single umbilical cord. Five cases are associated with Oligohydramnios, five cases with partial absent ribs, seven cases with spine anomalies and five cases with cardiac anomalies. Karyotype and postnatal autopsy show that three cases are male and two are female. MRI can help to diagnose sirenomelia, it's less useful to skeletal abnormalities. Although 3D-HCT is a gold standard in diagnosis of skeletal abnormalities, the fetuses can't be X-rayed, so it is restricted. The new fetal skeletal rendering seem to be useful complementary method to 2D-US, and can quickly provide skeletal imaging just as 3D-HCT, and may improve accuracy of the prenatal diagnosis of skeletal abnormalities in sirenomelia.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2970T

Comparative Study on Quantitative Parameters of Blood Perfusion and Microvessel Density in Placenta of Rat model of Preeclampsia Evaluated with Contrast-enhanced Ultrasound. *Y-Q. Xu[1], X-L. Chen[1], X-J. Lu[1], X-H. Yang[1], P-W. Chen[1], J-H. Huang[1], C. Huang[1], Z-L. Lu[1], N. Zhong[2].* 1) Department of Ultrasound, Hubei Provincial Maternal and Children's Hospital, Wuhan, Hubei, China; 2) New York State Institute for Basic Research in Developmental Disabilities, Staten Island, NY.

Objective: To investigate the correlation between quantitative parameters with contrast-enhanced ultrasound (CEUS) and microvessel density (MVD). Methods: Twenty placenta of rat model of Preeclampsia underwent CEUS examination. The peak intensity time curves on the enhanced images were analyzed quantitatively with computer to get Quantitative Parameters (The mean transit time, the maximal peak intensity, the area under curve and the mean perfusion volume) . These parameters were compared with MVD counted with immunohistochemistry and the correlation was statistically studied. Results: The peak intensity time in the enhanced images was (28.55±11.45) s, the mean transit time was (134.77±56.45) s, the maximal peak intensity was (84.83±20.15) dB, the area under curve was (7102.61±245.76) dB,The mean perfusion volume was (58.01±23.56) dB, and the MVD of placenta of rat model of Preeclampsia was (230.98±34.78) in 20 rats. The maximal peak intensity was correlated positively to MVD of histological section ($r= 0.885, P < 0.01$). The area under curve was correlated positively to MVD of histological section ($r= 0.883$, $P < 0.01$). The mean perfusion volume was correlated positively to MVD of histological section ($r= 0.877$, $P < 0.01$). There was no correlation between the peak intensity time and MVD in lesions. There was no correlation between the mean transit time and MVD in lesions. Conclusions: The maximal peak intensity, the area under curve and the mean perfusion volume calculated with CEUS were correlated to MVD of placenta of rat model of Preeclampsia immunohistochemistry. These Parameters were valuable index for quantitative evaluation of placental blood perfusion.

2971T

Prenatal Ultrasound and Autopsy Findings in Simpson-Golabi-Behmel Syndrome. *K. Chong[1], J. Jessen[1], M. Injeyan[1], K. Fong[2], S. Salem[2], S. Keating[3], P. Shannon[3].* 1) Dept . of Obstetrics/Gynecology, The Prenatal Diagnosis and Medical Genetics Program, Mount Sinai Hosp, Toronto, ON, Canada; 2) Dept. of Medical Imaging, Mount Sinai Hospital, Toronto, ON, Canada; 3) Dept. of Pathology and Laboratory Medicine, Mount Sinai Hospital, Toronto, ON, Canada.

Simpson-Golabi-Behmel (SGB) syndrome is an X-linked recessive genetic disorder characterized by prenatal and postnatal overgrowth, coarse facial features, polydactyly and an increased risk of embryonal tumors. Both a severe neonatal form with multiple anomalies and a less severe form have been described. There have been few reported cases of pregnancies affected with SGB syndrome prenatally and none with extensive pathological examination. We report two prenatal cases of congenital diaphragmatic hernia with multiple fetal anomalies that were diagnosed post delivery with SGB syndrome. Case 1 presented with normal nuchal translucency at 12 weeks gestation, but increased maternal serum AFP at 16 weeks gestation. There was on family history review, a maternal uncle who was “slow” and died in his teens. Case 2 presented with no previous antenatal screening in the pregnancy. This mother had a unilateral cleft lip and palate identified at birth. There was also a family history of male stillbirths. Upon identification of the multiple fetal anomalies including congenital diaphragmatic hernia on fetal ultrasound, both pregnancies were terminated after extensive counselling. Autopsy findings for both male fetuses revealed increased growth parameters, hypertelorism, congenital diaphragmatic hernia and some features not commonly observed in SGB syndrome: small omphalocele, renal anomaly and ovotestes. Follow up molecular analysis of the GPC3 gene confirmed a partial deletion of exon 3 (Case 1) and a splice site mutation in intron 7 (Case 2) confirming the diagnosis of SGB syndrome in both cases. Maternal DNA testing confirmed heterozygous carrier status for both mothers with X inactivation studies pending. These findings highlight the importance of family history, autopsy and molecular analysis in prenatal cases presenting with non isolated congenital diagphragmatic hernia on fetal ultrasound.

2972T

Testing the most frequent genetic diseases optimizes abnormal gene identity and test accuracy worldwide. *R. Lebo.* Dept Pathology, Akron Children's Hosp, Northeast Ohio Medical University, Akron, OH.

The most frequent worldwide genetic diseases have been selected to enable development and testing that identifies the largest proportion of disease alleles at the fewest number of disease sites analyzed. Targeted genomewide testing the most frequent genetic diseases provides the most reliable, conclusive results without testing relatives. Testing all the 145 most frequent genetic diseases each affecting at least 1 in 100,000 people among more than 1100 tested clinically in GeneTests would identify ~1 in 80 patients affected with a Mendelian disease and the 1 in 3 carriers of mutant alleles worldwide. The diseases on this list are further organized according to transmission category to enable selecting tests for any patient category including: patients selecting partners, pregnant couples, products of conception, fetuses, symptomatic patients of any age, and asymptomatic adults planning to optimize future medical interventions and lifestyle. Test platforms may be strengthened by also including the most frequent population-specific gene mutations. This approach will minimize confirmatory testing of known reported mutations as it optimizes test accuracy of the largest proportion of actionable positive and negative reports. The accompanying list can be analyzed by existing or to be developed targeted platforms that test all frequent disease categories and report only the appropriate category of patient results. Professionals who use this approach to develop, approve, order, test and/or report the most frequent diseases will facilitate genetic counseling and follow up patient care worldwide.

2973T

Single nucleotide polymorphism based method for fetal fraction quantification in maternal plasma using massively parallel sequencing (MPS). *G. Hogg[1], J. Tynan[1], M. Tang[1], S. Sovath[1], A. Maer[1], D. van den Boom[2], P. Oeth[1].* 1) Sequenom Center for Molecular Medicine (SCMM), 3595 John Hopkins Ct., San Diego, CA. 92121; 2) Sequenom, Inc. 3595 John Hopkins Ct., San Diego, CA. 92121.

BACKGROUND: The fraction of circulating cell-free (ccf) fetal DNA in maternal plasma can be measured by quantifying paternal Y chromosome markers or fetal specific methylation patterns of ccf fetal DNA. Here, we demonstrate the ability to detect and quantify ccf fetal DNA fraction in maternal plasma by measuring single nucleotide polymorphism (SNP) alleles using amplicon sequencing via MPS.

METHOD: ccf DNA was extracted from 4 mL plasma from 48 pregnant women using Qiagen ccf DNA extraction kit. A single-tube multiplex PCR was designed to amplify loci of 68 SNPs of high minor allele population frequency from the ccf DNA. A 2-step PCR approach was used to incorporate adapter and index sequences into the amplicons and allow direct capture of amplicon on the flow cell surface for sequencing using the HiSeq™ 2000 sequencer (Illumina, San Diego CA). Flow cells were clustered and sequenced according to the standard protocol for 36 cycles. Reads were aligned to the human genome (hg19) and matched read counts for expected SNP alleles were used to calculate the allele ratio of each SNP within each ccf DNA. Allele ratios based on sequencing counts above the homozygous maternal background were used to calculate the presence of paternal alleles from the fetal fraction not present in the maternal genome across multiple loci per sample.

RESULTS: 48 plasma samples were assayed. All 68 amplicons were detected in each sample. The average paternal allele frequency was used to calculate fetal fraction and ranged from 2 to 14%. These results correlated with a methylation based method of fetal fraction quantification of the same samples. The 68 amplicons were also used to sequence and accurately quantify fractions of a genomic DNA model system. Quantification was accurate using as little as 100,000 reads per sample.

CONCLUSION: This method allows accurate fetal fraction determination in maternal plasma by MPS and allows fetal fraction to be determined concurrently when multiplexed with other types of libraries. The low number of reads required to accurately quantify fetal fraction suggests high levels of sample multiplexing are possible.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2974T

The Pregnancy and Health Profile Pilot Project: Evaluating the impact of integrating a novel family history and genetic screening tool on patients, providers, and clinical practice. *E. Edelman[1], B. Lin[2], T. Doksum[3], B. Drohan[4], K. Hughes[4], S. Dolan[2], V. Edelson[5], J. O'Leary[5], L. Vasquez[6], S. Copeland[6], J. Scott[1].* 1) NCHPEG, Lutherville, MD; 2) March of Dimes, White Plains, NY; 3) Doksum Consulting, Boston, MA; 4) Partners Healthcare, Boston, MA; 5) Genetic Alliance, Washington, DC; 6) HRSA, Rockville, MD.

Effective patient-entered and EHR-compatible collection tools are needed to support translation of family health history (FH) into prenatal care. However, there are limited data about the impact of computerized tools on patient and provider acceptance and outcomes. “The Pregnancy and Health Profile,” (PHP) is a free prenatal genetic screening and risk assessment software for primary prenatal providers that collects FH during intake and provides point-of-care CDS for providers and education for patients. Our objective was to evaluate patient and provider outcomes of PHP in clinical practice. The study population included diverse prenatal providers and patients at 4 sites in IN, ME, NC, and NY. Evaluation included provider pre and post knowledge and usability surveys, pre and post chart audits to assess FH risk assessment, and patient satisfaction surveys. Quantitative survey results were entered into a SPSS database and analyzed using descriptive statistics, Fisher's exact test, and paired t-tests. Qualitative thematic analysis was used for responses to open-ended questions. 513/618 (83%) patients of diverse age and education levels provided feedback; 81% were white, 11% black, and 9% Latina. Patients felt PHP was easy to use (96%) and understand (98%). Ninety-six percent were not worried about the confidentiality of entering FH into PHP. Twenty of 65 providers (10 OBs and 6 family medicine physicians, 2 nurse midwives, 1 nurse, and 1 dietician educator) provided feedback. Providers felt the data collection and patient education aspects of PHP were useful (60%; 67%); there was mixed feedback on the usefulness of the CDS report. Provider confidence in identifying and managing genetic risks improved after using PHP in practice ($p<0.05$). In 2 of 3 sites with performance measure data, PHP improved documentation of a 3-generation FH ($p<0.001$). PHP improved documentation of race and ancestry for patient and/or father of the baby in all 3 sites ($p<0.001$). This study is one of the first to report on the integration of FH into primary prenatal practice. These data demonstrate that PHP is acceptable to patients and providers for providing FH and that PHP collects FH equal to or better than standard practice. Our results can inform future strategies that use point-of-care tools or online portals to improve patient participation around FH and provider management of genetic risks. Additional research is needed to validate patient-entered data in the prenatal setting.

2975T

Chromosomal Structural Abnormalities among Filipino Couples with Recurrent Pregnancy Losses. *E.C. Cutiongco-de la Paz[1,2], A.D. Berboso[1], E.G. Salonga[1], C.M. David-Padilla[1,2].* 1) Institute of Human Genetics, National Institutes of Health, University of the Philippines Manila, Manila, Metro Manila, Philippines; 2) Department of Pediatrics, College of Medicine, Philippine General Hospital, University of the Philippines Manila.

Introduction: Recurrent pregnancy loss (RPL) is a devastating reproductive problem that affects 5% of couples trying to conceive. RPL is classically defined as the occurrence of three or more consecutive pregnancy losses. However, several clinicians start evaluation with two or more because of the recent increase in childless miscarriages. Majority of the cases are due to cytogenetic errors. Objective : This study determines the prevalence of chromosomal structural abnormalities in Filipino couples who presented with 2 or more pregnancy losses. Methods: Results from chromosomal analysis of couples referred for 2 or more miscarriages done at the Institute of Human Genetics-National Institutes of Health-University of the Philippines, Manila on peripheral blood samples from 1991 to 2010 were retrospectively reviewed. Results and Discussion: There were 356 couples with a history of 2 or more miscarriages sent for chromosomal analysis from 1991–2010 included in this study. Among these 356 couples, 17 couples (4.8%) were found to be carriers of different chromosomal abnormalities, 1 of whom had both of them affected with chromosomal abnormalities. From a total of 18 cases, there were 13 (3.6%) translocations, 1 (0.3%) insertion, 2 (0.6%) with marker chromosomes, 1 (0.3%) pericentric inversion and 1 (0.3%) deletion. The results of this study are similar to that of previous large-scale studies which have demonstrated that parental chromosomal abnormalities are associated with RPL. Conclusion: The overall frequency of chromosomal structural abnormalities among patients with RPL in this study is 4.8% with translocations being the most common type detected. It highlights the recommendation that chromosomal analysis be an integral part of the evaluation of couples with two or more pregnancy losses because detection of chromosomal abnormalities is vital in genetic counseling and in the management of subsequent pregnancies.

2976T

Factor II G20210A and factor V G1691A mutations and methylenetetrahydrofolate reductase C677T polymorphism in 155 women with repeated pregnancy loss. *S. Seyedhassani[1,2], M. Houshmand[1], M. Neshan[2], F. Saeb[2].* 1) Medical Genetics, National Institute oF Genetic Engineering and Biot, Tehran, Tehran, Iran; 2) Dr. Seyedhassani Genetic Center, Yazd, Iran.

Introduction: pregnancy is the process from the fertilized ovum to the fetus with capability of extra uterine survival. Pregnancy loss is the most common complication of pregnancies. About 1 in 300 couples and 0.5–2% of women are involved in repeated pregnancy loss (RPL). Various etiological factors involve in RPL and the main part of them remains unknown. Among them the thrombophilic factors are important. Material and methods: Genetic counseling program was done for 158 couples suffering from RPL. Three molecular genetic variations were investigated in main thrombophilic agents: G20210A in factor II, G1691A in factor V and C677T in MTHFR gene. The method was PCR-RFLP. Results: No G20210A mutation was found in Factor II gene. Heterozygote G1691 mutation in factor V gene was found in 3 women (1.94%). But, C677T polymorphism in MTHFR gene was found in 33 women (21.3%). Among them, 4 cases (12.12%) were homozygote and 29 cases (87.88%) were heterozygote. Discussion: Assessment of variations in thrombophilic related genes can be useful in etiologic evaluation and planning of effective treatment in RPL women. Genetic counseling, clinical aspect of abortions and genotype-phenotype correlation should be considered for request of molecular thrombophilic tests in RPL.

2977T

Functional consequences of miscarriage Copy Number Variations (CNVs). *J. Wen[1], CW. Hanna[2], S. Martell[1], PCK. Leung[3], WP. Robinson[2], M. Stephenson[4], E. Rajcan-Separovic[1].*1) Pathology and laboratory medicine, University of British Columbia/Child & Family Res Inst, Vancouver, BC, Canada; 2) Medical genetics, University of British Columbia/Child & Family Res Inst, Vancouver, BC, Canada; 3) Obstetrics and gynecology, University of British Columbia/Child & Family Res Inst, Vancouver, BC, Canada; 4) Obstetrics and gynecology, University of Chicago, Chicago, IL, USA.

Our recent whole genome CGH array-based studies of sporadic and recurrent chromosomally normal miscarriages (N=43) identified unique CNVs in ~40% of miscarriages. The majority of the CNVs were small (<250kb) in size and inherited from either parent. We performed functional analysis of the TIMP2 gene, containing a unique maternal duplication (47kb), that was detected in 4/5 tested miscarriages from a couple with recurrent pregnancy loss (RPL). TIMP2 is of interest as it highly expressed in the placenta, pregnant endometrium and maternal-fetal interface, and plays a role in placental development and extracellular matrix (ECM) remodelling. The expression of TIMP2 was assessed using mRNA and protein extracted from primary cultures of placental tissue from miscarriage placentae (N=4) in comparison to cell cultures from elective abortions (N=4).As the TIMP2 gene has been proposed to be maternally expressed in the placenta, methylation of the promoter CpG island shore and an upstream enhancer was assessed using bisulfite pyrosequencing. Decreased mRNA and protein expression of TIMP2 was detected in miscarriages that carried the CNV. The TIMP2 promoter CpG island shore was 32% methylated in both blood and placenta, and showed no difference between CNV carriers and controls. The upstream enhancer was 96% methylated in blood and 56% methylated in placenta, suggesting a potential role as a tissue-specific enhancer. Similarly, no difference in DNA methylation at the TIMP2 enhancer was observed between CNV carriers and controls. However, there was a trend towards a negative correlation of enhancer methylation with mRNA and protein expression of TIMP2 ($r=-0.47$;$p=0.24$ and $r=-0.44$; $p=0.27$,respectively). Our analysis shows that miscarriage CNVs could result in abnormal RNA and protein expression for their integral genes. In the case of TIMP2 no dramatic methylation abnormalities were detected for the sites nearby the CNV examined in carriers. Aberrant expression of TIMP2 in the female carrier may impair ECM remodelling during decidualization, and/or in the embryo, it may affect placental development, trophoblast invasion or vascularisation. Both events could lead to miscarriage and possibly explain its recurrence. Microarray analysis and functional follow-up of CNVs could therefore enrich our knowledge of the genetic causes of miscarriage and potentially identify couples predisposed to RPL.As a consequence, CNVs are emerging as potential biomarkers for RPL.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2978T

The Fetal Fraction of cell-free DNA in Maternal Plasma is Not Affected by a priori Risk of Fetal Trisomy. *T. Musci[1], H. Brar[2], E. Wang[1], B. Brar[2], A.Y. Batey[1], A. Oliphant[1], M.E. Norton[3].* 1) Ariosa Diagnostics, San Jose, CA; 2) Riverside Perinatal Diagnostic Center, Riverside, CA; 3) Stanford University/Lucille Packard Children's Hospital, Stanford, CA.

Background: Non-invasive prenatal tests analyze cell-free DNA (cfDNA) for the detection of fetal trisomy. The amount of fetal DNA in a sample is a major determinant of the ability to detect trisomy. Objective: To examine possible clinical factors that might influence the fetal fraction of cfDNA in maternal plasma. **Methods**: A comparative analysis was performed on subjects from the multicenter Non-Invasive Chromosomal Evaluation (NICE) Study. Pregnant women 18 yrs and older, at 10 or more weeks gestational age with a singleton pregnancy who underwent invasive testing for any indication, were enrolled. Subjects with confirmed euploid pregnancies were stratified into a priori risk groups for having a fetus with trisomy. “Low-risk” (LR) subjects were less than 35 yrs of age and had low risk results based on serum screening. “High-risk” subjects were stratified by the most advanced maternal age (HR-MA), NT measurements >2.5mm (HR-NT), or those with the highest serum screening risk scores for fetal trisomy 21 or 18 (HR-SS). Fetal fraction was measured using the Ariosa Harmony™ test via targeted sequencing using 192 assays for polymorphic loci on chromosomes 1–12. Fetal percent comparisons were analyzed using analysis of variance (ANOVA) controlling for gestational age. **Results**: The median fetal fraction of cfDNA in each group was LR: 9.7% (n=208), HR-MA: 9.5% (n=208), HR-NT: 9.3% (n=152), and HR-SS: 10.6% (n=208). No significant differences for fetal fraction were observed when comparing women in the LR group to women in the HR-MA group (p=0.56) the HR-NT group (p =0.38) or the HR-SS group (p=0.29). **Conclusions**: Maternal age, serum screening and NT measurements show no significant effect on the fetal fraction of cfDNA in maternal plasma. These data suggest that the non-invasive Harmony™ test for aneuploidy detection is appropriate for a wide range of risk categories in the prenatal patient population.

2979T

Role of E2F1 in Testicular Descent and Fertility. *C. Jorgez[1,2], J. Addai[1], J. Newberg[3], V. Vangapandu[1], A. Sahin[1], J. Rosenfeld[4], L. Lipshultz[1,2], D. Lamb[1,2,3].* 1) Department of Urology, Baylor College of Medicine, Houston, TX; 2) Center for Reproductive Medicine, Baylor College of Medicine, Houston, TX; 3) Department of Molecular and Cell Biology, Baylor College of Medicine, Houston, TX; 4) Signature Genomic Laboratories, PerkinElmer, Inc., Spokane, WA.

E2F1 is a potent transcription factor required for DNA synthesis and cell proliferation. Mice lacking e2f1 are viable and fertile, yet display testicular atrophy with age, exocrine gland dysplasia and a spectrum of tumors. On the other hand, mice overexpressing e2f1 are infertile with testicular atrophy as a result of p53-independent apoptosis occurring within individual tubules resulting in germ cell, but not Sertoli cell depletion. After aCGH analysis of a series of men with non-obstructive azoospermia (NOA), we found a microduplication at 20q11.22 encompassing E2F1 in one NOA man. For confirmation E2F1-CNV-Taqman assays were performed that not only validated the aCGH finding, but also identified 10 of 86 (11.6%; p=0.002) NOA men with microduplications or microdeletions in E2F1. Sequence analysis of E2F1 revealed a SNP at position Ala102Thr in one infertile man that upregulates E2F1 expression in the human testis. Since humans lacking or having an extra copy of E2F1 are infertile, we reevaluated the phenotype of e2f1 deficient mice. The mice are subfertile with a decrease in the number of litters per month, although the number of pups per litter was not affected. By three weeks of age there is a significant decrease in testis weight. The declines in testicular weight, sperm count and motility continue throughout their lifespan. The mice have abnormal testicular histology with mislocalized cell associations, as well as a significant increase in circulating FSH levels. By five months of age fertility is significantly reduced. However, the most surprising finding is that these mice have unilateral (80%) and bilateral (10%) inguinal cryptorchidism. A search of the Signature Genomic Laboratories database of 26,706 individuals tested revealed an infant with an undescended right testis with a microduplication in 20q that includes E2F1. The DECIPHER database also contains a cryptorchid individual with a microdeletion encompassing E2F1 (of 5404 subjects deposited in this database). Additional analysis using E2F1-CNV-Taqman assays identified 2 of 60 (3.3%) cryptorchid individuals with microdeletions in E2F1. Our data suggest a role of E2F1, not only in spermatogenesis, but also in testicular descent. E2F1 plays an important role in male reproductive development and function.

2980T

Study of four SNPs in SLC6A14, INSR, TAS2R38 and OR2W3 genes for association with Iranian idiopathic infertile men. *E. Siasi[1], A. Aleyasin[2].* 1) microbiology, north tehran branch of azad university, Trhran, Iran; 2) molecular and clinical genetic, National institut of genetic engineering and biotechnology, Tehran, Iran.

Back ground- Genome wide SNP association study (GWAS) has reported several SNPs that had potential relevance to oligospermia and azoospermia in idiopathic infertile men in European population. Four of these releavant SNPs are in INSR (T132903C), SLC6A14 (C109869T), OR2W3 (T824C) and TAS2R38 (T886C) genes have reported to cause oligospermia and azoospermia in European populations. Objective- This is the first association study to evaluated their relationships with oligospermia and azoospermia in idiopathic infertile men in Iranian population. Methods- Their polymorphismshave been comparad in 96 male cases with idiopathic infertility and 100 normal control fertile men. SNPs analysis were performed using Real Time High Resolution Melt analysis (PCR-HRM) and confirmed by PCR-RFLP and sequencing analysis. Results- The frequency of two SNPs, T132903C in INSR gene and C109869T in SLC6A14 gene were statistically significant between infertile patients and fertile control groups. Their statistical analysis showed significant associations with idiopathic male infertility with P=0.02 and P= 0.04 respectively. The SNPs frequency for T824C in OR2W3 gene and T886C in TAS2R38 gene were approximately similar among case and control groups with P=0.2 and P=0.9 respectively. Conclusion- These results indicated that T132903C SNP in INSR gene and C109869T SNP in SLC6A14 gene play role in spermatogenesis defect in idiopathic infertility in Iranian case with oligospermia and azoospermia similar to that observed is Europian population.

2981T

NLRP7 subcellular localization in human oocytes and early cleavage stages. *E. Akoury[1,2], L. Zhang[2], M. Seoud[3], A. Ao[2], R. Slim[1,2].* 1) Department of Human Genetics, McGill University Health Centre, Montreal, Quebec, Canada; 2) Department of Obstetrics and Gynecology, McGill University Health Centre, Montreal, Quebec, Canada; 3) Department of Obstetrics and Gynecology, American University of Beirut, Beirut, Lebanon.

Hydatidiform mole is an abnormal human pregnancy with no embryo and cystic degeneration of placental villi. Maternal mutations in NLRP7, a gene that is expressed in oocytes, early cleavage embryos, and endometrial tissues, are responsible for recurrent moles and reproductive wastage. In addition, post-zygotic aneuploidies during early embryo development have been demonstrated in patients with NLRP7 mutations. However, to date the subcellular localization of NLRP7 in human oocytes and early developmental stages have not been reported. Using immunofluorescence experiments and confocal microcopy, we performed an extensive analysis to determine the subcellular localization of NLRP7 in leftover human oocytes and early cleave stage from patients undergoing pre-implantation genetic diagnosis. Our findings demonstrate that NLRP7 is indeed a maternal-effect gene and are in agreement with previous data about two other maternal genes characterized in mice, MATER (NLRP5) and Filia. In addition, our data shed new lights on the role of NLRP7 in postzygotic aneuploidies, implicate NLRP7 in early tissue differentiation, and will be presented at the meeting.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2982T

Human spermatogenic failure purges deleterious mutation load from the autosomes and both sex chromosomes, including the gene DMRT1. *D.F. Conrad[1], A. Lopes[2], K.I. Aston[3], F. Carvalho[4], J. Goncalves[5], R. Mathiesen[2], N. Huang[6], A. Ramu[1], J. Downie[7], S. Fernandes[8], A. Amorim[2,8], A. Barros[9], M. Hurles[6], S. Moskovtsev[10], C. Ober[11], J. Schiffman[7], P.N. Schlegel[12], M. De Sousa[13], D.T. Carrell[3, 14].* 1) Dept Genetics, Washington Univ School Med, St Louis, MO; 2) IPATIMUP, Institute of Molecular Pathology and Immunology of the University of Porto, R. Dr. Roberto Frias S/N, 4200-465 Porto, Portugal; 3) Andrology and IVF Laboratories, Department of Surgery; 4) Department of Genetics, Faculty of Medicine, University of Porto, Porto, Portugal; 5) Centre for Human Genetics, National Institute of Health Dr. Ricardo Jorge, Lisbon, Portugal; 6) Genome Mutation and Genetic Disease Group, Wellcome Trust Sanger Institute, Cambridge, UK; 7) Department of Oncological Sciences; 8) Faculty of Science, University of Porto, 4099-002 Porto, Portugal; 9) Centre for Reproductive Genetics Alberto Barros, Porto, Portugal; 10) Department of Obstetrics & Gynaecology, University of Toronto; 11) Department of Human Genetics, Department of Obstetrics & Gynecology, The University of Chicago, Chicago, IL 60637, USA; 12) Department of Urology, Weill Cornell Medical College, New York-Presbyterian Hospital, New York, USA; 13) Laboratory of Cell Biology, UMIB, ICBAS, University of Porto, Porto, Portugal; 14) Department of Physiology, Department of Obstetrics and Gynecology University of Utah School of Medicine, Salt Lake City, Utah, 84108, USA.

Gonadal failure, along with early pregnancy loss and perinatal death, may be an important filter that limits the propagation of harmful mutations in the human population. We hypothesized men with spermatogenic impairment, a condition with unknown genetic architecture and a common cause of male infertility, are enriched for rare deleterious mutations compared to men with normal spermatogenesis. We assayed genomewide SNPs and CNVs in 327 men with spermatogenic impairment and >1100 controls, and estimated that a rare autosomal deletion multiplicatively changes a man's risk for this condition by 10% (OR 1.10 [1.05–1.15], $p < 4 \times 10^{-5}$), a rare X-linked CNV by 29%, (OR 1.29 [1.16–1.43], $p< 3 \times 10^{-6}$) and a rare Y-linked duplication by 64% (OR 1.64 [1.28–2.10], $p < 9 \times 10^{-5}$). Based on the population frequency of potential risk alleles, extent of homozygosity, and evidence for dosage sensitivity of genes disrupted in men with spermatogenic impairment, we propose that the CNV burden is polygenic and distinct from the burden of large, dominant mutations described for developmental disorders. Our study also identifies focal deletions of the sex-differentiation gene *DMRT1* as likely recurrent causes of idiopathic azoospermia, and generates hypotheses for directing future studies on the genetic basis of male infertility and IVF outcomes.

2983T

A homozygous deletion of DPY19L2 gene is a cause of globozoospermia in men from the Republic of Macedonia. *D. Plaseska-Karanfilska[1], P. Noveski[1], S. Madjunkova[1], I. Maleva[1], V. Sotiroska[2], Z. Petanovski[2].* 1) Macedonian Academy of Sciences and Arts, Research Centre for Genetic Engineering and Biotechnology " Georgi D. Efremov", Skopje, Republic of Macedonia; 2) Private Hospital "Re-Medika", Skopje, Republic of Macedonia.

Globozoospermia is a rare but severe teratozoospermia, characterized by ejaculates consisting completely of round-headed spermatozoa that lack an acrosome or, in partial globozoospermia, containing a variable proportion (20%–90%) of acrosomeless spermatozoa. Men that are affected with total globozoospermia are infertile, and even the application of intracytoplasmic sperm injection (ICSI) has met with disappointingly low success rates. Several case reports of affected brothers suggested a genetic component to this disease. Homozygous mutations in *SPATA16* (spermatogenesis-associated 16) and *PICK1* (protein interacting with C kinase 1) genes were described as responsible for a complete form of globozoospermia in an Ashkenazi Jewish and a Chinese family, respectively. However, no mutations in these genes were detected in other men with globozoospermia, thus suggesting that*SPATA16* and *PICK1* genes are not the main loci associated with this condition. Recently, a homozygous 200 kb deletion on chromosome 12 encompassing only *DPY19L2* gene was identified in several patients from different origins, thus suggesting that deletions involving this gene might represent a major cause of globozoospermia. It was shown that the breakpoints of the deletions are located on a highly homologous 28 kb low copy repeat (LCR) sequence present on each side of *DPY19L2*, indicating that the identified deletions were probably produced by nonallelic homologous recombination (NAHR) between these two regions. Here we describe the presence of a homozygous deletion of the *DPY19L2* gene in two men with complete form of globozoospermia from the Republic of Macedonia. The deletion was detected by array CGH analysis, using 180K Agilent Human Genome CGH Microarrays and Genomic Workbench software. The results were confirmed with classical PCR using seven loci, of which three were intragenic, located in exons 1, 11, and 22 of the *DPY19L2* gene, while four were located in the LCR regions. PCR fragments were obtained only from the first centromeric and last telomeric locus in the LCR regions. Both patients are of Macedonian ethnic origin, live in the same town (Tetovo), but are not related. No consanguinity was reported in any of the two families. Four ICSI cycles were attempted in the two patients, but no pregnancy was achieved. In conclusion, our results support the notion that a homozygous deletion of *DPY19L2* gene represents a major cause of globozoospermia.

2984T

Decreased Puberty and Fertility Development in NELF KO Mice due to impaired GnRH neuron Migration. *S. Quaynor, L. Chorich, R. Cameron, L. Layman.* Institute of Molecular Medicine and Genetics, Georgia Health Sciences University, Augusta, GA.

Normal puberty and reproduction require the proper development and function of the hypothalamic-pituitary-gonadal axis controlled by gonadotropin-releasing hormone (GnRH) neurons. GnRH neurons develop in the olfactory placode region and migrate into the brain, where they project their processes to the median eminence. Lack of migration of these neurons results in diseases such as Idiopathic Hypogonadotropic Hypogonadism (IHH) and Kallmann syndrome (KS). Patients with IHH/KS present with absent puberty, low gonadotropins and sex steroids in addition to anosmia in KS. Nasal embryonic luteinizing hormone releasing factor (NELF) is an important protein that has been identified to contribute to GnRH neuron migration, and mutations of the gene have been demonstrated in IHH/KS patients. We previously confirmed *Nelf*'s role in GnRH neuron migration using immortalized mice GnRH cell lines; and to date all the studies regarding *Nelf* have been done using immortalized neuronal cells and zebrafish models. To extend these studies into mammals, we generated homozygous Nelf knockout (KO) mice to further understand the phenotypic effect with regard to GnRH neuron function in the whole animal. We hypothesized that *Nelf -/-* mice will have impaired pubertal development and compromised fertility. To assess puberty of *Nelf* KO animals, we measured hormone panels (LH, FSH, testosterone, and estradiol) and determined vaginal opening and estrus onset. To determine the effect of *Nelf* KO upon fertility, we measured estrus cycles, and determined the number of days between litters and the litter sizes of both homozygous males and homozygous females when they were mated with wild type animals. Finally, we confirmed the defects in GnRH neuron migration and cell number through the performance of immunohistochemistry on serial brain slices using GnRH antibody. Our preliminary results indicate that female *Nelf -/-* mice have delayed vaginal opening by five days compared with wild type mice. In addition, both male and female *Nelf -/-* mice have impaired fertility as manifested by a decreased litter size and increased time between litters. Preliminary evidence suggests that GnRH neuronal migration distance is decreased in *Nelf* KO mice compared to wild type. These data support our clinical findings of human NELF mutations and show that NELF plays an important role in both normal pubertal development and fertility.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2985T

Deficiency of PRSS37, a putative Trypsin-like serine protease, causes male infertility from mouse to human. *Z.G. Wang[1,3,5], C.L. Shen[2], J.B. Liu[3], J.S. Feng[4], Y. Kuang[5], H.X. Zhang[1], W.T. Wu[5], J. Chi[5], L.Y. Tang[3], J. Fei[5].* 1) State Key Laboratory of Medical Genomics, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine (SJTUSM), Shanghai, Shanghai, China; 2) Institute of Health Sciences, Shanghai Institutes for Biological Sciences of Chinese Academy of Sciences and SJTUSM, Shanghai, China; 3) Department of Medical Genetics, SJTUSM, Shanghai, China; 4) Department of Histology and Embryology, SJTUSM, Shanghai, China; 5) Shanghai Research Centre for Model Organisms, Shanghai, China.

PRSS37 is a putative Trypsin-like serine protease. Human serine proteases are involved in numerous physiological processes and pathological states. However, the physiological function of PRSS37 and its potential association with human reproductive capacity remain unknown to date. In the present work, we describe the crucial role of PRSS37 in male fertility. We show that Prss37 is exclusively expressed in adult testis, especially in developing elongating spermatids during spermiogenesis, and almost vanished in mature sperm. Mice deficient for Prss37 display significantly decreased male fertility, but their mating activity, testicular weights and histology, sperm counts, morphology and motility in vitro remain unaffected. In vivo fertilization assay reveals that Prss37-/- mice exhibit a markedly decreased fertilization rate (2.3% vs. 70% of that in control mice), accompanied by the defect in sperm transit from uterus to oviduct. In vitro study further shows incapability of sperm in sperm-egg recognition/binding when exposure of zona-intact eggs to Prss37-/- sperm, in which mature Adam3 was completely undetectable. These data indicate that Prss37-deficiency causes male infertility due to the defects of sperm migration from uterus to oviduct and sperm-egg binding via regulating Adam3 maturation. The infertile phenotypes in Prss37-/- mice perfectly reproduce the situation described as “unexplained male factor infertility (MFI)” in human. Thus, we further investigate the potential implication of PRSS37 in human male infertility. We show that human PRSS37 protein exists in the acrosome region of ejaculated human sperm, which is different from the mouse. Importantly, PRSS37 protein levels in 98 sperm samples from the patients with MFI (all normal semen parameters) are significantly decreased on average when compared to that in 30 healthy sperm samples (PRSS37/GAPDH ratio, 3.8 vs. 0.9, p<0.0001). Most strikingly, PRSS37 in 7 out of 98 patients (7%) is undetectable (ratio<0.01). In combination of the data acquired from the mouse and human studies, we summary our work as follows: Prss37 is a highly specifically expressed protein in late spermatids during spermiogenesis in mice. It is crucial for sperm migration from uterus to oviduct and sperm-egg binding through regulation of Adam3 maturation. Absence or decreased expression level of PRSS37 in human sperm could be one of the causes for “unexplained MFI”.

2986T

Changes in placental DNA methylation may be associated with karyotypically normal miscarriage. *CW. Hanna[1], DE. McFadden[2], WP. Robinson[1].* 1) Department of Medical Genetics, University of British Columbia, Vancouver, BC, Canada; 2) Department of Pathology, University of British Columbia, Vancouver, BC, Canada.

Miscarriage, the spontaneous abortion of pregnancy before 20 weeks gestation, is the most common complication of pregnancy with a frequency of 15–20%. Chromosomal errors are observed in over 50% of miscarriages; however causes of karyotypically normal losses are less understood. Appropriate patterns of DNA methylation in the fetus and placenta are essential to maintain a healthy pregnancy. As altered DNA methylation may be a cause of pregnancy loss, we hypothesize that placental villi of karyotypically normal miscarriages will exhibit aberrant DNA methylation at imprinted loci and globally throughout the genome. DNA was extracted from 70 chromosomally normal first trimester placental villi samples: 33 miscarriages from women experiencing recurrent miscarriage (RM), 21 miscarriages from women with no history of miscarriage (M), and 16 control elective terminations (TA). DNA methylation at the imprinted loci*PLAGL1*, *H19/IGF2* ICR1, *SGCE*, *SNRPN*, *CDKN1C*, KvDMR1, *MEG3*, *AXL*, and *APC*, as well as at LINE-1 sequences, was assessed using bisulfite pyrosequencing. Ten of 33 (30.3%) RM placentae showed a loss/gain of methylation (defined as values greater than 1.5x the inter-quartile range) at any of the 9 imprinted loci, as did 1 of 21 (4.8%) M and 3 of 16 (18.8%) TA placentae (p=0.07). After correction for multiple comparisons, there was a significant difference in average methylation at *H19/IGF2* ICR1 (p<0.0001). Genome-wide DNA methylation was assessed in a subset of 20 samples (10 RM, 10 TA) using the Illumina Infinium HumanMethylation27 BeadChip array, which interrogates 27,578 CpG sites. Significance Analysis of Microarrays was utilized to identify candidate genes, using a <5% False Discovery Rate and >5% absolute difference between groups. Three of 14 candidate CpG sites, of functional importance, were followed up with pyrosequencing, including sites near the *CYP1A2* (cg04988473), *APC* (cg20311501) and *AXL* (cg14892768) genes. There was significantly increased methylation in RM (58.6%) and M (64.4%) compared to TA (50.8%) at *CYP1A2* promoter (p=0.002). We found no difference in DNA methylation at LINE-1 sequences or average array methylation between groups. This study suggests that altered DNA methylation at specific loci (*H19/IGF2* ICR1, *CYP1A2*) and not global dysregulation in the placenta, may contribute to karyotypically normal pregnancy loss in women, or be a consequence of placental insufficiency leading to miscarriage.

2987T

Co-culture of mouse embryonic stem cells with Sertoli cells promote in vitro generation of germ cells. *M. Miryounesi[1], K. Nayernia[2], M. Dianatpour[1], S. Savad[1], M. H. Modarressi[1].* 1) Department of Medical Genetics, School of Medicine, Tehran university of Medical Sciences, Tehran, Iran; 2) GENEOCELL, Advanced Molecular & Cellular Technologies, Montreal, Canada.

About 10 to 15 percent of couples worldwide have fertility problems, and half of them are infertile men. It seems that germ cells divide indefinitely to maintain sperm production in male. Sertoli cells support in vivo germ cell production; however the mechanism for this support has not been well understood. We designed the present study to analyze the effect of sertoli cells in differentiation of mouse embryonic stem cells (mESCs) to germ cells. Methods: To select differentiated mouse embryonic stem cells (mESCc) we used a fusion construct composed of the 1.4 kb promoter region of germ line specific gene Stra8 and Coding region of enhanced green fluorescence protein. Linearized vector (35 µg) was electroporated into ES cells. resistant colonies were selected and cultured 72 hours in the presence of Retinoic Acid (RA). GFP positive cells were selected by using fluorescence-activated cell sorting (FACS) and cultured for one month in non induced condition. To analyze sertoli cells effect in differentiation process, mESCs were separated in two groups: the first group was cultured on gelatin with retinoic acid treatment and the second group was co-cultured with sertoli cell feeder without retinoic acid induction. Expressions of pre-meiotic (Stra8), meiotic (Dazl and Sycp3) and post-meiotic (Prm1) genes were evaluated at different differentiation stages (+7, +12 and +18 days of culture). Results: In the first group, the expression of meiotic and post meiotic genes starts at 12 and 18 days after induction with RA respectively. In the second group, 7 days after co-culturing with Sertoli cells, expression of meiotic and post-meiotic genes were observed Discussion: In this study the effect of co-culturing of ESC and Sertoli cell line (TM4) in the process of differentiation of ESC was investigated. These results clearly show that differentiation process to germ cells is supported by Sertoli cells, which provide a novel effective approach for generation of germ cell in vitro and studying interaction of germ cells with their niche.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2988T

Androgen receptor CAG repeat length may influence the risk of polycystic ovarian syndrome (PCOS). *J. C. Silas[1,2], Singh. Rajender[3], S. Nirmala Sadasivam[4], M.N. Sadasivam[4], Singh. Lalji[5], K. Thangaraj[1].* 1) Centre for Cellular and Molecular Biology, Hyderabad, India; 2) Centre for Genetics and Inherited Diseases (CGID), Taibah University, Al- Madinah, Kingdom of Saudi Arabia; 3) Division of Endocrinology, Central Drug Research Institute, Lucknow, India; 4) Maaruthi Medical Centre and Hospitals, Erode, Tamil Nadu, India; 5) Banaras Hindu University, Varanasi, India.

Polycystic ovarian syndrome (PCOS) refers to an inheritable endocrine disorder in which the multiple fluid-filled sacs or cysts grow under the surface of one or both of the ovaries of women. Now PCOS is considered as an androgen excess disorder or hyperandrogenism. Since the effect of androgen is through the androgen receptor (AR), CAG triplet repeat length in the coding region of the AR gene could influence the risk of PCOS. It has been established in vitro that AR activity is inversely proportional to the length of the CAG repeats. We hypothesized that length of the CAG repeat could affect the androgenic response and hence the risk of PCOS. Methods used: In total, 169 patients following strictly defined inclusion criteria for PCOS were recruited. A total of 175 control samples were collected from ethnically similar normal women who volunteered for participation in the study. CAG repeat region of the AR gene was screened by PCR amplification followed by gene-scan and direct DNA sequencing methods. Distribution of the allele data and the mean allelic length was compared between cases and controls and between obese and non-obese cases. Along with the present study from India, we also conducted a meta-analysis on data from 11 studies including a total of 1777 cases and 1918 controls. Results: We found that the CAG repeat length was between 7 and 29 for both cases and controls. Mean and median CAG values were almost same for cases (17.4±3.3 and 18 repeats, respectively) and controls (17.4±3.31 and 18 repeats, respectively). CAG allele distribution pattern also looked similar between these two groups. Upon categorizing the cases into obese and non-obese groups, the allele distribution and mean repeat length looked indifferent. We observed a strong correlation between the DHEAS level and the presence of hirsutism; however, CAG repeat length was not associated with hyper-androgenism. Overall, meta-analysis showed no association between CAG length and PCOS; however, upon removal of two sensitive studies, data was almost homogenous (Q-value=8.727, df (Q) = 8.000, P-value=0.366, I-squared=8.330) and a strong association between CAG length and PCOS was seen. Overall, increased CAG length appeared to increase the risk of PCOS (fixed and random effect model p values, 0.001 and 0.002, respectively). Conclusion: CAG length does not associate with PCOS risk in the studied Indian population; however, overall, increased CAG length appears to increase the risk of PCOS.

2989T

An analysis of the initial experience offering cell-free fetal DNA testing to pregnant women. *J. Taylor[1], A. Ables[2], L. Hudgins[3].* 1) Perinatal Genetics, Lucile Packard Children's Hospital, Stanford, CA; 2) Master's Program in Human Genetics and Genetic Counseling, Stanford University School of Medicine, Stanford, CA; 3) Department of Pediatrics, Division of Medical Genetics, Stanford University School of Medicine, Stanford, CA.

Noninvasive prenatal testing (NIPT) for aneuploidy via maternal serum cell-free fetal DNA (cffDNA) recently became clinically available. In our Perinatal Genetic Counseling Clinic, NIPT is offered for indications including positive aneuploidy screening, ultrasound markers, and ultrasound abnormalities as well as to those who request a discussion of prenatal testing options including invasive procedures. Patients receive genetic counseling which includes a review of the cffDNA technology, and emphasizes the screening nature of this testing and limitation of screening to aneuploidy involving chromosomes 21, 18 and 13. The purpose of this report is to share our initial experience, highlighting the clinical utility and limitations of this novel test. Between 3/1/12 and 4/30/12, 76 women were offered NIPT. The test was accepted by 48 (63%) and declined by 28 (37%). For those who accepted NIPT, the median a priori Down syndrome (DS) risk was 1/94. For those who declined NIPT, the median a priori DS risk was 1/51. Of those patients who accepted testing, 43 (90%) had a “low risk” result (<1/10,000). Of those who had a low risk result, 4 continued on to have amniocentesis with normal results. Three patients (6%) received “high risk” results (>99% risk): 2 for trisomy 21 (T21) and 1 for trisomy 18 (T18). T21 was confirmed by amniocentesis in 1 patient and the other 2 patients declined confirmation but are using this information for delivery management. The 2 remaining patients (4%) received “no result” reports because of a low fetal fraction of DNA. Of the 28 patients who declined NIPT testing, 20 (78%) pursued diagnostic testing (median a priori DS risk: 1/29). Fourteen had amniocentesis and 6 had CVS. Of these, 7 (35%) were abnormal (median a priori DS risk: 1/5), including 5 cases of T21, 1 case of T18, and 1 case of tetrasomy 12p. Eight of the 76 (11%) patients declined all testing (median a priori DS risk: 1/510). In our initial experience the majority of women with relatively lower risk chose to pursue NIPT or declined all testing while women with relatively higher risk chose to proceed with diagnostic testing. In this initial review of our data, we find that NIPT is pursued by many women in place of diagnostic testing. However, it is essential that patients have genetic counseling to review the limitations of this technology, as illustrated by the tetrasomy 12p result that would not have been detected through NIPT.

2990T

Identification of messenger RNA of fetoplacental source in maternal plasma of women with normal pregnancies and pregnancies with intrauterin growth restriction. *P. Ayala Ramírez[1], R. García Robles[1], J. D. Rojas[2], M. Bermúdez[1], J. Bernal Villegas[1].* 1) Instituto de Genética Humana, Pontificia Universidad Javeriana, Bogotá, Colombia; 2) Department of Obstetrics and Gynecology Hospital Universitario San Ignacio, Pontificia Universidad Javeriana. Bogotá, D.C., Colombia.

Objective: to quantify placenta-specific RNA in plasma of women carrying foetuses with intrauterine growth restriction and pregnant women with normal pregnancies. Materials and methods: 8 pregnant women with foetuses with intrauterine growth restriction were studied as well as 18 women with uncomplicated pregnancies in the third pregnancy trimester. Total free RNA was quantified in maternal plasma by spectrophotometry and the gene expression of hPL (Human Placental Lactogen) at the messenger RNA level through technical Real Time-Chain Reaction Polymerase. Results: plasma RNA of fetoplacental origin was successfully detected in 100 percent of pregnant women. There were no statistically significant differences between the values of total RNA extracted from plasma ($p = 0.5975$) nor in the messenger RNA expression of hPL gene ($p = 0.5785$) between cases and controls. Conclusion: messenger RNA of fetoplacental origin can be detected in maternal plasma during pregnancy.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2991T

Instant familial haplotyping in conjunction with embryo analysis in preimplantation genetic diagnosis using DNA microarrays. *G. Altarescu[1], D. A. Zeevi[1], S. Zeligson[1], S. Perlberg[2], T. Eldar-Geva[2], E. J. Margalioth[2], E. Levy Lahad[1], P. Renbaum[1].* 1) Zohar PGD Lab, Medical Genetics Institute, Shaare Zedek Medical Center, Jerusalem, Israel; 2) IVF Unit, Shaare Zedek Medical Center, Jerusalem, Israel.

Introduction: PGD for molecular disorders involves development of gene and family specific protocols that are performed prior to the PGD cycle. Whole genome amplification (WGA) in conjunction with microarray platforms allows analysis of a very large number of SNPs. The aim of this study was to validate instant familial haplotyping in conjunction with blastomere analysis during PGD cycles. Methods: A PGD cycle was performed for a couple with paternal Charcot Marie Tooth type 1A disease using multiplex nested PCR for five informative STRs surrounding the duplicated 17p12 region. All mutant embryos from the case were blindly reanalyzed using WGA and microarray SNP analysis. WGA was performed on single blastomeres and blastocyst biopsies (2–5 cells) using the Repli-G kit. Genotype analysis was performed using Affymetrix Gene Chip Human Mapping 250K Nsp Arrays. Results: One blastomere and one blastocyst biopsy from two different embryos found to be mutant during the PGD cycle were reanalyzed along with genomic DNA samples from the CMT1A patient, his spouse, and affected mother using SNP microarrays. Wild type and mutant paternal haplotypes were assigned based on the affected mother. Out of 4844 total SNP probes on chromosome 17, 649 were identified as paternally informative, and 378 SNPs remained informative for the wild type or mutant allele when the maternal genotype was introduced. Qualitative analysis showed a 96% SNP call rate in blastocyst biopsies and 90% in blastomeres. Allele drop out (ADO) rates ranged between 3–18% (blastocyst) and 27–33% (blastomeres). After ADO analysis 103 to 137 SNPs were available for haplotype construction in blastocyst and 97 to 109 SNPs in blastomere. The whole chromosome haplotypes in both blastomere and blastocyst biopsies clearly showed that both embryos received the mutant paternal chromosome, reconfirming the initial analysis. Two recombination events were identified in all samples at +13.88 Mb and -54.94 Mb from the disease associated gene, and included 69–87 informative SNPs. Conclusions: Instant haplotyping using microarrays in conjunction with PGD analysis appears to be accurate and feasible during the time frame required for embryo transfer. This universal methodology drastically reduces pre-case haplotype preparation. Whole chromosome analysis allows fine-mapping of recombination events, thereby facilitating inclusion of a much larger number of genetic markers for haplotype construction.

2992T

Role of Cytokines in Recurrent Miscarriages. *F. Parveen, S. Agrawal.* Medical Genetics, Sanjay Gandhi PGIMS, Lucknow, Uttar Pradesh, India.

Introduction: Different single nucleotide polymorphisms within promoter and regulatory regions of different cytokines may influence the level of cytokines. It is now a well-established fact that during recurrent fetal losses there is a imbalance between Th1 and Th2 cytokines. Keeping this in mind we have studied the association of gene polymorphisms in the Th1 and Th2 cytokines i:e interferon-γ gene (IFN-γ +874 T/A), interleukin-4, interleukin-6, interlekin-10 gene (IL10; -1082 G/A,819, 592) and transforming growth factor-β gene (TGFB; +869 T/C) with RM cases and normal fertile women without the history of any fetal loss.. Materials and Methods: The polymorphism of the genes IFN-γ and TNF-α and interleukin IL-6, IL-4, IL-10 and TGF-β were assessed in 200 North Indian RM cases and 300 ethnically matched true negative controls by using RFLP and allele-specific oligonucleotides-polymerase chain reaction method. Results: We applied dominant and recessive model to test the above hypothesis. Our results revealed that there was no associations inTh2: IL-10(592 C/A) and the TGF-β gene anda significant association was observed inTh2: IL-4C590T, IL-6G174C, IL-10 (1082A/G&819 C/T) and Th1: IFN-γ +874 A/T between RM versus controls. Conclusions: This study indicates that both the Th2: IL-4C590T, IL-6G174C, IL-10 1082A/G and Th1: IFN-γ +874 A/T dominant cytokines are associated with RM.

2993T

Developmental neuropsychological assessment of 4–5 year old children born after preimplantation genetic diagnosis: A pilot study. *G. Sacks[1], J. Guedalia[3], T. Eldar-Geva[2], T. Gilboa[4], E. J. Margalioth[2], E. Levy-Lahad[1], G. Altarescu[1].* 1) PGD Lab, Medical Genetics Institute, Shaare Zedek Medical Center, Jerusalem, Israel; 2) IVF Unit, Shaare Zedek Medical Center, Jerusalem, Israel; 3) Neuropshychology Unit, Shaare Zedek Medical Center, Jerusalem, Israel; 4) Peiatric Neurology Unit, Shaare Zedek Medical Center, Jerusalem, Israel.

Introduction: Preimplantation genetic diagnosis (PGD) is a relatively new technique that enables the identification of affected embryos prior to transfer in couples at risk for a genetic disorder. Due to its relative invasiveness it is essential to identify any adverse outcomes and rule out potential insults to the long term development of children born as a result of PGD. The aim of this study was to measure and evaluate developmental neuropsychological profiles of 4–5-year-old children born after PGD and compare them to the norm. Materials and methods: Twenty seven, 4–5year-old, PGD children were recruited. Participants were subjected to a battery of neuropsychological assessments including: WPPSI - III (cognitive development), PLS-4-heb (language development), WRAVMA (visual motor abilities), CARS2 (a screening test for autistic spectrum disorders) and the Miles ABC Test (ocular dominance). Parental questionnaires included: BRIEF-P (executive function), CBCL and The CTS-BSQ (socio-emotional development). In addition, parents were interviewed to determine independent functional abilities of the children (VABS2-Interview edition). Subject's tests results were compared to each test's norms as provided by the test authors. Results: Children born after PGD showed scores within the normal range for all developmental outcomes. The mean WPPSI Full Scale Intelligence Quotient (FSIQ) was 102.7 (±13.3). Moreover, the children's mean Verbal Intelligence Quotient (VIQ) was 107.4 (±14.4); statistically significantly higher (P=0.013) than the general population, although not clinically significant (100±15). The mean total PLS-4 language score was 113.2 (±12.4), the mean WRAVMA scores for the visual-motor (drawing test), and fine-motor skills (pegboard test) were 104.4 (±13.4) and 97.7 (±15.5) respectively, the mean VABS Adaptive Behavior Composite was 106.1 (±11.4); all within normal test ranges. Socio-emotional development and executive function were all clinically normal. 40.7% of PGD cases were mixed eye-hand dominance. One of the subjects was found to show possible signs for autistic spectrum disorders and was sent for further evaluation, although a family history of autism was noted. Conclusions: This pilot study shows that children conceived after PGD show normal developmental neuropsychological outcomes at age 4–5 years when compared to the norms representing the general population of children of the same age.

2994T

Terminological clarification in PGD: Could reframe the debate? *S. Côté[1, 2], P. Hamet[1,2,3].* 1) Medical Genetic service, Montreal University health center, Montreal, Canada; 2) Biomedical sciences, faculty of medicine, University of Montreal, Montreal, Canada; 3) Canada Research Chair in predictive genomic, Montreal, Canada.

Although, most cases of preimplantation genetic diagnosis (PGD) are undertaken for medical reasons, as some authors specify, the choice of language in discussing the process is important in order to avoid stigmatization and discrimination of patients who use PGD (Bouffard et al. 2009). These authors add that when experts discussing PGD use terms such as eugenics or designer baby, there is a risk of shifting the debate and therefore preventing a deeper questioning about the possible misuse of PGD for social reasons. Therefore we are proposing a terminological clarification. Genetic diagnosis is originally defined as a genetic test to identify a genetic condition or predisposition to this condition. Actually, as the use of PGD goes far beyond the frameworks of medicine to better frame discussions and to respect parents seeking PGD for medical reasons, we propose that the term PGD be reserved exclusively for requests based on diagnostic indications. We suggest that PGD not be used in cases of social sex selection because identifying the sex of an embryo is not a genetic diagnosis and as such cannot, under any circumstances, be considered as a disease. Therefore, for social uses we suggest using terms that would be close to PSS (preimplantation social selection). However, should PGD be used in cases where sex selection is done in the context of an X-linked genetic disorder? If the embryo selected were a carrier of the mutation, would it not by definition be a generally healthy carrier? With regard to PGD used for the health benefits of a family member, we use the HLA typing as a component of genetic diagnosis. In this case, we could use PGD-HLA typing. Unfortunately, the vague terminology in this field risk to obscure objectives sought by a legitimate parental wish to avoid the burden of transmitting a genetic mutation. Echoes of these concerns can be heard not only in academic discussions but also in debates in the general population using terms designer baby and perfect child in a misguided fashion. Finally, in a context where social uses of PGD are increasing, we believe that a reflection on the use of a fair and appropriate terminology is necessary in order to could foster a climate of respect for patients who use PGD for medical reasons, facilitating health providers and insurances reimbursements.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2995T

A novel, generic, preimplantation genetic diagnosis protocol applied to cystic fibrosis involving mutation detection through high resolution melting analysis and simultaneous haplotype analysis through QF-PCR. *A. Destouni, M. Poulou, G. Kakourou, C. Vrettou, J. Traeger-Synodinos, I. Fylaktou, E. Kanavakis, M. Tzetis.* Medical Genetics, University of Athens, St.Sophia's Children's Hospital, Athens, Attica, Greece.

Cystic Fibrosis (CF) is the most common, autosomal recessive severe monogenic disease in Caucasians and a common indication for PGD. It is characterized by an extremely heterogeneous mutation spectrum (>1800 mutations http://www.genet.sickkids.on.ca/). Most PGD protocols to preclude CF described to date apply single-cell multiplex fluorescent PCR for analysis of linked polymorphic short tandem repeats (STR-linkage approach), including testing for the commonest CFTR mutation, p.Phe508-del, when indicated. This potentially has limiting application for cases with other mutations when phase analysis is unavailable. Here we report the development and successful clinical application of a novel, flexible, generic PCR PGD protocol to facilitate direct detection of any CFTR nucleotide variation(s) by HRM, and simultaneous confirmation of diagnosis through STR-haplotype analysis. A touch-down, multiplex PCR was optimized supporting co-amplifcation of any CFTR exon-regions, along with linked 6 STRs. The primers used for CFTR HRM analysis were previously described (Montgomery et al. 2007). Diluted (1/1500) 1st PCR products were used for a nested PCR to amplify separately the CFTR regions for HRM analysis (Idaho's LightScanner). CFTR genotypes were confirmed by STR analysis of the 1st PCR products. STR analysis provides an indication of embryo ploidy and facilitates contamination monitoring. The protocol was validated pre-clinically according to PGD guidelines, by testing single lymphocytes, isolated (by micromanipulation) from whole blood samples of candidate PGD patients. Four clinical PGD cycles were performed, for 4 CF carrier couples with the following CFTR genotypes: a) p.Phe508del and p.Phe508del, b) p.Phe508del and c.489+1G>T; c) p.Arg334Gln and c.489+3A>G; d) p.Phe508del and p.Leu732X, (HGVS CFTR reference sequences NM000492.3 and NG016465.1). 34 embryos were biopsied and a total of 34 blastomeres were analyzed. Five samples failed to amplify at all loci. Genotypes were achieved in 29/29 amplified samples, of which 17 were suitable for embryo transfer (unaffected CFTR genotypes). Transfer of 9 embryos led to 3 pregnancies (2 twin and 1 singleton). PGD genotypes were confirmed following conventional amniocentesis or trophoblast PND testing. The reported PGD method is a flexible and robust tool which facilitates direct CF genotype analysis, genotype confirmation and contamination detection in single cells, with minimal family work-up prior to PGD (generic).

2996T

PGD and heteroplasmic mitochondrial DNA point mutations: A systematic review estimating the chance of healthy offspring. *D.M.E.I. Hellebrekers[1], R. Wolfe[2], A.T.M. Hendrickx[1], I.F.M. de Coo[3], C.E. de Die[1,4], J.P.M. Geraedts[1,4], P.F. Chinnery[5], H.J.M. Smeets[1,4].* 1) Clinical Genetics, Maastricht University Medical Centre, Maastricht, Netherlands; 2) School of Public Health and Preventive Medicina, Monash University, Melbourne, Australia; 3) Pediatric Neurology, Erasmus MC-Sophia Children's Hospital, Rotterdam, The Netherlands; 4) Research Institute GROW, Maastricht University Medical Centre, Maastricht, Netherlands; 5) Institute of Human Genetics, Newcastle University, Newcastle upon Tyne, UK.

Background: Mitochondrial disorders are often fatal multisystem disorders, partly caused by heteroplasmic mitochondrial DNA (mtDNA) point mutations. Prenatal diagnosis is generally not possible for these maternally inherited mutations because of extensive variation in mutation load among embryos and the inability to accurately predict the clinical expression. The aim of this study is to investigate if PGD could be a better alternative, by investigating the existence of a minimal mutation level below which the chance for an embryo of being affected is acceptably low, irrespective of the mtDNA mutation. Methods: We performed a systematic review of muscle mutation levels, evaluating 159 different heteroplasmic mtDNA point mutations derived from 327 unrelated patients or pedigrees, and reviewed three overrepresented mtDNA mutations (m.3243A>G, m.8344A>G and m.8993T>C/G) separately. Results: Mutation levels were included for familial mtDNA point mutations only, covering all affected (n=195) and unaffected maternal relatives (n=19) from 137 pedigrees. Mean muscle mutation levels were comparable between probands and affected maternal relatives and between affected individuals with tRNA- versus protein-coding mutations. Using an estimated a-priori prevalence of being affected in pedigrees of 0.477, we calculated that a 95% or higher chance of being unaffected was associated with a muscle mutation level of 18% or less. At mutation level 18%, the predicted probability of being affected is 0.00744. The chance of being unaffected was lower only for the m.3243A>G mutation (P<0.001). Most carriers of mtDNA mutations will have oocytes with mutation levels below this threshold. Conclusions: Our data show, for the first time, that carriers of heteroplasmic mtDNA mutations will have a fair chance of having healthy offspring, by applying PGD. Nevertheless, our conclusions are partly based on estimations and, as indicated, do not provide absolute certainty. Carriers of mtDNA should be informed about these constraints.

2997T

Successful pre-implantation genetic diagnosis in isolated sulfite oxidase deficiency: A happy end to the saga of a devastating neurogenetic disorder. *A. Kondkar[1], K. Abu-Amero[1,4], M. Salih[3], T. Bosley[2].* 1) Department of Ophthalmology, College of Medicine, King Saud University, Riyadh, Saudi Arabia; 2) Department of Ophthalmology and Pediatrics, College of Medicine, King Saud University, Riyadh, Saudi Arabia; 3) Department of Neurology, College of Medicine, King Saud University, Riyadh, Saudi Arabia; 4) Department of Ophthalmology, College of Medicine, University of Florida, Jacksonville, FL, USA.

Isolated sulfite oxidase deficiency (ISOD) is an autosomal recessive disorder that typically causes seizures shortly after birth followed by cystic cerebral cortical degeneration and death at an early age. This report presents a consanguineous Sudanese family who had three children who were biochemically diagnosed with ISOD. Another child died shortly after birth without definitive diagnosis; a fetus was aborted because of anencephaly; and a normal fetus aborted after diagnostic chorionic villus sampling. Sequencing revealed a novel two base successive deletion c.1232–1233 delTG in the sulfite oxidase gene on chromosome 12. Pre-implantation genetic diagnosis (PGD) and in vitro fertilization resulted in the birth of a healthy child. Patients with ISOD generally develop intractable seizures within days of birth followed by what appears on neuroimaging to be acute brain ischemia affecting predominantly cerebral white and grey matter, basal ganglia, and other deep grey structures. Various dietary manipulation and molybdenum cofactor treatment have been suggested as partial treatments for molybdenum cofactor deficiency for this type of genetic disorder but only OGD offers a definitive reproductive option for families with ISOD.

2998T

Outcomes of 687 in vitro fertilization (IVF) cycles and 5871 embryos evaluated using 23-chromosome single nucleotide polymorphism (SNP) microarray preimplantation genetic screening (PGS) for recurrent pregnancy loss (RPL). *K.J. Tobler[1], P.R. Brezina[1], A.T. Benner[2], L. Du[2], B. Boyd[2], W.G. Kearns[1,2].* 1) Department of Gynecology and Obstetrics, Johns Hopkins Medical Institutions, Baltimore, MD; 2) Genetics, Center for Preimplantation Genetics, Lab Corp, Rockville, MD.

Objective: We determined the pregnancy and miscarriage rates of patients undergoing IVF and dense microarray PGS for RPL.

Design: Retrospective review

Materials and Methods: A retrospective review was conducted of all embryos that underwent PGS by 23 chromosome SNP microarrays from January 2010 to April 2012. Patients underwent standard IVF and PGS primarily due to a history of ≥ 2 spontaneous miscarriages. Embryo biopsy was performed at either the cleavage or blastocyst stage. Sample DNA was amplified and analyzed using Human CytoSNP-12 DNA beadchips and GenomeStudio and KaryoStudio software. Embryos derived from parents with known translocations or inversions were excluded from the study. Data was split into four patient groups based on maternal age (<35, 35–37, 38–40, >40). Microarray data was also compared to previously published fluorescence in situ hybridization (FISH) PGS data from our laboratory (previously published). Binomial confidence intervals for proportions were calculated.

Results: A total of 5871 molecular karyotypes were obtained from 687 IVF cycles. The overall pregnancy rate per IVF cycle for women in this trial undergoing 23-chromosome SNP microarray PGS for RPL was 57% with a miscarriage rate of 9%. In blastocyst stage PGS, the pregnancy rate, 75%, was significantly (p<0.001) higher than the cleavage stage pregnancy rate, (47%). Miscarriage rates were significantly (p=0.002) higher in the cleavage stage (12%) than the blastocyst stage (5%). Both cleavage and blastocyst stage PGS yielded significantly (p<0.001) higher pregnancy rates than FISH PGS. Pregnancy rates decreased and miscarriage rates increased in all groups with advancing maternal age.

Conclusions: PGS using 23 chromosome SNP microarrays in women with RPL was superior to FISH PGS. Furthermore, performing the biopsy at the blastocyst stage was superior to cleavage stage biopsy. The application of dense microarray 23 chromosome PGS at the blastocyst stage may substantially improve pregnancy outcomes in couples suffering from RPL.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

2999T

Cleavage-stage blastomere biopsy significantly impairs human embryonic reproductive potential while blastocyst-stage trophectoderm biopsy does not: A paired randomized controlled trial utilizing SNP microarray-based DNA fingerprinting. *N.R. Treff[1,2,3], K.M. Ferry[1], T. Zhao[1], J. Su[1], E.J. Forman[1,2], R.T. Scott Jr.[1,2].* 1) Reproductive Medicine Associates of New Jersey, Morristown, NJ; 2) Department of Obstetrics, Gynecology, and Reproductive Sciences, UMDNJ-Robert Wood Johnson Medical School, New Brunswick, NJ; 3) Department of Genetics, Rutgers-The State University of New Jesery, Piscataway, NJ.

Preimplantation genetic diagnosis (PGD) requires embryonic DNA which necessitates embryo biopsy. Despite the use of PGD for over 2 decades, the risks associated with embryo biopsy itself, and whether biopsy is safest at the cleavage or the blastocyst stage of embryo development, have not been well characterized. This study uses a novel paired randomized controlled trial (RCT) to accurately assess the impact of embryo biopsy at these 2 stages of development for the first time. Patients received routine care inclusive of embryo selection. Once the 2 best embryos were selected for transfer, one was randomized to biopsy and the other to control. Biopsy was performed with the most commonly used methodology involving a diode laser and gentle aspiration of either a blastomere (cleavage) or trophectoderm cells (blastocyst). The sibling biopsied and non-biopsied embryos were then transferred together. With resulting singleton deliveries, SNP microarray based fingerprinting of the embryo biopsy and newborn DNA was used to determine whether the child was derived from the biopsied (self) or non-biopsied (sibling) embryo. A McNemar paired test of proportions was used to evaluate significance. Results were obtained from 100% of blastomere and trophectoderm biopsies. 117 patients completed the study, including 55 cleavage stage and 62 blastocyst stage transfers. There was a significant reduction in delivery rates amongst biopsied compared to the non-biopsied cleavage stage embryos (29% compared to 46%, respectively; P=0.049); a loss of delivery potential of 39%. In contrast, the delivery rate of biopsied and non-biopsied blastocysts was not significantly different (47% and 48%, respectively; P=1.00). Optimizing clinical outcomes following any type of embryo genetic analysis (aneuploidy, single gene disorder, or translocation screening) might therefore be best realized by performing biopsies at the blastocyst stage. This is supported by the observation that the only RCTs to demonstrate a clinical benefit from aneuploidy screening utilized blastocyst stage biopsy. Potential explanations as to why blastocyst biopsy is safe when cleavage stage biopsy is not include the removal of a smaller proportion of the embryo, certainty that only extra-embryonic cells are biopsied, and better embryonic tolerance of manipulation after genomic activation.

3000T

Comparison of two whole genome amplification methods for preimplantation genetic diagnosis with HLA typing on beta-thalassaemia. *Q. Wang, J.F.C. Chow, W.S.B. Yeung, E.H.Y. Ng, P.C. Ho.* Dept. of Obs. & Gyn., The University of Hong Kong, Hong Kong.

Preimplantation genetic diagnosis (PGD) combined with HLA typing is a preventive and therapeutic approach allowing the birth of unaffected children who are also HLA identical to the affected siblings with life-threatening disorders such as beta thalassaemia. This study aimed to compare the performance of two whole genome amplification (WGA) methods used in PGD, SurePlex (BlueGnome) and REPLI-g (Qiagen) WGA kits for the HLA typing on single lymphocytes from HBB mutation carriers with known serological HLA types. HLA typing was determined by the CTS-PCR-SSP TRAY KITS (UCH). The genotype patterns of HLA regions A, B, and DRB1 obtained from patients' genomic DNA samples were used as references. WGA products generated through the SurePlex and REPLI-g kits were analyzed and the results of single cell HLA typing were compared. Our data revealed that following manufacturer's instructions, the Qiagen REPLI-g Kit was able to amplify DNA templates specifically, produce longer products, and detect the three mentioned alleles, while the SurePlex kit caused non-specific amplification in negative controls containing no DNA templates and failed to detect the HLA-DRB1 region properly. The detection failure of the HLA-DRB1 region with the SurePlex kit was also observed on the WGA products amplified from diluted genomic DNA samples (equivalent to the amount of DNA estimated in a single cell). These findings indicate that besides the well-known phenomena of allelic drop-out (ADO, PCR failure of one allele) and preferential amplification (PA, hypo-amplification of one allele), the inability of WGA kits to amplify target regions (i.e., amplification failure of both alleles) could occur, and might cause misdiagnosis. Therefore, a careful evaluation of the ability of WGA kits to amplify target regions is essential to ensure diagnostic efficacy and accuracy.

3001T

Preimplantation genetic diagnosis for a carrier of complex chromosome rearrangement resulting in a healthy offspring. *K. Writzl, A. Veble, B. Peterlin.* Institute of Medical Genetics, Ljubljana, Slovenia.

Complex chromosome rearrangements (CCR) are structural aberrations involving more than two breakpoints on two or more chromosomes. Carriers of CCR are at increased risk for infertility, miscarriages or chromosomally unbalanced offspring. Preimplantation genetic diagnosis (PGD) may avoid the transfer of embryos carrying unbalanced rearrangements, therefore increasing the chance of pregnancy. This is the first report for a live birth after PGD in the CCR carrier associated with translocation and inversion. A carrier of CCR was ascertained through the birth of an infant who had an unbalanced karyotype and died at 10 days. Using conventional cytogenetic methods the rearrangement was characterized as a translocation between the long arms of chromosomes 11 and 18. Whole chromosome paint fluorescent in situ hybridisation (FISH) analyses revealed additional rearrangement - paracentric inversion of derivative chromosome 11. Blastomeres were biopsied from 18 embryos in two PGD cycles, and 17 embryos (94.4%) were successfully diagnosed by FISH method. The probes were chosen to detect unbalanced translocated products and dicentric and acentric recombinant products in the event of recombination. The segregation mode most frequently found was adjacent - 1 (47%), followed by alternate (41%). No evidence for adjacent-2 segregation or 3:1 segregation mode was found. A pregnancy was achieved in the second PGD cycle, and the karyotype of amniocentesis was normal. A healthy baby was delivered. This report supports the benefit of PGD in CCR couples and adds to understanding of the mechanism of meiotic segregation in CCR. It also suggests that FISH method should be used to detect CCRs in apparently balanced translocation carriers.

3002T

Preimplantation genetic diagnosis (PGD) in Genomic Regions with Duplications and Pseudogenes: Long Range PCR in the Single Cell Assay. *D.A. Zeevi[1], R. Ron El[2], P. Renbaum[1], E. Kasterstein[2], D. Strassburger[2], D. Komarovsky[2], B. Maslanski[2], I. Ben-Ami[2], O. Bern[2], A. Komski[2], E. Levy Lahad[1], G. Altarescu[1].* 1) Medical Genetics Institute, Shaare Zedek Medical Center, Jerusalem, Israel; 2) IVF and Infertility unit, Assaf Harofeh Medical Center, Tel Aviv University.

Introduction: Long and short range nested PCR are generally employed for the analysis of disease-causing mutations in genes with homologous pseudogene copies. However, long range PCR is challenging when performed on single cells, as for PGD of monogenic disorders. PGD on single cells requires concurrent analysis of a mutation together with multiple linked polymorphic markers from close family members in order to reduce misdiagnosis. In PGD cases involving childless de novo mutation carriers, linkage cannot be performed based on family members but rather must first be identified in single gametes. However, if the mutation to be assayed lies in a duplicated genomic region, the design of an appropriate single cell assay, incorporating a long and short range-coupled multiplex PCR, can be an especially difficult task. Here we describe a novel method by which accurate PGD of de novo pseudogene-homologous mutations can be achieved. Methods: Whole genome amplification (WGA) was performed on single sperm using a multiple displacement amplification kit. Subsequently, the WGA samples were subjected to long range PCR followed by nested PCR and sequencing of a mutation in PKD1. In parallel, the samples were analyzed with gene-proximal polymorphic STR markers in order to determine allelic association. Results: Haplotypes were constructed around a pseudogene homologous mutation in PKD1 (c.701delC; de novo autosomal dominant male affected) in single sperm. Direct PCR mutation analysis could not be performed on the mutation carrier because the gene-specific allele was heavily diluted in 6 pseudogene copies. In contrast, WGA on 16 single sperm samples followed by long range PCR-mediated analysis clearly identified the PKD1 gene-specific locus and mutation. In addition, linkage analysis identified two distinct haplotypes; 11 sperm with haplotype A and 5 sperm with haplotype B. Surprisingly, this analysis showed that the PKD1 mutation carrier is a mosaic since only 3 out of 5 sperm with haplotype B possessed the mutation, unlike 11/11 sperm with haplotype A. Conclusions: Whole genome multiple displacement amplification enables the simultaneous analysis of short polymorphic marker, and long range gene-specific, PCR fragments. This original analytical scheme is universal and facilitates accurate PGD analysis in cases where mutation-linked haplotypes must be constructed from gametes (single sperm/polar body) of pseudogene-homologous mutation carriers.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3003T

TAR Syndrome Diagnosis Confirmed by aCGH Analysis. *S. Ceylaner[1], Y.K. Terzi[2], S. Kalyoncu[3], F.I. Sahin[2].* 1) Intergen Genetics Center, Ankara, Turkey; 2) Baskent University Faculty of Medicine Department of Medical Genetics, Ankara, Turkey; 3) Jineclinic Women's Health Center, Ankara, Turkey.

Thrombocytopenia-absent radius (TAR) is a rare congenital syndrome affecting less than 1 in 100,000 newborns, and is characterized by bilateral radial aplasia and thrombocytopenia. Also a number of additional features such as skeletal and cardiac anomalies, dysmorphic features, renal malformations, and intellectual deficit have been observed in TAR syndrome. In the present study we analyzed a 22-weeks-old male fetus with severe micrognathia, phocomelia of upper extremities, bilateral joint limitations of hip, knee and ankle of lower extremities and diastasis recti. Absence of the long bones of the upper extremities was also confirmed by X-ray imaging. Parents are healthy non-consanguineous individuals. To uncover the genetic cause of these findings array comperative genomic hybridization (aCGH) analysis was performed. DNA was isolated from amniotic fluid sample. Roche NimbleGen Human CGH 3×720K Whole-Genome Tiling v3.0 Array was used for aCGH analysis. The data were analyzed by using web based Genoglyphix program (Signature Genomics). As a result of aCGH analysis, variations in 12 regions were thought to be clinically significant, eleven of them were copy gain on chromosomes 2, 7, 11, 17, 19, 20 and X, and one of them was copy loss on chromosome 1. Deletion, located at 1q21.1, encompassed 346 kb of genomic DNA. This region comprises 16 genes, 9 of which are OMIM genes. Deletion in this region was found to be related to thrombocytopenia-absent radius syndrome. Chromosome analysis was also done from amniotic fluid and the result was normal constitutional karyotype (46,XY). Recent findings showed that deficieny of the RBM8A gene protein product Y14 may be responsible for development of TAR syndrome. Also other identified copy gains may have contributed to the accompanying dysmorphic findings. Array CGH analysis is a valuable tool in uncovering submicroscopic chromosomal rearrangements that cannot be discerned by conventional cytogenetic analysis.

3004T

Prenatal detection of anophthalmia in three cases with significantly different neonatal outcomes. *T. Bardakjian, D. Dorsainville, A. Schneider.* Dept Peds/Genetics, Albert Einstein Med Ctr, Philadelphia, PA.

Anophthalmia and microphthalmia can be detected prenatally via ultrasound. Genetic counseling after identification of this rare and severe birth defect is complicated given its etiological heterogeneity and phenotypic variability. We describe three prenatally identified cases where parents contacted the Anophthalmia /Microphthalmia Registry at Einstein Medical Center, Philadelphia after the 20 week “anatomy survey” ultrasound identified anophthalmia in the fetus. Two of the fetuses were identified with unilateral anophthalmia and the third with bilateral anophthalmia. All three had follow-up fetal MRI which confirmed the findings. Additional testing offered by the obstetrician varied. None were offered molecular testing for eye development genes. These cases demonstrate the variability seen in this condition and the difficulty encountered when a prenatal diagnosis of anophthalmia is made. We will propose a detailed prenatal work-up to help elucidate the situation for a family facing this diagnosis in utero and make genetic counseling more targeted.

3005T

Antenatal detection of a fetus with bilateral anophthalmia and pulmonary hypoplasia - a rare case report. *C.C. Albu[1,2], D.F. Albu[1,2,4], M. Dumitrescu[2,3], E. Severin[1].* 1) University of Medicine and Pharmacy Carol Davila Bucharest, Romania; 2) Alco San Clinic, Bucharest, Romania; 3) Marius Nasta Hospital, Bucharest, Romania; 4) Panait Sirbu Hospital, Bucharest, Romania.

Background: Matthew-Wood syndrome is a rare disorder including as main characteristics anophthalmia/microphthalmia, and pulmonary hypoplasia/aplasia, diaphragmatic hernia/eventration and cardiac congenital malformation. Case report: A 31-year-old Caucasian female, pregnant for the first time, was referred at 16 weeks' gestation for a routine prenatal ultrasound. The couple had normal general health and was not consanguineous. There was no family history of chromosomal anomalies. Methods: Routine ultrasonography at 16 weeks of pregnancy, triple test (AFP,uE3, hCG), selective ultrasonography for detection of fetal abnormalities and amniocentesis were performed. Results: Ultrasound examination revealed a single fetus with bilateral anopthalmia, pulmonary hypoplasia and diaphragmatic hernia. Triple test was not sensitive to the presence of trisomy 13, 18 and 21. The fetal karyotype indicated a normal cytogenetic female: 46,XX. Because of prenatal ultrasound detection of bilateral anophthalmia, the parents decided to terminate the pregnancy. An autopsy showed bilateral anopthalmia, pulmonary hypoplasia and diaphragmatic hernia. Conclusions: Matthew-Wood syndrome can be diagnosed prenatally by routine ultrasound examination. Prenatal ultrasound diagnosis of a sporadic case of Matthew-Wood syndrome was useful in the management of this case.

3006T

Outcomes following antenatal diagnosis of ultrasonographically isolated cleft lip with or without cleft palate. *A. Lehman[1], L. Burnell[1], C. Verchere[2], D. Pugash[3], S. Robertson[4], A. Loo[5].* 1) Dept Med Gen, Univ British Columbia, Vancouver, BC, Canada; 2) Dept Surgery, Univ British Columbia, Vancouver, BC, Canada; 3) Dept Radiology, Univ British Columbia, Vancouver, BC, Canada; 4) Cleft Palate / Craniofacial Program, BC Children's Hospital, Vancouver, BC, Canada; 5) Faculty of Dentistry, Univ British Columbia, Vancouver, BC, Canada.

Cleft lip with or without cleft palate (CLP) is one of the most common congenital malformations with an average incidence of 1 / 700 worldwide. Clefting of the fetal lip is routinely screened for during the second trimester detailed ultrasound scan. We reviewed postnatal outcomes following an antenatal diagnosis of ultrasonographically isolated CLP in a ten period between January 2000 and December 2009. All pregnancies in British Columbia referred to our tertiary care centre for a level III ultrasound were ascertained and then matched with the corresponding pediatric chart in our Cleft Palate / Craniofacial Program. Isolated CLP was defined as the absence of other fetal structural malformations on our level III scan. A minimum of 2 years of follow-up was reviewed. Ninety-seven cases were identified, 1 of which had no outcome data. Fifteen pregnancies led to termination and 13 had some amount of post-mortem examination or investigation recorded. Seventy-six of the 81 liveborn infants had at least 2 years of postnatal follow-up data available for review. The majority had no significant medical or developmental issues beyond effects from the clefting. Eleven of the 76 had a diagnosed or strongly suspected complex developmental disorder. Two of these children had inherited a clefting syndrome from a parent (Ectrodactyly ectodermal dysplasia and Blepharocheilodontic dysplasia). One had Trisomy 21. The other children had Noonan syndrome, Kabuki syndrome, autism with intellectual disability, autism, suspected autism spectrum disorder, a CNS malformation, and the last had a combination of short stature, minor anomalies, and significant behavioural difficulties. Another 11 children had minor anomalies not requiring any therapy (eg., transient hydronephrosis, non-closure of the fontanelle, small ventricular septal defect). With a normal karyotype and negative family history for a clefting syndrome, the frequency of an associated complex developmental disorder following an otherwise reassuring fetal ultrasound is around 10 percent. Combining our data with comparable studies on fetal karyotyping for isolated CLP, the yield is about 1 percent.

3007T

cfDNA is not compromised by temperature stress of maternal blood collected in Streck Cell-free DNA BCT. *A. Huang, A. Srinivasan, V. Nguyen, J. Cheng, J. Sprowl, R. Van Luchene, A. Aziz, D. Comstock.* Verinata Health, Inc., 800 Saginaw Drive, Redwood City, CA 94063.

Objective: Simulating temperature stresses that patient blood tubes endure during transit from a collection site to the clinical testing facility, this study examined the use of plasma cfDNA isolated from temperature-stressed maternal blood in massively parallel sequencing-based prenatal diagnostic screening assays. Methods: Four blood tubes each from 121 pregnant donors were received in Streck Cell-free DNA BCTs by overnight shipping. The first tube was the control (D1) while the remaining three tubes (D2, D4, D6) were subjected to repeat cycles of 24hr temperature stress (15hrs at 25C, 4hrs at 30C, 5hrs at 40C). The stress regimen mimicked typical “summer extreme” industry standard for shipping temperature profile. Plasma cfDNA was isolated from these samples, barcoded sequencing libraries generated and the libraries sequenced on the Illumina HiSeq2000. For each donor specimen, normalized chromosome ratios were derived from the sequencing data and compared across the four conditions. A normalized chromosome ratio is the number of unique tags aligning to a specific chromosome divided by the sum of tags aligning to one or more “normalizing” chromosomes. This ratio, with %CVs of less than one across hundreds of samples, is a quantitative metric for copy number variation that accounts for GC and sample-to-sample biases. Normalized chromosome ratio is very sensitive to small changes in copy number. Results: Plasma cfDNA quality was evaluated by comparing average Ratio data for Chrs 13, 18, and 21 for all conditions. Fetal fraction representation was evaluated by comparing R2 of linear correlation plots for Chrs X and Y for maternal samples with male fetuses. Average Ratio_18 and Ratio_21 from the 121 samples were largely unaffected by the temperature stress; average Ratio_13 showed a slight upward trend. Fetal fraction was not compromised by the temperature stress. Conclusion: Repeated temperature stress did not compromise cfDNA quality or fetal fraction in maternal blood collected in Streck BCT.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3008T

Effective epigenetic biomarker in first-trimester maternal plasma for non-invasive fetal trisomy 21 detection. *J. Lim[1], D. Lee[1], S. Kim[1], M. Kim[1], S. Park[1], Y. Han[2], M. Kim[2], M. Kim[2], K. Choi[2], H. Ryu[1,2].* 1) Laboratory of Medical Genetics, Medical Research Institute, Cheil General Hospital and Women's Healthcare Center, Seoul, South Korea; 2) Department of Obstetrics and Gynecology, Cheil General Hospital and Women's Healthcare Center, KwanDong University college of Medicine, Seoul, Korea.

Purpose: Noninvasive prenatal diagnosis of trisomy 21 (T21) is being actively researched using epigenetic biomarkers for specific detection of fetal DNA in maternal plasma. The *HLCS*(holocarboxylase synthetase) and *RASSF1A* (Ras association domain family member 1) genes are located on chromosome 21 and 3, respectively. Their promoters have regions which are hypermethylated in the placenta and hypomethylated in maternal blood cells. We estimated the accuracy of non-invasive fetal T21 detection using a tissue-specific epigenetic characteristic of*HLCS* and *RASSF1A* during the first-trimester of pregnancy.

Methods: A nested, case-control study was conducted with maternal plasma collected from 40 pregnant women carrying 30 normal and 10 T21 fetuses. In this study, we used the methylated DNA immunoprecipitation (MeDIP) methodology for specific isolation of methylated DNA in maternal plasma. The maternal plasma levels of *HLCS* and *RASSF1A* were simultaneously measured by methylation-specific multiplex real-time quantitative PCR.

Results: The *HLCS* and *RASSF1A* levels were obtained in all cases. These levels were not different according to fetal gender (P>0.05 in both). The *HLCS* level was significantly increased in women with a T21 fetus compared with controls (P<0.05). The *RASSF1A* level was not different between the T21 cases and controls (P>0.05). The specificity of HLCS level and ratio of*HLCS* and *RASSF1A* levels for non-invasive fetal DS detection showed 66.7% and 90.0% at 90% sensitivity, respectively.**Conclusion:** The findings of this study suggest that ratio of *HLCS* and *RASSF1A* levels may be useful as potential biomarkers for non-invasive fetal T21 detection during the first trimester of pregnancy, regardless of fetal gender.

3009T

Prenatal diagnosis of a fetus with mosaic trisomy 18, omphalocele and severe intrauterine growth restriction: case report. *D.F. Albu[1,2], C.C. Albu[1,2], A. Oncescu[2,4], E. Severin[1], M. Dumitrescu[2,3].* 1) University of Medicine and Pharmacy Carol Davila, Bucharest, Romania; 2) Alco San Clinic, Bucharest, Romania; 3) Marius Nasta Hospital, Bucharest, Romania; 4) Panait Sirbu Hospital, Bucharest, Romania.

Background: Omphalocele is a midline abdominal wall defect with extrusion of abdominal viscera, covered by a membranous sac, into the base of the umbilical cord and is the one of the most common congenital malformation of the anterior abdominal wall. Omphalocele is frequently associated with other congenital malformations. However, the frequency of the reported associated malformations for omphalocele ranges from 27%; to 63%;. Objective: To determine wether omphalocele and intrauterine growth restriction (IUGR) are syndromic or non syndromic. Patients and methods: A 30 year old, Caucasian woman was referred at 17 weeks' gestation for a routine prenatal ultrasound. The couple had normal general health and was not consanguineous. There was no family history of congenital malformations and chromosomal abnormalities. Routine ultrasonography at 17 weeks of pregnancy, triple test (AFP,uE3, hCG), selective ultrasonography for detection of fetal abnormalities, and amniocentesis were performed. Ultrasound evaluation revealed severe intrauterine growth restriction (IUGR). Immediate amniocentesis was performed. It revealed the elevation of the AFP in the amniotic fluid. FISH analysis showed mosaic trisomy 18. The pregnancy was terminated one week later. Conclusions: Our case provides further evidence for the benefit of advanced imaging for prenatal diagnosis and the understanding of fetal condition. More important for our patient was her better understanding of fetal disorder due to 3D findings.

3010T

Chromosomal abnormalities in prenatal diagnosis: An Indian perspective. *B.B. Ganguly[1], N.N. Kadam[2].* 1) Genetics, MGM Center for Genetic & Diagnosis, New Mumbai, Maharashtra, India; 2) Department of Pediatrics, MGM Institute of Health Sciences, New Mumbai, India.

Statement of purpose: In cases with family history of mental retardation or suspected genetic illness, previous child with chromosome anomaly and advanced maternal age are generally considered for prenatal screening of double/triple/quadruple markers in maternal serum, which collects information about possibility of genetic abnormality in the fetus and calculates risk of Down syndrome, Edward syndrome and neural tube defects. The perceived risk directs prenatal chromosome analysis. Ultrasound imaging also prompts for prenatal diagnosis. Chromosome analysis is generally not considered as a straightforward approach for prenatal diagnosis in cases with significant background information. In the present paper, an analysis has been carried out for ascertaining the incidence of chromosomal anomaly in cases where screening result indicated high risk. Method: Amniotic fluid was cultured in replicate sets following standard technique for 140 cases. One of the primary cultures with 10–15 colonies was considered for harvesting and chromosome preparation approximately on day 8–10. Conventional G banding and chromosome analysis was carried out with the help of IKAROS software. Result: The cases registered in the study were comprised of 28.6% of 21–30 years of age, 69% of 31–40 years and 2% of 41–48 years. Of them, 65% had high risk of Down syndrome, 10% was with advanced age, 9% had family history, 4% with previous child carrying chromosome anomaly, and 1.4% each with IVF pregnancy, parental chromosome aberration, and cystic hygroma or echgenic bowel. Trisomy 21 was observed in 4.4% suspected cases only, though significant risk was detected in 65% cases on maternal serum screening. One case showed both inv(9) and t(15;16). Altogether, constitutive abnormality was observed in 9.2% cases. Other abnormalities, including inv(9), t(1;3), der(9)t(9;21)mat, der(22)t(9;22)mat, del(9q13), and polymorphic variability such as 14p+, 1qh+, 9qh+, 21p+, 22p+, etc. were recorded in 6% cases. Therefore, it has been realized that the practice of maternal serum screening is popular; however, traumatize patients with false positive result. And it may be imperative that false negative result may lead to delivery of chromosomally abnormal children. Chromosome analysis, in any circumstances, appears to be a more reliable and may be adopted as a direct approach for prenatal diagnosis.

3011T

A Powerful "Smart Tip" Method for Nucleic Acid Extraction and Enrichment of Cell-Free Fetal DNA from Maternal Plasma. *A. Gindlesperger, T. Stokes, P. Belgrader, R. Holmberg.* Akonni Biosystems, Frederick, MD.

Background: Non-invasive prenatal diagnostics (NIPD) is a much safer alternative to invasive techniques, but is more challenging due to the low concentration of highly fragmented fetal DNA in a high background of maternal DNA circulating in maternal plasma during early gestation periods. These conditions require the processing and concentration of large sample volumes to achieve adequate amounts of fetal DNA for analysis. The high ratio of maternal to fetal DNA can also confound the downstream detection method when trying to discern copy number variations such as aneuploidy. Akonni Biosystems' patented "smart-tip" technology, called TruTip®, uses a porous nucleic acid binding matrix inserted into a disposable pipette tip. This technology offers many advantages over the current extraction methods for this application, including the ability to process large volumes of plasma and enrich specific sized DNA. Its format is also amenable to automation to provide high-throughput processing capabilities. Methods: For this study, we designed a custom TruTip for large volume samples and developed a manual protocol demonstrating feasibility to extract circulating DNA from up to 10 mL of plasma. Conditions were optimized to maximize the yield of the male fetal DNA and minimize the recovery of the less fragmented female DNA in a final elution volume of 50 µL. Extraction yields were compared to Qiagen's Circulating Nucleic Acid Kit and assayed using a duplex quantitative real-time PCR assay to detect 1) male fetal DNA using a multi-copy Chrom Y assay and 2) total DNA, fetal and maternal, using a Chrom 1 assay. Results: The final optimized protocol provided for a 200-fold concentration of the plasma DNA in approximately 2.5 hrs. Real-time PCR data demonstrated similar recoveries of low levels of fetal DNA between TruTip and the Qiagen Circulating Nucleic Acid Kit, but TruTip exhibited a 3–4% higher enrichment of fetal DNA. This enrichment is highly significant when applied to detection methods, such as digital PCR and sequencing, and could allow for more accurate testing results earlier in the pregnancy than are currently available. Development is in progress to further enhance the enrichment and automate this TruTip application on the Hamilton liquid handling system for efficient, higher throughput processing.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3012T

Prenatal detection of a supernumerary dicentric chromosome 15 through cytogenetics and array CGH after fetal ultrasound findings of micrognathia and small stomach bubble. *J.L. Giordano, MS, L. Cohen, MD, MPH, V. Pulijaal, PhD, S. Gelber, MD, PhD.* Division of Human Genetics, Weill Cornell/NYP, New York, NY.

Array Comparative Genomic Hybridization (array CGH) is increasingly used in the prenatal setting to detect unbalanced chromosomal abberations in patients with abnormal ultrasound findings.A 35-year prima gravida had a routine fetal anatomy survey at 19+ weeks which suggested micrognathia and a small stomach bubble. There was normal growth, normal AFI, and no other anomalies noted. The sequential screening was within the normal limits. Pregnancy history was significant for IVF conception with ICSI using banked sperm given partner's history of Hodgkins lymphoma. Conventional cytogenetic analysis of amniotic fluid cells revealed an abnormal female karyotype with a supernumerary bisatellited marker chromosome. FISH analysis determined the marker was composed of chromosome 15 and included the Prader-Willi/Angelman critical region: 47,XX,+mar.ish idic(15)(q12)(D15Z1++,SNRPN+,PML-). Further characterization of the marker chromosome by whole genome oligonucleotide microarray confirmed a complex rearrangement consisting of a 5.70 Mb two-copy gain from the centromere of chromosome 15 to 15q13.1, comprising 39 genes including SNRPN and UBE3A, and a 1.96 Mb single copy gain, containing at least 19 genes:arr 15q11.2q13.1(20,410,890-26,106,556)x4,15q13.1q13.2(26,180,806-28,138,979)x3. Parental karyotypes were normal. This finding in the fetal DNA was consistent with a well-described idic (15) syndrome which is characterized by developmental delay and intellectual disability of moderate to profound degree, autistic behavior, epilepsy, early hypotonia, and some mild dysmorphic features. The family subsequently terminated the pregnancy after 21 weeks. While the idic(15) syndrome has been well described, this is the first report in which ascertainment occurred due to an ultrasound finding of micrognathia and a small stomach bubble. Accurate clinical interpretation of abnormal prenatal ultrasound findings requires comprehensive assessment and testing. This case illustrates the importance of using both conventional cytogenetics and aCGH to uncover atypical findings and provide information on the clinical heterogeneity of known genomic aberrations. Diagnostic information from prenatal aCGH earlier in pregnancy may improve genetic counseling and help parents make informed decisions.

3013T

Prenatal Diagnosis and Fetal Autopsy Findings in Aicardi syndrome. *M.C. Injeyan[1], I. Miron[1], P. Shannon[2], A. Malinowski[3], S. Blaser[4], D. Chitayat[1,3,5].* 1) The Prenatal Diagnosis and Medical Genetics Program, Mount Sinai Hospital, Toronto, Canada; 2) Department of Pathology and Laboratory Medicine, Mount Sinai Hospital, Toronto, Canada; 3) Department of Obstetrics and Gynecology, Mount Sinai Hospital, Toronto, Canada; 4) Department of Diagnostic Imaging, Division of Neuroradioloy, The Hospital for Sick Children, Toronto, ON, Canada; 5) Division of Clinical and Metabolic Genetics, The Hospital for Sick Children, University of Toronto, Toronto, Canada.

Aicardi syndrome is a neurodevelopmental disorder purported to have an X-linked dominant mode of inheritance. The causative gene on the X chromosome has not been identified and diagnosis is based solely on clinical features. This disorder has been classically characterized by a triad of infantile spasms, agenesis of corpus callosum (ACC), and distinctive chorioretinal lacunae. Largely due to improved imaging techniques a number of additional cerebral features have been described, which together likely represent a specific malformation complex. Features include cortical abnormalities with heterotopias, cystic lesions, optic nerve hypoplasia, and hemispheric asymmetry. Patients with Aicardi syndrome may also have costovertebral anomalies, embryonic tumors and eye abnormalities such as colobomas. We report on a patient presenting at 19 weeks gestation with fetal ultrasound findings of a interhemispheric cystic structure in the region of the cavum septum pellucidum and third ventricle. Follow-up ultrasound at 21 weeks gestation showed ACC and one or more interhemispheric cerebral cysts. Fetal MRI performed at 24 weeks gestation found complete ACC, left cortical dysplasia, periventricular heterotopia, and an interhemispheric cyst. Amniocentesis was done and aCGH was normal and female. The couple elected to interrupt the pregnancy via induction of delivery. Autopsy showed a 27 week gestation female infant with ACC, interhemispheric cystic mass arising from tela choroidea, nodular periventricular and subcortical heterotopias, and abnormal cerebral gyration. The eyes showed numerous chorioretinal defects with absence of pigmentation of the overlying pigment epithelium, staphylomas and retinal dysplasia. X-rays showed T10 sagittal clefting, T7 wedged vertebral body and abnormal facies with flattened midface, prominent infraorbital creases, upturned nose, microstomia, and micrognathia. This constellation of cerebral, ocular and vertebral abnormalities was consistent with Aicardi syndrome. Prenatal diagnosis of ACC is a counselling challenge as the prognosis ranges between normal development to severe delay. A diagnosis of Aicardi syndrome may impact reproductive decision making due to low recurrence risk. Since the fetal ultrasounds/MRI findings are not specific, an autopsy including eye histopathology should be performed if the pregnancy is interrupted and ophthalmological examination should be done postnatally if parents decide to continue the pregnancy.

3014T

High-Throughput Massively Parallel Sequencing for Fetal Aneuploidy Detection from Maternal Plasma. *T.J. Jensen[1], T. Zwiefelhofer[1], R.C. Tim[1], Z. Dzakula[1], S.K. Kim[1], A.R. Mazloom[1], G. McLennan[2], C. Deciu[1], D. van den Boom[2], M. Ehrich[2].* 1) Research and Development, Sequenom Center for Molecular Medicine, San Diego, CA; 2) Research and Development, Sequenom Inc., San Diego, CA.

Background: Circulating cell-free fetal (ccff) DNA comprises 3–20% of all cell-free DNA present in maternal plasma. Numerous research and clinical studies have described the analysis of ccff DNA using next generation sequencing for detecting fetal aneuploidies with greater than 99% sensitivity and specificity. We sought to extend these studies using semi-automated library preparation, higher multiplexing during sequencing, and improved bioinformatics tools. This will enable a higher throughput and a more cost effective assay while maintaining optimal clinical performance. **Methods:** We performed initial studies to determine technical feasibility using an increased multiplexing level as well as a semi-automated library preparation method, following the processing of more than 3,000 maternal plasma samples. Subsequently,these processes were combined and we performed a clinical evaluation for detection of fetal aneuploidy using 1,269 samples and incorporating improved bioinformatic procedures for classification of fetal aneuploidy status. **Results:** Of the 1,269 clinical samples, 1,093 were confirmed by fetal karyotype analysis to be from a euploid fetus and 176 were confirmed to be from an aneuploid fetus (trisomy 21 = 134 cases; trisomy 18 = 36 cases; and trisomy 13 = 6 cases). Detection rates were 100% for all aneuploid cases. False positive rates were .08%, 0%, and .08% for trisomy 21, 18, and 13, respectively. **Conclusions:** These data suggest that increased multiplexing and semi-automated DNA library preparation methods in concert with improved bioinformatic procedures enable higher sample throughput without adversely impacting fetal aneuploid classification accuracy.

3015T

Isolation of selected, single chromosomes from the isolated circulating fetal cells for clonal expansion, molecular profiling, and next generation sequencing. *M. Malecki[1,2,4], R. Malecki[2,3].* 1) Phoenix Biomolecular Engineering Foundation, San Francisco, CA, USA; 2) Western University of Health Sciences, Pomona, CA, USA; 3) San Francisco State University, San Francisco, CA, USA; 4) University of Wisconsin, Madison, WI, USA.

Introduction. Circulating fetal cells (CFCs) present in the maternal circulation carry a great potential for diagnosis of genetic disorders, as well as for regenerative therapy after their clonal expansion. Although, their low concentrations - at 1 fetal cell per 10^7 maternal blood cells - makes them difficult to obtain through negative selection. The specific aim of this work was genetic engineering of the superparamagnetic and fluorescent, single chain variable fragments targeting the fetal cells, which would allow us to isolate circulating fetal cells. **Methods.** We have recently developed effective methods, including superparamagnetic and fluorescent chromosome-specific probes for isolating single chromosomes. We also engineered, based upon the cDNA library, the cell surface receptor targeting superparamagnetic and fluorescent scFvs (EGFR, TfR, SSEA-3 & 4, TRA-1-60 & -1-81) for isolating the selected cells [1]. For this study, we used these novel approaches of biomolecular engineering to generate the scFv targeting OSMR and COL1 proteins, which are present exclusively on fetal cells, but not on maternal cells [2].**Results.** With the native immunoblotting and fluorescence confocal, we demonstrated the high specificity of the scFv generated by biomolecular engineering. These scFvs allowed us to capture the fetal and stem cells with the FACS and MACS, as well as analyze them with FCM, XRF, and NMR. With the liquid phase FISH, we demonstrated the specificity of the engineered chromosome-specific probes (CSPs). Superparamagnetic and fluorescent properties of these CSPs allowed us to isolate single, selected chromosomes from single fetal cells. This followed by WCA and sequencing. **Conclusion.** The scFvs and probes generated by biomolecular engineering in this project create the solid foundations for prenatal diagnosis through molecular profiling of the circulating fetal cells. **Bibliography.** [1]. Malecki M, Szybalski W. Isolation of single, intact chromosomes from single, selected ovarian cancer cells for in situ hybridization and sequencing. Gene. 2012 Feb 1;493(1):132-9. Epub 2011 Dec 1. [2] Götherström C, Chan J, O'Donoghue K, Fisk NM. Identification of candidate surface antigens for non-invasive prenatal diagnosis by comparative global gene expression on human fetal mesenchymal stem cells. Mol Hum Reprod. 2010 Jul;16(7):472-80. Epub 2010 Mar 3.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3016T

Noninvasive prenatal sex determination using massively parallel sequencing in samples from a large clinical validation study. *P. Oeth[1], A. Mazloom[1], T. Wang[1], G.E. Palomaki[3], J.A. Canick[3], A. Bombard[2], D. van den Boom[2], M. Ehrich[2], C. Deciu[1].* 1) Sequenom Center for Molecular Medicine, San Diego, CA; 2) Sequenom, Inc., San Diego, CA; 3) Women & Infants Hospital, Alpert Medical School of Brown University, Providence, RI.

Background: Early determination of fetal sex can be an important aspect of diagnosing and treating congenital X-linked disorders in the fetus including: hemophilia, X-linked mental retardation, Duchenne muscular dystrophy and congenital adrenal hyperplasia (where prenatal treatment of the female fetus can minimize or prevent virilization early in pregnancy). The current standard for fetal sex determination involves karyotyping of samples obtained via an invasive procedure (chorionic villus sampling or amniocentesis), which pose some risk of fetal loss. An equally accurate, but noninvasive approach is desirable. Many studies have demonstrated that noninvasive fetal sex determination using circulating cell-free DNA (ccfDNA) is highly accurate when using PCR-based assays or sequencing technologies. However, none have been conducted on a large clinical cohort. Materials and Methods: Plasma samples for the training cohort (Training Cohort) were part of a nested case-control study of pregnant women at high risk for fetal aneuploidy, collected between 10.5 and 20 gestational weeks (Palomaki, et al. GiM, 2011). Sequencing reads from a randomly selected subset of 1242 pregnancies, 623 male and 619 female fetuses, allowed for optimization of the classification algorithm based on normalized sequencing counts for select regions of the Y-chromosome. An independent blinded clinical validation cohort (Validation Cohort) consisting of another 397 samples was then tested (210 male and 187 female fetuses) using the Training Cohort classification algorithm. Results: The Training Cohort yielded a karyotype concordance rate of 99.4% (95% CI: 98.9–99.7%); eight discordant calls on 5 male and 3 female samples were made, respectively. The Validation Cohort yielded a karyotype concordance of 100% (95% CI: 99.0–100.0%). The combined concordance of all samples in both cohorts was 99.5% (95% CI: 99.0–99.8%) with a total of eight discordant classifications based on reported karyotypes. Conclusion: Noninvasive fetal sex determination using massively parallel sequencing on ccfDNA purified from maternal plasma demonstrated a high predictive rate compared to the reported karyotype in these clinical cohorts. This noninvasive molecular karyotyping methodology correctly classifies fetal sex for at least 99% of samples tested and can aid in the detection and management of inherited congenital sex-linked, X-linked, and autosomal recessive genetic disorders of the fetus.

3017T

MeDIP real time qPCR of maternal peripheral blood reliably identifies trisomy 21. *P. Patsalis[1], E. Tsaliki[1,2,3], C. Spyrou[2], G. Koumbaris[1,2], E. Kypri[2], S. Kyriakou[2], C. Sotiriou[2], E. Touvana[2], A. Keravnou[2], A. Karagrigoriou[4], K. Lamnissou[3], V. Velissariou[5], C. Sismani[1], E.A. Papageorgiou[1,2].* 1) Cytogenetics & Genomics, Cyprus Inst Neurology Gen, Nicosia, Cyprus; 2) NIPD Genetics Ltd, P.O.Box 27954, 2434, Nicosia, Cyprus; 3) Department of Genetics and Biotechnology, Faculty of Biology, School of Science, National and Kapodistrian University of Athens, Panepistimiopolis, 15771, Ilissia, Athens, Greece; 4) Department of Mathematics and Statistics, University of Cyprus, P.O. Box 20537, Nicosia, Cyprus; 5) Department of Genetics and Molecular Biology, Mitera Hospital, 15123, Athens, Greece.

Objective: One of the most promising technologies for non-invasive prenatal diagnostic (NIPD) of trisomy 21 is the application of the Methylated DNA Immunoprecipitation (MeDIP) real-time qPCR-based approach developed by Prof. Philippos Patsalis and his team. Thousands of Differentially Methylated Regions (DMRs) were generated and the best were used for the development and validation of the first universal non-invasive prenatal diagnosis for trisomy 21 (NIPD21), using MeDIP real-time qPCR for chromosome 21. To re-evaluate the efficiency of the 12 DMRs used in the MeDIP real time qPCR based approach, develop an improved version of the diagnostic formula and perform a larger validation study. Methods: Twelve selected DMRs were checked for CNVs in the DGV database. The DMRs located within CNVs were excluded from the analysis. One hundred and seventy five maternal peripheral blood samples were used to re-construct and evaluate the new diagnostic formula and for a larger scale blinded validation study using MeDIP real time qPCR. Results: Seven DMRs entered the final model of the prediction equation and a larger blinded validation study demonstrated 100% sensitivity and 99.2% specificity. No significant evidence for association was observed between cffDNA concentration and D value. Conclusion: MeDIP real time qPCR method for NIPD of trisomy 21 was confirmed and re-validated in a 175 samples with satisfactory results demonstrating that it is accurate and reproducible. We are currently working towards simplification of the method to become more robust and therefore easily, accurately and rapidly reproduced and adapted by other laboratories. Nevertheless, larger scale validation studies are necessary before the MeDIP real time qPCR based method could be applied in clinical practice.

3018T

Highly accurate non-invasive detection of fetal aneuploidy of chromosomes 13, 18, 21, X and Y by targeted sequencing. *M. Rabinowitz[1], G. Gemelos[1], M. Hill[1], B. Levy[2], S. McAdoo[1], M. Savage[1], Z. Demko[1].* 1) Natera, Inc., San Carlos, CA; 2) Department of Pathology and Cell Biology, Columbia, New York.

Objective: To develop a non-invasive prenatal aneuploidy test based on analysis of cell-free DNA in maternal blood that is capable of detecting all relevant whole chromosomal abnormalities of chromosomes 13, 18, 21, X and Y with high accuracy.

Materials and Methods: 176 maternal plasma samples (including 154 euploid and 22 aneuploid samples) were collected from patients at greater than 9 weeks of gestation under an institutional review board (IRB) approved protocol. Cell free DNA was isolated, and a targeted multiplex PCR amplification of 11,000 loci on 13, 18, 21, X and Y was performed. Sequencing data was analyzed using novel Parental Support (PS) technology that uses Bayesian statistics to analyze multiple copy number hypotheses and determine the Maximum likelihood a posteriori (MAP) hypothesis given the sequencing data. Unlike other reported methods, the PS method determines confidence, or estimated accuracy, for each sample. Similar confidences are produced for each of the five chromosomes, and can be incorporated into a DNA quality threshold metric to enable draws early in the pregnancy without increasing the probability of false positives or negatives.

Results: 155 samples passed a DNA quality threshold, and the PS approach yielded a call at all five chromosomes for these cases; all calls in this group were correct (775/775) including 12 trisomy 21, 3 trisomy 18, 2 trisomy 13, 1 monosomy X, and 2 XXY samples.Conclusions: The PS informatics approach is able to detect fetuses with a chromosomal abnormality from a maternal blood sample with high accuracy at chromosomes 13, 18, 21, X and Y. This method performs equally well at all chromosomes, does not require the use of a reference chromosome, and calculates a confidence call that is personalized for each sample.

3019T

Accuracy of prenatal diagnosis in elective termination of pregnancy. *J. Saraiva[1,2,3], F. Ramos[1,2], S. Maia[1], M. Branco[4], J. Raposo[5], J. Sá[1], S. Sousa[1], M. Venâncio[1,2], R. Pina[5], E. Galhano[4], L. Ramos[1].* 1) Servico Genetica Medica, Hospital Pediatrico, CHUC, Coimbra, Portugal; 2) Faculdade de Medicina da Universidade de Coimbra, Portugal; 3) Compete, Projecto PEst-C/SAU/UI3282/2011, IBILI, Coimbra, Portugal; 4) Centro de Diagnóstico Pré-natal, Maternidade Bissaya Barreto, CHUC, Coimbra, Portugal; 5) Serviço de Anatomia Patológica, Centro Hospitalar de Coimbra, CHUC, Coimbra, Portugal.

Objectives To evaluate the quality of the results of prenatal results in all cases of abortion due to fetal abnormalities. Method All 385 cases of therapeutic abortion due to fetal abnormalities recorded in a tertiary prenatal diagnosis centre from 2000 to 2007 were evaluated. Data from the prenatal results and post-mortem autopsy and genetics consultation was collected and compared regarding the abnormalities identified, the etiological diagnosis and its prognosis. The diagnosis were classified according to the etiology (Wellesley et al.) and the EUROCAT groups. Results In 81 of the 385 fetus (21%) there was full agreement between the prenatal data and the post-mortem examinations. In 81 fetus (21%) the etiologic diagnosis was changed, The results of the post-mortem examinations changed the prognosis in only 1 fetus. Conclusion In the 385 abortions due to fetal abnormalities the prenatal evaluation was excellent regarding the assessment of the prognosis as it was wrong in only 1 case. However the post-mortem examinations were useful as they identified further abnormalities in 79% and made possible a more precise or different diagnosis in 21% of the cases. [Compete, Projecto PEst-C/SAU/UI3282/2011].

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3020T

Non-Invasive Chromosomal Evaluation (NICE) Study: Results of a Multicenter, Prospective, Cohort Study for Detection of Fetal Trisomy 21 and Trisomy 18. *K. Song[1], H. Brar[2], J. Weiss[3], A. Karimi[4], L.C. Laurent[5], A.B. Caughey[6], M.H. Rodriguez[7], J. Williams, III[8], M.E. Mitchell[9], C.D. Adair[10], H. Lee[11], B. Jacobsson[12], M.W. Tomlinson[13], D. Oepkes[14], D. Hollemon[1], A.B. Sparks[1], A. Oliphant[1], M.E. Norton[15], NICE Study Group.* 1) Ariosa Diagnostics, San Jose, CA; 2) Riverside Perinatal Diagnostic Center, Riverside, CA; 3) East Bay Perinatal Medical Associates, Oakland, CA; 4) Institute of Prenatal Diagnosis and Reproductive Genetics, San Gabriel, CA; 5) University of California - San Diego, San Diego, CA; 6) Oregon Health Sciences University, Portland, OR; 7) San Gabriel Valley Perinatal Medical Group, West Covina, CA; 8) Cedars-Sinai Medical Center, Los Angeles, CA; 9) Medical College of Wisconsin, Milwaukee, WI; 10) University of Tennessee, Chattanooga, TN; 11) University of California - San Francisco, San Francisco, CA; 12) Sahlgrenska University Hospital, Gothenburg, Sweden; 13) Northwest Perinatal Center, Portland, OR; 14) Leiden University Medical Centre, Leiden, Netherlands; 15) Stanford University/Lucile Packard Children's Hospital, Stanford, CA.

Background: Several groups have demonstrated in case-control studies, the ability to detect fetal trisomy 21 (T21) and trisomy 18 (T18) via analysis of cell-free DNA (cfDNA) using massively parallel DNA shotgun sequencing (MPSS). MPSS randomly analyzes cfDNA from the entire genome, resulting in higher cost and complexity than is currently practical for widespread clinical adoption. **Objective:** To evaluate performance of a directed cfDNA analysis method for detection of fetal T21 and T18. **Methods:** A multicenter, international, prospective cohort study was performed. Pregnant women of at least 10 weeks gestational age with a singleton pregnancy who underwent invasive testing for any reason were enrolled. A directed cfDNA analysis method was used to selectively analyze chromosomes 21 and 18 from maternal plasma of enrolled subjects. An algorithm was utilized that incorporated chromosomal cfDNA counts, fetal fraction, and maternal and gestational age to calculate a trisomy risk score for each sample. Samples were classified as High Risk or Low Risk for a given trisomy based on a 1 in 100 (1%) risk score cut-off. **Results:** All eligible subjects underwent analysis. Of the 81 T21 cases, all were classified as High Risk for T21 and there was one false positive among the 2888 normal cases, for a sensitivity of 100% (95% CI: 95.5–100%) and a false-positive rate of 0.03% (95% CI: 0.002–0.20%). Of the 38 T18 cases, 37 were classified as High Risk and there were 2 false positives among the 2888 normal cases, for a sensitivity of 97.4% (95% CI: 86.5%–99.9%) and a false-positive rate of 0.07% (95% CI: 0.02–0.25%). The presence of chromosomal abnormalities other than T21 or T18 did not affect the analysis of T21 or T18 risk. **Conclusions:** Directed cfDNA analysis and application of an individualized risk algorithm is an efficient and effective method for the detection of fetal T21 and T18.

3021T

Detection of fetal sub-chromosomal alterations in maternal plasma cell-free DNA (cfDNA) using massively parallel sequencing (MPS). *A. Srinivasan, X. Li, J. Sprowl, A. Huang, A. Sehnert, R. Rava.* Verinata Health, Inc., Redwood City, CA.

Objective: It has been well established that massively parallel sequencing (MPS) of cfDNA isolated from the plasma of pregnant women can be utilized to accurately detect whole chromosome fetal aneuploidies. The objective of this study is to further investigate and demonstrate detection by MPS of fetal chromosomal copy number changes at a sub-chromosomal level. Methods: To establish an algorithm to detect fetal sub-chromosomal aneuploidies, maternal-fetal DNA mixtures were generated by combining sheared genomic DNA (gDNA) from beta-lymphocytes of a child with Greig Cephalopolysyndactyl Syndrome caused by a 16 Mbp deletion in Chromosome 7 with sheared gDNA from beta-lymphocytes of the mother. Maternal-child mixtures containing 5% or 10% gDNA from the deletion sample were analyzed utilizing MPS with 36-mer sequence tags that were mapped to approximately 600 million unique sites on the hg19 build of the human genome. The sites were counted in non-overlapping 1 Mbp bins across the genome at approximately 190 thousand sites per bin. An algorithm was developed to quantitatively estimate a “fetal fraction” in each bin relative to the normalized counts from unaffected samples. Signal of the estimated fetal fraction above the background noise threshold was applied to detect amplifications or deletions at sub-chromosomal levels. The method was optimized on the prepared mixtures and then applied to clinically obtained samples from pregnant women with known fetal sub-chromosomal alterations. Results: The analyzed MPS data quantitatively identified the “fetal” 16 Mbp deletion on Chromosome 7 in the prepared maternal-child mixtures. The method was also able to detect an amplification of approximately 35 Mbp on Chromosome 6 in a clinical maternal sample where microarray data from chorionic villus sampling confirmed this finding in the fetus. Additional samples are now being tested. Conclusion: MPS data from cfDNA isolated from the plasma of pregnant women can be analyzed to detect fetal chromosomal alterations at a sub-chromosomal level and provide further karyotype information beyond whole chromosome aneuploidies.

3022T

Approach to prenatally diagnosed esophageal atresia/tracheoesophageal fistula. *B. Suskin[1,3], S. Klugman[1,3], T. Goldwaser[1,3], A. Roe[2,3], K. Bajaj[1,2,3].* 1) Obstetrics and Gynecology, Montefiore Medical Center, Bronx, NY; 2) Department of Obstetrics & Gynecology, North Bronx Health Network, Bronx, NY; 3) Albert Einstein College of Medicine, Bronx, NY.

BACKGROUND: Esophageal atresia (EA) describes a condition in which there is abnormal development of the upper gastrointestinal tract. EA can be isolated or associated with other anomalies including tracheoesophageal fistula (TEF), an abnormal connection between the esophagus and trachea. With an overall birth incidence of 1:3500 live births, only 44–56%; of cases are diagnosed prenatally. Management of prenatally diagnosed EA/TEF involves a multi-disciplinary approach. CASE: 31 year-old multiparous woman, without significant medical or family history, was found to have an appropriately grown female fetus with a small stomach at the time of anatomic survey at 19 weeks gestation. No other anomalies were seen. Subsequent sonograms, including a fetal echocardiogram, did not reveal additional structural anomalies. The patient underwent genetic counseling and opted to undergo an amniocentesis which revealed 46, XX, inv (4) (p12q21.1). The paternal karyotype demonstrated the same inversion. At 21 weeks gestation, polyhydramnios was present. The patient has been linked with an interdisciplinary team including clinical geneticists, neonatalogists and pediatric surgeons. [Patient has not yet delivered]. DISCUSSION: Esophageal atresia/tracheoesophageal fistula can be isolated or may be associated with other anomalies. When an absent or small fetal stomach is found, detailed ultrasound should be performed as up to 60% of EA's are seen with other anomalies. Thorough genetic counseling is indicated as EA can be associated with a wide range of genetic aberrations. The presence of additional anomalies and an underlying genetic etiology significantly impacts prognosis. We highlight the sonographic diagnosis of EA/TEF. Furthermore, we will also review common heritable etiologies as well as genetic counseling/management approaches to EA/TEF with an emphasis on currently available diagnostic testing.

3023T

Quantification of fetal DNA in maternal plasma by massively parallel sequencing using fetal-specific methylation markers. *M. Tang[1], J. Tynan[1], S. Sovath[1], A. Maer[1], G. Hogg[1], D. van den Boom[2], P. Oeth[1].* 1) Sequenom Center for Molecular Medicine, 3595 John Hopkins Ct., San Diego, CA 92121; 2) Sequenom, Inc., 3595 John Hopkins Ct., San Diego, CA 92121.

BACKGROUND: The presence of circulating cell-free (ccf) fetal DNA in maternal plasma allows for noninvasive prenatal diagnostic (NIPD) applications such as detection of fetal trisomies or monogenic diseases. Given the low fractional concentration of ccf fetal DNA in maternal plasma, metrics are needed to ascertain if sufficient ccf fetal DNA is present in order to accurately determine the fetal genetic status. We previously developed a MALDI-TOF Mass Spectrometry (MS) based fetal quantification assay (FQA), using fetal-specific methylation DNA markers (Nygren et. al., 2010, Clin Chem 56: 1627-35). Here we extended the FQA to its detection by massively parallel sequencing (MPS). **METHOD:** The unmethylated maternal fraction of ccf DNA was digested by methylation-sensitive restriction enzymes, and the undigested methylated fetal DNA fraction was amplified in the presence of synthetic oligonucleotides to permit competitive PCR. The amplicons were ligated to sequencing adaptors via two protocols. One directly ligated the 3' A-overhang amplicons to sequencing adaptors with 3' T-overhangs. The other utilized PCR primers containing deoxyuridine that were selectively cleaved by uracil glycosylase after PCR to generate specific overhangs, allowing sequencing adaptors to ligate with high efficiency. The resulting libraries were sequenced on HiSeq™ 2000 (Illumina, San Diego CA). Copy numbers of individual marker were calculated from the ratio of sequence reads uniquely mapped to the target regions with expected reference or competitor alleles, and fetal fraction was derived by comparing fetal DNA markers to total DNA markers. **RESULTS:** We sequenced FQA libraries from 48 maternal plasma samples. When compared to those detected by MALDI-TOF MS, copy numbers of individual markers were highly correlated (R^2>0.98), although platform-specific bias in allele ratio was observed in certain targets. Importantly, fetal fractions were consistent between MPS and MALDI-TOF MS detection. Furthermore, calculated fetal fraction did not change when FQA libraries were sequenced with other whole genome libraries in multiplex setting. **CONCLUSION:** We demonstrated feasibility of targeted sequencing of methylated DNA and its application to accurately detect fetal fraction in ccf DNA from maternal plasma. Combined sequencing of FQA and other NIPD application libraries allows NIPD test results be called with higher confidence.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3024T

Curated gene set outperforms GWAS Data on Pathway-Based Genetic Analysis. *J. Padbury[1,2], A. Uzun[1,2], A. Dewan[3], S. Istrail[2].* 1) Pediatrics, Women & Infants Hospital, Providence, RI; 2) Brown University, Providence, RI; 3) Yale University, New Haven, CT.

Genome-wide association studies (GWAS) have the potential of identifying novel genetic variants. They have become a popular approach to the investigation of complex diseases but they have had only a measured success. Complex mechanisms of gene-gene interactions underlie the pathogenesis of most complex diseases. We are interested in the genetic contribution(s) to preterm birth. Preterm birth is an important, poorly understood clinical problem. The incidence of preterm birth (PTB) in the United States is now 12.7%. Multiple genes, gene networks and variants have been associated with PTB. Simple patterns of inheritance are not sufficient enough to explain the pathogenesis of the majority of preterm births. We have developed an approach for identifying a parsimonious set of candidate genes for the study of preterm birth validated by a priori biological information. We took the genes involved in preterm birth from the published literature and from available expression arrays. In addition to these initial steps we included pathway analysis to impute additional genes likely to be involved in these pathways. Results from a genome-wide study of preterm birth, “GENEVA,” became available in dbGAP. The dataset includes phenotypic information (i.e. gestational age) and complete genotype data on nearly 4000 patients. We used our curated set of genes to analyze the GENEVA dataset. The results of the refined candidate genes were further analyzed by gene set enrichment analysis. We identified a total of 30 pathways with high confidence values (FDR<0.05) based on our curated set of genes. From a whole genome based pathway analysis 39 pathways with high confidence were identified. There were 17 shared pathways which were significant and the analysis based on curated genes showed greater significance in 15 out of the 17 shared pathways. Interestingly, most of the pathways involving inflammation were either shared or identified only by curated set of genes based pathway analysis. In conclusion the curated set of genes gave much stronger associations than the genome wide analysis in all but a few of these pathways. These results provide important confirmation of the role of genetic architecture in the risk of preterm birth. The results of the pathway analysis show highly significant associations and suggest several, important mechanistic implications.

3025T

Analysis of etiology and complications among 1,128, including 493 iatrogenic and 635 spontaneous, Chinese preterm births. *M. Xiao[1], W. Ren[1], N. Zhong[2].* 1) Dept. Obstetrics, Hubei Provincial Maternal and Children's Hospital, Wuhan, Hubei, China; 2) New York State Institute for Basic Research in Developmental Disabilities.

Objective: To investigate the singleton iatrogenic preterm birth and perinatal outcome. Methods: To retrospectively, from January 2007 to December 2011, analyze the etiology of iatrogenic preterm birth, which is compared to that of singleton spontaneous preterm birth, and to analyze the mode of delivery and perinatal outcome. A total number of 493 iatrogenic preterm births versus 635 cases of singleton spontaneous preterm birth in the same period were studied. Results: The etiology for the 493 cases of iatrogenic preterm birth studied is in an order of placenta previa (n=169, 34.3%), hypertensive disorders in pregnancy (n=152, 30.8%), intrahepatic cholestasis of pregnancy (n=74, 15%), placental abruption (n=47, 9.5%), fetal distress (n=35, 7.1%), and oligohydramnios (n=16, 3.2%). Compared to the singleton spontaneous preterm birth, the following rate among singleton iatrogenic preterm births was significantly ($p<0.05$ to $p<0.001$) increased: cesarean section (99.2% versus 11.7%), neonatal mortality (2.6% versus 0.9%), neonatal asphyxia (13.6% versus 6.8%), aspiration pneumonia (27.8% versus 15.4%),and neonatal respiratory distress syndrome (13.6% versus 6.8%). No statistical difference ($p>0.05$) was found in the rate of neonatal necrotizing enterocolitis (3.9% versus 3.6%) and of neonatal jaundice (44.4% versus 36.5%). Conclusion: iatrogenic preterm births is a severe condition that influences maternal and infant health. Active prevention and intervention of maternal health condition may have an important impact on the perinatal outcome.

3026T

Admixture Mapping Identifies Susceptibility Loci Associated with Preterm Birth in African Americans. *H. Tsai[1,2], X. Liu[1,3], X. Hong[4], G. Wang[4], T. Chang[2], Q. Chen[1], C. Pearson[5], D. Caruso[4], L. Heffner[6], B. Zuckerman[5], X. Wang[4].* 1) Dept Pediatrics, Northwestern Univ Sch Med, Chicago, IL, USA; 2) Division of Biostatistics and Bioinformatics, Institute of Population Health Sciences, National Health Research Institutes, Taiwan; 3) Department of Preventive Medicine, Feinberg School of Medicine, Northwestern University; Chicago, IL, USA; 4) Center on the Childhood Origins of Disease, Department of Population, Family and Reproductive Health, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA; 5) Department of Pediatrics, Boston University School of Medicine and Boston Medical Center, Boston, MA, USA; 6) Department of Obstetrics and Gynecology, Boston University School of Medicine and Boston Medical Center, Boston, MA, USA.

Objectives: Remarkable ethnic disparities have long been observed in preterm birth (PTB), particularly between African Americans (17.8%) and European Americans (11.5%). We screened and identified genomic regions associated with PTB using admixture mapping methods in 1,106 African American women among PTB cases and term controls. In addition, we also explored susceptibility loci associated with PTB subgroups, specifically spontaneous PTB. Methods: A total of 1,106 African American women were genotyped using the Illumina African American Panel, consisted of 1,509 ancestry informative markers. We performed admixture mapping analysis using the statistical methods implemented in the ANCESTRYMAP program. Results: The highest peak was observed on rs4782812 locating on the chromosomal region of 16q23.3 for PTB (LOD score = 1.80) and for spontaneous PTB (LOD score = 1.78), respectively. In addition, the chromosomal region of 7q22 with LOD score equal to 1.66 was observed in PTB. Fourteen known genes were found within the 7q22.1 region, including 2 genes, PDGFA (platelet-derived growth factor alpha polypeptide) associated protein 1 and ATP synthase, H+ transporting, mitochondrial Fo complex, subunit F2, which are potentially functional relevant to PTB. Similarly, fifteen known genes were found within the 16q23.3 region, including ATPase, Ca++ transporting, type 2C, member 2 and oxidative stress induced growth inhibitor 1, which are potentially functionally-associated with PTB. Conclusion: Our findings suggest that potential novel PTB related disease genes reside within the identified peak regions (7q22.1 and 16q23.3, respectively) in the inner-city study population. Further investigation of these promising regions, including fine-mapping and sequencing, is warranted.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3027T

Single Nucleotide Polymorphism in Toll-Like Receptor 6 is Associated with Risk for *Ureaplasma* Respiratory Tract Colonization and Bronchopulmonary Dysplasia in Preterm Infants. *A.H. Winters[1], T.D. LeVan[5,6], S.N. Vogel[2], K.L. Chesko[3], T.I. Pollin[1,4], R.M. Viscardi[3].* 1) Human Genetics, University of Maryland, Baltimore, Baltimore, MD; 2) Microbiology and Immunology, University of Maryland, Baltimore, Baltimore, MD; 3) Pediatrics, University of Maryland, Baltimore, Baltimore, MD; 4) Endocrinology, Diabetes and Nutrition, University of Maryland, Baltimore, Baltimore, MD; 5) Departments of Internal Medicine/Pulmonary and Epidemiology Nebraska Medical Center, Omaha, NE; 6) VA Nebraska-Western Iowa Health Care System Research Service, Dept. of Veterans Affairs Medical Center, Omaha, NE.

Introduction: Respiratory colonization with the genital mycoplasmas *Ureaplasma parvum* and *U. urealyticum* is associated with pulmonary inflammation and an increased risk for the chronic lung condition bronchopulmonary dysplasia (BPD) in preterm infants. We hypothesized that variants in genes regulating the inflammatory response, including toll-like receptors (TLRs), are associated with altered risk for *Ureaplasma* respiratory colonization and BPD. *Ureaplasma* is known to bind to TLRs, specifically heterodimers of TLR2 with TLR1 or TLR6, and cause downstream activation of the transcription factor NF-κB and upregulation of gene expression of pro-inflammatory mediators. *Methods:* 35 common SNPs from *CD14, IL6, IL8, MyD88, NOD2, TLR1, TLR2, TLR4,* and *TLR6* were assayed in 254 preterm infants <33 weeks gestation cultured for *Ureaplasma* species and evaluated for BPD. The majority of subjects (N=191[75%]) were African-American, and the remaining subjects were non-Hispanic Caucasian (N=63 [25%]). *Results:* SNPs in *NOD2* (rs2066844, OR=7.3 [1.46–36.52], total sample), *TLR2* (rs1439164, OR=3.7 [1.32–10.04] and rs3804099, OR=0.42 [0.18–0.96], non-Hispanic Caucasians), and *TLR6* (rs5743827, OR=0.54 [0.34–0.86], total sample and OR=0.55 [0.32–0.95], African-Americans; rs5743788, OR= 2.68 [1.11–6.43], non-Hispanic Caucasians) were significantly associated with *Ureaplasma* respiratory tract colonization. SNPs in *CD14, IL-8,* and *TLRs 2, 4,* and *6* were associated with BPD. *TLR6* SNP rs5743827 was associated with both *Ureaplasma* respiratory tract colonization and BPD, and there was evidence of an interaction between genotype and*Ureaplasma* colonization. Compared to infants without either risk factor, infants with both the risk genotype (GG) at rs5743827 and *Ureaplasma* colonization have a greater increased odds of BPD (OR=3.46 [1.58–7.61]) than would be expected from addition of the risks of genotype and *Ureaplasma* colonization (OR=1.33 [0.61–2.94] and OR=1.26 [0.46–3.45] respectively).*Discussion:* Polymorphisms in host defense genes may alter susceptibility to *Ureaplasma* infection and severity of the inflammatory response, which contributes to BPD. Specifically, the GG genotype at *TLR6* SNP rs5743827 was associated both with increased risk for *Ureaplasma* respiratory tract colonization and with BPD, particularly in *Ureaplasma*-colonized infants. These observations implicate host genetic susceptibility as a major factor in BPD pathogenesis in *Ureaplasma*-infected preterm births.

3028T

Genome-wide association study of preterm delivery in mothers. *F. Geller[1], B. Feenstra[1], M. Melbye[1], R. Myhre[2], S. Myking[2], The Early Growth Genetics Consort.[3], T. I. A. Sørensen[4], P. Magnus[2], M. L. Marazita[5], J. C. Murray[6], B. Jacobsson[2].* 1) Statens Serum Institut, Copenhagen, Denmark; 2) Norwegian Institute of Public Health, Oslo, Norway; 3) c/o M. McCarthy, Wellcome Trust Centre for Human Genetics, Oxford, UK; 4) Copenhagen University Hospital, Copenhagen, Denmark; 5) University of Pittsburgh, PA; 6) University of Iowa, IA.

Preterm delivery is a condition that puts mother and child at risk, additionally to the possible perinatal complications long term developmental problems are frequently observed. Based on two genome-wide association studies from Denmark and Norway, each starting out with 1,000 mothers with preterm delivery and 1,000 control mothers with term delivery, we tried to replicate SNPs with p-values below 0.0001 in additional cohorts of European descent from the Early Growth Genetics Consortium. We investigate five different phenotypes, preterm delivery (all deliveries before 37 weeks of gestation), premature rupture of membranes, spontaneous preterm labor without rupture of membranes, early preterm delivery (before gestational week 34), and gestational age as a quantitative trait. So far four additional cohorts have provided results based on more than 10,000 individuals. Preliminary results suggest that genetic regions relevant for fasting glucose levels, age at menarche and growth also play a role in preterm delivery.

3029T

Evidence of SNP variation in the folic acid metabolic pathway associates with preterm birth (PTB). *Y. Chen[1], B-J. Wang[1], M-J. Liu[1], Y. Wang[1], J. Mao[1], S-N. Wang[1], J-R. Dai[1], H. Li[1], N. Zhong[2].* 1) Ctr Reproduction and Genetics, Suzhou Municipal Hospital, Suzhou, Jiangsu, China; 2) New York State Institute for Basic Research in Developmental Disabilities, Staten Island, NY.

Preterm birth (PTB; birth before 37 weeks in humans) is one of the most common conditions in perinatal health. It is the leading cause for neonatal death. The rate of PTB is increasing globally. Presently, the etiology and pathogenesis are yet unclear. It has been proposed that metabolism of folic acid might have some pathophysiological role in the PTB and outcome of related pregnancy. We hypothesize that the genetic variation of folic acid metabolism may associate with PTB. To test which, a subset of single nucleotide polymorphisms (SNPs), selected from 12 genes/loci involved in the folic metabolic pathway, were subjected to SNaPshot analysis in a case-control study. Fourteen SNPs (CBS-C699T, DHFR-c594+59del19, FOLH1-T1561C, GST01-C428T, MTHFD1-G878A, MTHFD-G1958A, MTHFR-A1298C, MTHFD-C677T, MTR-A2756G, MTRR-A66G, NAT1-C1095A, NFE2L2-ins1+C11108T, RFC1-G80A, SHMT1-C1420T, TCN2-C776G and TYMS-1494del6) were simultaneously tested, within a single PCR reaction, among 527 DNA samples that include 340 from PTBs and 187 controls. Four SNPs from this subset of loci were found to associate with PTBs, with a p value significantly less than 0.05 to less than 0.001. Our study demonstrate that a panel of SNP variations related to the folic acid metabolism pathway could be screened through a flexible, fast and inexpensive procedure with a purpose of analyzing their association to PTB, as we described elsewhere. In addition to the fact that folic acid metabolism pathway is associated with birth defects including neural tube defects and congenital heart diseases, we further provided evidence here that SNP variations may also associate with PTB. Testing with a larger sample size of PTBs and controls are underway to validate the potential application that these loci may be applied as a clinical biomarker for prenatal screening for risk of PTB in a longitudinal cohort study.

3030T

Investigation of genetic risk factors for chronic adult diseases in preterm birth. *N. Falah[1], J. McElroy[1], V. Snegovskikh[2], C. Lockwood[2], E. Kuczynski[2], E. Norwitz[2], J. Murray[3], R. Menon[4], K. Teramo[5], L. Muglia[1], T. Morgan[1].* 1) Pediatrics, Div Med Gen & Genomic Med, Vanderbilt Univ, Nashville, TN; 2) Department of Obstetrics and Gynecology, Yale University School of Medicine, New Haven, CT; 3) University of Iowa; 4) The Perinatal Research Center, Nashville, TN; 5) Departments of Obstetrics and Gynecology, University of Helsinki, Helsinki, Finland,.

Background: Preterm birth (PTB) is the leading cause of infant mortality. The putative pathogenetic pathways of preterm birth overlap with those of adult cardiovascular, immune and metabolic disorders (CIMD), with common mechanisms including inflammation, immunotolerance, thrombosis, and nutrient metabolism. Whereas many genetic factors for CIMD have been identified, progress in PTB has lagged behind. We hypothesized that highly validated genetic risk factors for CIMD (P < 1 × 10-30) may also be associated with PTB, via pleiotropic effects. Result: Case-control study of women with spontaneous preterm delivery (n=673) versus term delivery (n=1119). There were four diverse cohorts (US White, US African-American, US Hispanic, and Finnish). Of 35 independent SNPs genotyped, there were 13 statistically significant associations (P < 0.05), which was more than expected (binomial test; P = 0.02). In US White mothers (307 cases/342 controls), the G allele of HLA-DQA1 (A/G) rs9272346 was protective for PTB in the initial discovery cohort (P = 0.02; OR = 0.65; 95% CI: 0.46, 0.94). This protective association replicated (P = 0.02; OR=0.85(95% CI; 075- 0.97) nominally in the Danish National Birth Cohort (883 cases, 959 controls), but lost significance upon multiple testing correction. Conclusion: We observed more statistically significant associations between CIMD-related genetic risk variants and PTB than expected, suggesting that chance is an unlikely explanation for one or more of the associations. In particular, a protective association of the G allele of HLA-DQA1 was found in two independent cohorts of mothers, and in previous studies, this same allele was found to protect against type I diabetes (meta-analysis P value 5.52 × 10-219). Previous investigations have implicated HLA phenotypic variation in recurrent fetal loss as well as in chronic chorioamnionitis. Given the limited sample size in the present study, we suggest larger studies to further investigate possible HLA genetic involvement in PTB.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3031T

Maternal SNPs in the p53 pathway: risk factors for aneuploidy 21? *A.P.C. Brandalize[1, 2, 3], J.A. Boquett[1, 2, 3], L.R. Fraga[1, 2, 3], L. Schuller-Faccini[1, 2, 3].* 1) Genetics, UFRGS, Porto Alegre, Rio Grande do Sul, Brazil; 2) Instituto de Genética Médica Populacional, INAGEMP, Rio Grande do Sul, Brazil; 3) Programa de Pós-Graduação em Genética e Biologia Molecular, UFRGS, Porto Alegre, Rio Grande do Sul, Brazil.

p53 family and its regulatory pathway play an important role as regulators of developmental processes, limiting the propagation of aneuploid cells. Its dysfunction or imbalance can lead to pathological abnormalities in humans. Our working hypothesis was based on two evidences: 1) The loss of p53 function or genes that regulate this metabolic pathway may be related to the accumulation of aneuploid cells and then decrease the action of apoptotic mechanisms that would eliminate aneuploid embryos in women with wild alleles, increasing the risk of DS in women with these polymorphisms and; 2) Loss of p73 function due to polymorphisms in its encoding gene may interfere in the control of the meiotic spindle during oogenesis, increasing the risk of aneuploidy in the offspring. Based on these hypothesis we analyzed through real time PCR (using taqman probes) the role of TP53 R72P polymorphism (rs1042522), TP73 G4A (rs2273953) and C14T (rs1801173), MDM2 SNP309 (rs2279744), MDM4 T>C in intron 9 (rs1563828) and USP7 G>A in intron 25 (rs1529916) as maternal risk factors for 21 aneuploidy. A case-control study was conducted with 263 mothers of DS children and 196 control mothers. The distribution of these genotypic variants was similar between case and control mothers. However, the combined alleles TP53 C and MDM2 G, and TP53 C and USP7 A increased the risk of having offspring with DS (OR = 1.84 and 1.77; P < 0.007 and 0.018, respectively). These results indicate a synergistic effect between genes that act in the same pathway in a multifactorial way. The allele P72 reduces the efficiency of p53 to induce apoptosis. The G allele increases the expression of MDM2, degrading more p53 and negatively influencing the induction of apoptosis as well. It is known that a large amount of p53 protein is produced by the human placenta in abnormal pregnancies, so p53 can be an important factor in the pathogenesis of diseases through the induction of throphoblast apoptosis. The interaction of these polymorphisms could decrease the levels of the pro-apoptotic p53 protein, making it less functional in response to cell damage. As a consequence, the reaction would be attenuated by trophoblastic apoptosis and promote greater tolerance of aneuploidy. This work is the first to establish a relationship between polymorphisms in the TP53 gene family and its regulatory pathway as risk factor for aneuploidy of 21. Future studies in other populations should be conducted to confirm our findings.

3032T

Disomy 21 mosaicism in sperm. *E. Iwarsson[1], U. Kvist[2], M.A. Hultén[1,3].* 1) Dept of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden; 2) Centre for Andrology and Sexual Medicine, Karolinska University Hospital Huddinge, Stockholm, Sweden; 3) Warwick Medical School, University of Warwick, United Kingdom.

The mechanism underlying the low paternal origin of trisomy 21 Down syndrome (less than 10%) is still unknown. To elucidate this question we have now initiated fluorescent in situ hybridisation (FISH) studies to record the copy number of chromosome 21 at different stages of spermatogenesis. We here present data on the incidence of disomy 21 sperm mosaicism by FISH analysis in a sample of 11 men with normal spermiograms. Due to the risk of false positive and false negative signals using a single FISH probe we have applied two chromosome 21q probes. We review previous published work on the incidence of disomy 21 in sperm samples and conclude that all previous studies in this respect have been applying a single chromosome 21 probe only. Thus, our study is the first using more than one probe on chromosome 21 for the reliable establishment of disomy 21 in sperm. The average incidence of chromosome 21 disomy in the sperm samples from our cohort of 11 men was 0.13%, with a range of 0.00%–0.25%. On this score we may therefore expect around 1 per 1000 conceptions to be trisomic for chromosome 21 of paternal origin.

3033T

DiGeorge sequence due to uncontrolled maternal gestational diabetes. *A. Ludtke[1], T. Mucci[2], M. Aquino[2], E. Graber[3], L. Mehta[1].* 1) Department of Genetics and Genomic Sciences, Mount Sinai School of Medicine, New York, NY; 2) Department of Allergy/Immunology, Winthrop University Hospital, New York, NY; 3) Pediatric Endocrinology and Diabetes, Mount Sinai School of Medicine, New York, NY.

DiGeorge syndrome is commonly defined by the presence of congenital heart defect, absent or hypoplastic thymus with immune deficiency and hypocalcemia due to hypoplastic or absent parathyroid glands. Additional features include craniofacial abnormalities such as velopharyngeal incompetence, facial dysmorphism, skeletal abnormalities, growth retardation, learning difficulty, and renal anomalies. The etiology of DiGeorge syndrome is heterogeneous, with 90 percent of cases due to microdeletion of chromosome 22q11.2 and rarely of chromosome 10p13. Non-genetic factors causing features of DiGeorge syndrome, referred to as DiGeorge sequence or anomaly, include fetal exposure to teratogens such as alcohol, Vitamin A, or maternal diabetes. We report a female infant delivered to a 26 year old mother with uncontrolled insulin dependent diabetes, who was non compliant with prenatal care. The infant was large for gestational age, with birth weight of 4.08kg and typical appearance of an infant of diabetic mother. She developed persistent hypocalcemia and was flagged by positive newborn screen for SCID. Evaluation revealed absent thymus. Lymphocyte panel revealed no T-lymphocytes but normal numbers of NK cells and elevated B cells. Cardiology evaluation showed no structural heart defect, but biventricular hypertrophy and PDA, both of which resolved at 2 weeks age. In addition, the infant had rib and vertebral anomalies. Chromosome analysis and FISH to detect a chromosome 22q11 deletion was unsuccessful due to absence of T cell lymphocytes. Array CGH was performed with a normal result. In the absence of 22q11 and 10p13 deletion, uncontrolled maternal diabetes was the most likely cause of these abnormalities. The increased risk of congenital malformations in infants of diabetic mothers is well documented. In particular, heart defects such as truncus arteriosus, caudal dysplasia, renal and musculoskeletal findings are well described with several reports of DiGeorge syndrome spectrum of branchial arch abnormalities. At least 9 individuals with DiGeorge anomaly including thymic aplasia are reported in association with maternal diabetes. Particularly noted in these cases is the associated presence of cardiac and renal defects. The mechanism by which maternal hyperglycemia causes birth defects is ill understood but is likely to be complex. Early and rigorous control of hyperglycemia is an important public health measure to prevent the occurrence of serious birth defects.

3034T

Investigation of telomere length in new newborns and correlation with gestation age and birthweight. *E.C. Tan[1], S.N. Lim[2], Z. Yahya[2], D. Zeegers[2], T. Moe[3], E.E.P. Kyaw[3], G.S.H. Yeo[3], M.P. Hande[2].* 1) KK Research Ctr, KK Women's & Children's Hospital, Singapore; 2) Department of Physiology, National University of Singapore; 3) Maternal-fetal Medicine, KK Women's & Children's Hospital, Singapore.

Telomeres are critical for genomic stability; they provide a means for chromosome replication and become shorter with each cell division. Telomere length appears to be a predictor of organismal fitness in the human population. Although there is some inter-individual variation in telomere length, shorter telomeres are associated with increased risk of some cancers and cardiovascular disorders. As there are only a few small studies on telomere lengths in newborns, we investigated the variability for newborns of different birthweight and gestation age in our hospital. Cord blood from births with Chinese parents was collected at delivery. Data was also collected on birthweight, gestation age, and mother's age. Terminal restriction fragment (TRF) assay was performed on genomic DNA extracted from whole blood to measure the telomere length of each sample. Variance analyses and correlation coefficients were calculated using SPSS. All study variables have normal distribution. The birthweight of the newborns in the study ranged from 2.2 kg to 4.2 kg with gestation age from 245 to 288 days. There was statistically significant correlation between mother's age with gestation age, and between birthweight and gestation age. Telomere length from 134 samples (68 males, 66 females) showed a wide range from 8.2 kb to 19.2 kb. The mean TRF length was 12.6 kb (males: 12.3 kb ± 2.6; females: 13.1 kb ± 2.4). Overall, there was marginal correlation of telomere length with gestation age with increased significance after controlling for maternal age. For males, the correlation was highly significant (r: -0.349; p: 0.004). This correlation of telomere length with gestation age is likely due to the result of telomere attrition in utero with increased growth.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3035W

A Case of Arthrogryposis and Mosaic Turner Syndrome. *L. Mora, A. Lopez, I. Zarante.* Instituto de Genética Humana, PontificiaUniversidad Javeriana, Bogotá, Colombia.

Arthrogryposis Multiplex Congenital is a heterogeneous condition defined as non-progressive contractures in more than two joints and in multiple body areas, the pathogenesis is impaired fetal movements. Arhrogryposis with central nervous system involvement includes chromosomal and other syndromes. The cause can be pathology in the peripheral or central nervous system, in muscles or in connective tissue, defects in neuromuscular transmission, compromised space in utero, maternal disease, external factors like medication or drugs, or compromised vascular supply to the fetus. Turner syndrome is a disorder of females, is a chromosomal aberration associated with partial or complete absence of one X chromosome. It is characterized by short stature, primary amenorrhea and some phenotypical features such as webbed neck, short stature, broad chest with widely spaced nipples, cubitus valgus, and lymphedema at birth. The most frequent chromosome constitution in Turner syndrome is 45,X without second sex chromosome. However, about 50 percent of cases have other karyotypes. About one quarter of Turner syndrome cases involve mosaic karyotypes, in which only a proportion of cells are 45,X. The most common karyotypes and their approximate relative prevalence are as follows; 50 percent of cases 45,X, 15 percent of cases 46,X,i(Xq), 15 percent of cases 45,X/46,XX, about 5 percent of cases 45,X/46,X i(Xq), about 5 percent of cases 45,X, other X abnormality and about 5 percent of cases other 45,X/? Mosaics. Contractures in multiple areas are quite rare in Turner Syndrome. Here is a case of newborn girl with Arthrogryposis multiplex congenital, and mosaic with a X chromosome monosomy 45,X/46,XX.

3036F

Paternal uniparental disomy 6 and 2q13 deletion detected prenatally in a patient with cardiac defects, transient neonatal diabetes mellitus, hydrocephalus and dysmorphic features. *D. Niyazov[1], C. Tillis[2], D. Feliperamirez[3], A. Robichaux[4].* 1) Dept Pediatrics, Div Med Genet, Ochsner Clinic Foundation, New Orleans, LA; 2) Dept Pediatrics, Div Neonatol, Ochsner Clinic Foundation, New Orleans, LA; 3) Dept Pediatrics, Div Endocrin, Ochsner Clinic Foundation, New Orleans, LA; 4) Dept Obst & Gynecol, Div Mat Fet Med, Ochsner Clinic Foundation, New Orleans, LA.

Introduction: Paternal uniparental disomy 6 (pUPD6) is associated with transient neonatal diabetes mellitus (TNDM), macroglossia and dysmorphic features. To our knowledge, pUPD6 was not reported with 2q13 deletion. We present a patient with multiple anomalies and prenatally ascertained pUPD6 and 2q13 deletion. We discuss mechanisms elucidating effects of pUPD6 and 2q13 deletion that resulted in patient's phenotype. Case Report: Fetal hydrocephalus and IUGR were detected at 18 weeks gestation. Karyotype showed 46,XX but SNP array revealed homozygosity of chromosome 6 which was paternal. SNP array also detected a 469 kb 2q13 deletion which was paternally inherited. NPHP1 sequencing showed a heterozygous deletion. Proband was born SGA at term and developed TNDM which resolved at 5 weeks. She had macroglossia, shallow orbits, infraorbital creases, low-set ears, micrognathia and long fingers. She also had ASD/VSD and hypotonia. Her brain MRI revealed hydrocephalus consistent with aqueductal stenosis requiring shunting. Discussion: Paternal UPD6 is associated with TNDM due to abnormal imprinting of 6q24 locus which may be due to overexpression of PLAGL1. The latter regulates pituitary adenylate cyclase-activating polypeptide receptor which stimulates insulin secretion. PLAGL1 is also an important member of a cellular network of imprinted genes involved in fetal growth. ZFP57 is another gene in 6q24 region which may play a role in structural brain abnormalities, developmental delay and CHD. It's unclear if the paternally inherited 2q13 deletion contributed to the phenotype. NPHP1 sequencing and MLPA did not detect another mutation along with the heterozygous deletion which could cause Joubert syndrome type 4. Positional and downstream effects of the 2q13 deletion could potentially modify expression of paternal UPD6 but there's no literature to support this. UPD6 was not described with 2q13 deletion. Our case underscores importance of chromosomal microarray in prenatal diagnosis. The SNP array in our patient appeared to have an advantage over oligonucleotide array because the latter cannot detect UPD. Prenatal identification of paternal UPD6 was crucial in diagnosis and management of TNDM and in estimation of prognosis and recurrence risks.

3037W

Congenital acquired mosaicism for monosomy 7. *J. Hiemenga, S. Klemm, J. Foley, H. Toriello.* Medical Genetics, Cancer Genetics, Spectrum Health, Grand Rapids, MI.

The molecular genetic etiology for Acquired Monosomy 7 and inevitable progression to myelodysplastic syndrome (MDS) or acute myelogenous leukemia (AML) is elusive, in part because of the rarity of affected individuals and families. We present the only reported congenital case of acquired Mosaic Monosomy 7. An infant male with a pregnancy complicated by polyhydramnios and antenatally diagnosed multiple congenital anomalies had a normal 46, XY prenatal karyotype. At birth he was found to be mosaic for monosomy 7 in 15% of peripheral leukocytes on a 180K oligonucleotide array, additionally he had an 11.6 Mb microduplication of the 9q33.3q-q34.3 region and a 5.18 Mb microdeletion of the 11q24.1-q24.3 region. He remained negative for mosaic monosomy 7 by FISH analysis on an alternate tissue source, a newborn buccal swab. This infant's 11q24.1 5.18 Mb deletion overlaps partially with that of Jacobsen syndrome, a contiguous gene deletion syndrome involving chromosome 11 distal to 11q23. This infant's clinical presentation was consistent with both Mosaic Monosomy 7 and Jacobsen syndrome. He had pancytopenia, was transfusion dependent, and supported with granulocyte-colony stimulating factor. He had multiple physical findings consistent with Jacobsen syndrome. Additionally, this infant had duodenal atresia, micro retrognathia, long thin neck, lumbar hemivertebrae with gibbus malformation, and a unilateral dysplastic kidney. Reports of acquired Mosaic Monosomy 7 in the literature are scant with a familial confirmation of the diagnoses made during bone marrow transplant screening of potential donor siblings. Affected parents have not been reported. The earliest reported onset is 9 months of age but acquired Mosaic Monosomy 7 presents most commonly throughout childhood into early adulthood. One of this infant's four elder siblings was found to have 3% monosomy 7 in peripheral leukocytes, below the reporting laboratory's control confidence interval range. She did not have thrombocytopenia; her initial bone marrow evaluation was normal and had no cells reported with monosomy 7. We have recommended close monitoring of this sibling and annual follow up with other family members.

3038F

Overgrowth as a presenting symptom of the 22q11.2 duplication syndrome - a novel association. *D.M. McDonald-McGinn[1], E. Bratton[1], K. Dickinson[1], A. Kohut[1], A. Bailey[1], A. Wilkens[1], I. Krantz[1], B. Emanuel[1], S. Saitta[2], M. Deardorff[1], E.H. Zackai[1].* 1) Div Human Gen, Children's Hosp Philadelphia, Philadelphia, PA; 2) Cedars Sinai Medical Center, University of California Los Angeles, Los Angeles, CA.

Introduction: 22q11.2 CNV's are well described entities involving deletions and duplications. Both are the result of non-allelic homologous recombination due to the presence of low copy repeats which flank the deletion/duplication and define the breakpoints. Although expected to occur with equal frequency we've found the deletion to be twice as prevalent as the duplication. Moreover, in several instances we and others have identified the duplication in seemingly unaffected adults, raising important questions regarding management, recurrence risk counseling, and prenatal diagnostic issues. Now, having previously evaluated 52 individuals with a 22q11.2 duplication, we report the novel association of overgrowth as a prominent feature - this following the identification of duplications in two children during routine evaluations for LGA. This new association may have important implications for ascertainment and further management. Methods: We retrospectively reviewed our database of 52 individuals with the 22q11.2 duplication syndrome following the identification of 22q11.2 duplications in two patients who were referred for evaluation to our overgrowth clinic. Results: 54 individuals (38 males; 16 females) with 22q11.2 duplications included 33 probands, 15 parents, 4 siblings, 1 aunt and 1 first cousin. Of these, growth data is reported on 34 children (27 males; 7 females) of whom 14/34 (41%) had a birth weight >90%. By gender, 11/27 (41%) males were >90% and 3/7 (43%) females were >90%. Birth length was >90% in 7/34 (21%) including 6/27 males (22%) and 1/7 (14%) females. 8/34 (24%) continued above the 90% beyond infancy including 6/27 (22%) males and 2/7 (29%) females. Head circumference mirrored height with 8/34 (24%) patients having macrocephaly including 7/27 (26%) males and 1/7 (14% females). Conclusion: Chromosome 22q11.2 is a hot spot for copy number variation with both deletions and duplications having similar structural and functional differences reported such as congenital heart disease, palatal differences, intellectual deficits and autism. However, recently published growth curves for the 22q11.2 deletion do not demonstrate overgrowth. Thus, here we report a novel finding in association with the 22q11.2 duplication. This may well be an important clue to the identification of the diagnosis. Moreover, it will be useful in providing appropriate care for these patients and perhaps in identifying candidate genes for overgrowth within this region.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3039W

Down syndrome patients nutritional evaluation at the Hospital Para El Nino Poblano outpatients. *M. Cortes[1], J. M. Aparicio[2,3], R. J. M. Perez[1], F. E. P. Romero[1].* 1) Nutrtion; 2) Genetics, Hospital para el Nino Poblano, Puebla, Puebla; 3) Estomatology, Benemerita Universidad de Puebla, Mexico.

From a total of 4617 karyotypes (100%) performed in 19 years period of time at a Pediatric Hospita, at Puebla city in Mexico; 33.6% (1553 patients) were diagnosed as chromosomal trisomy, were 32.8% (1511 patients) diagnosed as Down syndrome (trisomy 21). This alteration is usually caused by an extra copy of chromosome 21 (trisomy 21). Characteristics include decreased muscle tone, stockier build, asymmetrical skull, slanting eyes and mild to moderate developmental disability. 1127 patients have regular trisomy (47XX+21 or 47XY+21), caused by a meiotic nodisjunction event, also chromosomal translocations were observed; 260 patients 46XYt(14;21) or 46XXt(14;21q) were the long arm of chromosome 21 is attached to chromosome 14. 43 patients 46XYt(21;21) or 46XXt(21;21), and 81 patients with mosaicism. It was tried to determine nutritional status for growth, development and nutrition improvement in some Down patients . Looking at history, John Langdon Haydon Down identified and first described in 1866 the syndrome that now his name, it was not until 1932 when Davenport suggested that chromosomal irregularities could cause certain forms of mental retardation, including Down syndrome. Only in 1956 the available techniques allowed establish a definitive finding that the normal human chromosome number is 46. In 1959 Jerome Lejeune discovered that Down syndrome was extra chromosome belonging pair chromosomes 21 (HSA21). His finding was confirmed by Jacobs that year. After first cases of translocation and mosaicism. In 1970 Caspersson postulate genetic material that, when it tripled, causes appearance phenotypic characteristic of Down syndrome, is concentrated in distal long arm chromosome 21. Down Syndrome is principal cause of intellectual disability and the most common human genetic disorder: 1/600 to 1/1000 views. Approximately 40% children with Down syndrome born with congenital heart disease and 15% born with gastrointestinal malformations. Children with Down syndrome also at risk of other medical complications can affect nutritional status and their overall development. This study main goal carries assessment nutritional status in children with Down syndrome by determining nutritional status which may impact on growth, development and nutrition in children with this genetic alteration. Both rehabilitation therapy as nutritional evaluation, are both important for a better quality of life.

3040F

Follow-up of patient with non-supernumerary ring chromosome 7: Clinical manifestations, cytogenetic and molecular analysis. *C. Salas[1], P. Perez-Vera[1], D. Cervantes[2], R. Cruz-Alcivar[1], R. Daber[4], L. Conlin[4], L. Leonard[4], NB. Spinner[4], C. Duran-Mckinster[3], V. Del Castillo-Ruiz[2].* 1) Laboratorio de Cultivo de Tejidos, Instituto Nacional de Pediatria, Mexico city, Mexico; 2) Departamento de Genetica Humana, Instituto Nacional de Pediatria, Mexico city, Mexico; 3) Servicio de Dermatologia, Instituto Nacional de Pediatria, Mexico city, Mexico; 4) Department of Pathology and Laboratory Medicine of Children's Hospital of Philadelphia, PA.

Ring chromosome 7 (r7) is an uncommon cytogenetic finding that leads to variable phenotypes, resulting from variable terminal deletions, dynamic mosaicism and parental origin. The most frequent phenotype of the r7 patients are pre- and postnatal growth retardation, microcephaly, intellectual disability and dermatologic findings. Almost all reported cases present dynamic mosaicism, with the same apparent chromosomal breakpoints and ring structure, and in only two patients FISH analysis revealed cryptic subtelomeric or telomeric deletions. Here we report the clinical, cytogenetic and fine molecular mapping of a male child with a non-supernumerary r7. Clinical report. The propositus was the first child of healthy non-consanguineous parents. At 20 months of age, the patient had dysmorphic features, developmental delay and dermatologic lesions with variable pigmentation. We re-evaluated the patient at 6 years and physical examination showed hypogenitalism, lumbar dextroscoliosis, cerebellar and ophthalmological abnormalities and melanocytic congenital nevi (MCN). Methodology. Cytogenetic and molecular analysis was performed in peripheral blood (PB), fibroblasts of light (LS) and dark (DS) skin and in a MCN sample. Results. The cytogenetic analysis in PB showed a single cell line with r7; LS and DS revealed monosomy (-7) and r7 cell lines in different proportions. Almost 5 years later, a new cytogenetic analysis showed three different cell lines in PB: monosomy 7, r7 and duplicated r7 (idic r(7)); LS sample was similar to the previous study, with monosomy and r7 cell lines, and in MCN, monosomy, r7 and duplicated r7 cell lines were observed. FISH revealed subtelomeric loss in both chromosome arms in all tissues studied. Genome-wide SNP array showed a 0.8 Mb deletion in 7p22.3 (8 genes) and a 7.5 Mb deletion in 7q36 (29 genes). SNP array analysis also revealed low level monosomy 7 mosaicism in PB (less than 5%). Analysis of informative SNPs on chromosome 7 demonstrates the presence of maternal uniparental disomy (UPD) in the monosomy 7 cell line. This finding suggests that the r7 is paternally derived. Discussion. The SNP array analysis was useful to delineate the deleted r7 regions, which included genes related to genital and central nervous system development and for identification of the maternal UPD involving Silver-Russell region. The combination of all results obtained permits us to establish an accurate genotype-phenotype correlation.

3041W

MOWAT - WILSON SYNDROME: CASE REPORT. *J. Acosta Guio[1,2,3], A. Zarante[1].* 1) medical genetic, Pontificia Universidad Javeriana Instituto genética humana, Bogotá, Colombia; 2) Instituto de nutricion, genetica y metabolismo de la Universidad el Bosque, Bogotá, Colombia; 3) Instituto de Ortopedia Infantil Roosevelt, Bogotá, Colombia.

Mowat - Wilson syndrome (MWS), is a rare disease the prevalence is currently unknown approximately 180 cases have been reported until 2010. It is characterized by distinctive facial phenotype, moderate to severe mental retardation, epilepsy Hirschsprung disease and other congenital malformations. Case Report: We describe the case of 9 years old Colombian girl, second pregnancy, non-consanguineous marriage, born at term uncomplicated. Mother reports that the patient had hypotonia from birth. Later the patient present developmental delay, and epilepsy. Currently, low weight, microcephaly, triangular facies, bilateral ptosis, posteriorly rotated ears, short philtrum and full lips, hypothenar hypoplasia hands, cubitus valgus and joint hypermobility. The patient did not present Hirschsprung disease but she has a diagnosis of precocious puberty. The initial investigation included karyotype: the result was 46,XX and the comparative genomic hybridization (CGH) showed a microdeletion at chromosome 2q22.32-q22.3, for this reason the Fluorescent in situ hybridization (FISH) was made in her parents who had not the microdeletion. Finally, we concluded that the patient had a novo deletion. MWS is caused by mutations or deletion (≈ 17%) in that region is located the ZEB2 gene (zinc finger E-box-binding homeobox 2), who is involved in the development of neural crest derived cells, central nervous system and midline development. The majority of mutation are sporadic, therefore the risk of recurrence in the siblings is low or the same of the rest of population. The precocious intervention is necessary and it should include: physical therapy, occupational and speech therapy. We present the first case molecularly confirmed in Colombia.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3042F

The association of low socioeconomic status and the risk of having children with Down syndrome: a report from the National Down Syndrome Project. *J.E. Hunter, E.G. Allen, L.J.H. Bean, S.B. Freeman, S.L. Sherman*. Department of Human Genetics, Emory University, Atlanta, Georgia.

The majority of Down syndrome (DS) cases are due to the nondisjunction of chromosome 21 in oocytes. To date, advanced maternal age and altered recombination are the only well-established risk factors for the nondisjunction errors and these are known to differ in profile with respect to the type of error, meiosis I (MI) or meiosis II (MII). In this study, we examined the association between low maternal socioeconomic status (SES) and trisomy 21 stratified by the type of nondisjunction error among 1691 families recruited as part of the National Down Syndrome Project, a multisite, population-based case-control study of live births. Three potential SES risk factors (both mother and father had not completed high school, both maternal grandparents had not completed high school, and an annual household income of <$25,000) were used to calculate an SES index (0, 1, and ≥2 low SES factors present). We used logistic regression and adjusted for statistically significant confounders including maternal age, race, and parity. We detected an association of having a child with maternal MII error among families with one low SES factor (OR=1.81, 95% CI=1.07–3.05) and two or more low SES factors (OR=2.17, 95% CI=1.02–4.63) compared to families with no risk factors. We found that this association was driven primarily by having a low household income (OR=1.79, 95% CI=1.14–2.73). No association was found with maternal MI errors for 1 and ≥2 low SES factors (OR=1.11, 95% CI=0.80–1.56 and OR=1.31, 95% CI=0.81–2.10, respectively). Further studies are warranted to explore which aspects of low maternal SES, such as environmental exposures, may account for this association and why they influence this one type of nondisjunction error of chromosome 21.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3043W

Surgical Intervention for Esophageal Atresia in Patients with Trisomy 18. *E. Nishi[1,2], T. Nakamura[3], K. Iio[4], S. Mizuno[5], H. Kawame[6], Y. Fukushima[7], T. kosho[7].* 1) Dept Genetics, Nagano Children's Hosp, Azumino City, Nagano, Japan; 2) Department of Medical Genetics, Shinshu University Graduate School of Medicine, Matsumoto, Japan; 3) Department of Neonatology, Nagano Children's Hospital, Azumino, Japan; 4) Department of Surgery, Central Hospital, Aichi Human Service Center, Kasugai, Japan; 5) Department of Pediatrics, Central Hospital, Aichi Human Service Center, Kasugai, Japan; 6) Department of Genetic Counseling, Ochanomizu University School of Humanities and Sciences, Tokyo, Japan; 7) Department of Medical Genetics, Shinshu University School of Medicine, Matsumoto, Japan.

Trisomy 18 (T18) is the second most common autosomal trisomy characterized by multiple congenital anomalies and an extremly short life span, with the median survival time as 14.5 days and the survival rate at age 1 year from 5 to 10%. Esophageal atresia (EA) is a frequently observed and potentially letal complication in patients with T18. There has been no evidence about efficacy of surgical intervention for EA in patients with T18. To delineate the outcome of surgical intervention for EA with tracheoesophageal fistula (TEF) in T18, we reviewed detailed clinical data of 24 patients with full T18 who had EA admitted to the neonatal intensive care units of Nagano Children's Hospital from April 1993 to March 2008 and Central Hospital of Aichi Human Service Center from April 1982 to March 2009. They all underwent one or two operations for their EA. Nine of 24 patients (Groups 1 and 2, 37%) had only palliative surgery. Group 1 (n=6) had only gastrostomy or both gastrostomy and jejunostomy. Group 2 (n=3) had gastrostomy and TEF resection (TEFR). The other 15 patients (Groups 3 and 4, 63%) had radical surgery. Group 3 (n=10) had single-stage operation of esophago-esophagostomy and TEFR. Group 4 (n=5) had gastrostomy followed by esophago-esophagostomy and TEFR. Surgical complications included hemorrhage, chylothorax, pneumothorax, mediastinitis, respiratory infection, and insufficient suture requiring reoperation. No intraoperative death or anesthetic complications were noted. Enteral feeding was accomplished in 17 patients (71%), three of whom were fed orally. The median survival time was 29.5 days (range,1–1786): 12 days in Group 1 (1–133): 106 days in Group 2 (47–172): 23 days in Group 3 (2–694): and 518 days in Group 4 (32–1786). The survival rates at age 1 day, 1 week, 1 month, and 1 year were 100%, 92%, 58%, and 17%, respectively. Three patients could be discharged to home (12.5%), with the median hospital stay as 137 days (73–947). The common underlying factors associated with death were congenital heart defects and heart failure, followed by pulmonary hypertension. These data suggest that surgical intervention for EA in patients with T18, especially two-stage operation (gastrostomy followed by esophago-esophagostomy and TEFR), could contribute to longer survival, establishment of enteral feeding, and that EA might not always be an absolute poor prognostic factor in patients with T18.

3044F

Mosaic supernumerary ring chromosome 3: Does copy number gain of FOXP1 contribute to expressive speech impairment and intellectual disability? *I. Filges[1,2], A. Datta[3], E. Boehringer[3], L. Suda[4], B. Roethlisberger[4], P. Miny[1].* 1) Division of Medical Genetics, University Children's Hospital and Department of Biomedicine, Basel, Switzerland; 2) Department of Medical Genetics, Children's and Women's Hospital, Child and Family Research Institute, Vancouver, Canada; 3) Division of Neuropediatrics and Developmental Medicine, University Children's Hospital, Basel, Switzerland; 4) Medical Genetics, Center of Laboratory Medicine, Cantonal Hospital, Aarau, Switzerland.

Assessing the clinical significance of mosaic small supernumerary ring chromosomes (sSRC) still represents a major challenge in prenatal and postnatal diagnosis, and phenotype-genotype correlations are notoriously difficult. We report on the extremely rare finding of a mosaic sSRC of chromosome 3, initially identified prenatally, the current postnatal findings by array genomic hybridization (AGH) and the long-term clinical follow-up of the child. The now 7 year-old girl shows minor facial anomalies, microcephaly and is missing one upper central incisor. She has moderate intellectual disability and severe speech impairment, with an almost complete absence of expressive speech, uses simple sign language, whereas receptive speech is conserved. Motor development is within the normal range. She has repeated absences and had one generalized seizure. Cerebral MRI and basic metabolic investigations were normal. Repeated EEG and polysomnography recordings showed increasing epileptogenic activity of a spike and spike-wave type in the posterior region to currently generalized spike-wave activity continuing during slow- sleep (CSWS) but conserved sleep structure. High resolution AGH of DNA from stored cultured amniocytes confirmed a de novo gain (copy number state 2.5) of the genomic region 3p14.2q11.2 encompassing 35.1 Mb harboring 65 genes. The identical aberration was found in DNA from cultured fibroblasts but not from blood lymphocytes of the patient. Conventional karyotyping of fibroblasts showed a SRC in 25%, 2 SRC's in 25% and a normal result in 50% of cells reflecting array results. Due to the size of the aberration and developmentally relevant gene content we conclude that the mosaic SRC 3 may explain our patient's phenotype although we could not identify patients with the same duplication of the affected region for comparison. The region contains FOXP1 (3p13) haploinsufficiency of which causes ID with severe language impairment. FOXP1 encodes a transcription factor and is closely related to FOXP2, a gene also implicated in rare forms of speech disorders. Functional evidence of heterodimer formation and overlapping neural expression patterns suggests that FOXP1 and FOXP2 can co-regulate gene expression in the brain. We hypothesize that also increased levels of FOXP1 transcription may affect these interactions and thus contribute to speech and neurodevelopmental impairment in our patient.

3045W

Cat eye syndrome: Wide clinical variability in three patients from the same family. *M.I. Melaragno[1], S.I. Belangero[1], A.N. Pacanaro[1], F.T. Belucco[1], D.M. Christofolini[2], L.D. Kulikowski[3], R.S. Guilherme[1], A. Bortolai[4], A.R. Dutra[1], F.B. Piazzon[5], M.C. Cernach[1].* 1) Morphology and Genetics Department, Universidade Federal de São Ppaulo, São Paulo, São Paulo, Brazil; 2) Ginecology and Obstetricy Division, Faculdade de Medicina do ABC, Brazil; 3) Department of Pathology, Cytogenomics Laboratory, LIM 03, Universidade de São Paulo, SP, Brazil; 4) Genetics Division. Hospital do Servidor Público do Estado de São Paulo, Brazil; 5) Genetics Unit. Instituto da Criança, Universidade de São Paulo, Brazil.

A small supernumerary marker chromosome (sSMC) derived from chromosome 22 is a relatively common cytogenetic finding. The 22 sSMC can result in tetrasomy for a chromosomal region that spans the chromosome 22p arm and the proximal 2 Mb of 22q11.2. The majority of cases are sporadic, although some familial cases have been reported. Here, we report three patients from the same family (mother, daughter, and son), with sSMCs derived from chromosome 22, resulting in cat eye syndrome. There are no previous reports of this syndrome occurring within three patients in the same family. Classical cytogenetics and FISH with probes from chromosome 22 revealed a supernumerary dicentric and bisatellited chromosome derived from chromosome 22 in all metaphase cells analyzed. MLPA using the SALSA MLPA P250 DiGeorge Syndrome kit and high resolution array showed four copies of the proximal region of chromosome 22, described as arr 22q11.1q11.21(14,435,171-17,190,083)×4 (NCBI 36/hg 18), indicating a breakpoint at LCR-A. Thus, all patients presented as 47,XY or XX,+idic(22)(pter→q11.21::q11.21→pter), resulting in cat eye syndrome due to chromosome 22 partial tetrasomy 22pter-q11.21. Although all subjects presented the same chromosomal abnormality, they showed a wide range of phenotypic differences, even being from the same family. While the daughter presents a more severe phenotype than her brother and her mother, the mother presents only subtle abnormalities and was diagnosed only because of her children. Our findings emphasize the great clinical variability in CES even in familiar cases with the same marker chromosome and also highlight the importance of performing chromosome examination even in asymptomatic parents to determine the risk of recurrence and provide more conclusive genetic counseling for these families. Financial support: FAPESP, Brazil.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

† Trainee Award Finalist

3046F

Maternally inherited Xq26.2 duplication in a male offspring with severe prenatal and postnatal growth deficiency, dysmorphic facial features, hydrocephalus and developmental delay resembling Russell-Silver syndrome. *L. Tuzovic, J. Wynn, L. Rohena, K. Anyane-Yeboa, A. Iglesias.* Division of Clinical Genetics, Columbia University Medical Center, New York, NY.

Xq duplications in males are extremely rare. Less than 100 cases with different breakpoints on Xq have been described in the literature. To our knowledge, duplications in the distal region of Xq26.2 have been described in only a few patients. Madrigal et al. reported a similar duplication as the one in our case in two brothers with prenatal and postnatal growth deficiency and developmental delays, hypotonia, craniofacial disproportion and dental malocclusion. We further expand the clinical and molecular phenotype in a 5 year old male with severe prenatal and postnatal growth deficiency, hydrocephalus, dysmorphic facial features and developmental delay resembling Russell-Silver syndrome. He was born to a phenotypically normal 19 year old Hispanic female G1P0 at 35 weeks gestation via emergent cesarean section due to fetal bradycardia. Prenatal course was complicated with severe intrauterine growth restriction and oligohydroamnios which prompted induction of labor. Prenatal ultrasound revealed bilateral ventriculomegaly. Fetal echocardiogram showed mild muscular VSD. Amniocentesis was performed and showed 46, XY male karyotype. His birth weight was 1365 grams (<5th percentile). Exam at 2 months of age was significant for dysmorphic facial features including relative macrocephaly, frontal bossing, small and flat face, epicanthal folds, hypertelorism, down slanted palpebral fissures, small cupped ears, down turned corners of mouth and micrognathia. His postnatal course was further characterized by severe growth deficiency and motor and cognitive delay. MRI of brain at 2 years of age showed stable mild hydrocephalus secondary to partial aqueduct stenosis and thinning of the corpus callosum. Genetic workup included SNP-microarray which revealed a 3.9 MB microduplication in Xq26.2q26.3 region. Parental studies revealed that phenotipically normal mother is a carrier of the same duplication. Maternal X-inactivation studies showed 100 percent skewed X-inactivation pattern suggesting that the inherited duplication is likely pathogenic. At the age of 5 years, he remains below the 3rd percentile for height and weight. He shows mild developmental delay and attends special class in kindergarten where he receives speech and physical therapies. He now speaks in full sentences however still has problems with fine motor skills. This report will help to further expand the clinical and molecular phenotype associated with rare distal Xq duplications.

3047W

***De novo* translocation disrupting Mediator complex subunit in a patient with Pierre-Robin sequence and developmental delay.** *K.H. Utami[1,4], A.M. Hillmer[2], E.G.Y. Chew[2], C.L. Winata[1], V. Korzh[3], S. Mathavan[1], P. Sarda[5], S. Davila[1], V. Cacheux[1,4].* 1) Human Genetics, Genome Institute of Singapore, Singapore, Singapore, Singapore; 2) Genome Technology and Biology, Genome Institute of Singapore, Singapore, Singapore; 3) Fish Developmental Biology, Institute of Molecular and Cell Biology, Singapore, Singapore; 4) Cytogenetics Department, Genome Institute of Singapore, Singapore, Singapore; 5) Department of Genetics, Montpellier University Hospital, Montpellier, France.

Pierre-Robin sequence (PRS) is a clinically well-defined subgroup of cleft lip and palate, micrognathia, and respiratory distress with unknown etiology. Here we report a case of a girl diagnosed with Pierre-Robin sequence at birth. She was diagnosed with developmental delay by the age of 4, and shown to have dysmorphic features, strabismus, hirsutism and multiple contractures. Karyotyping analysis revealed a *de novo* translocation between chromosomes 12q24 and 19q12. By using long paired-end tag (DNA-PET) sequencing we narrowed down the breakpoint to a disrupted Mediator complex subunit gene, *MED13L* on chromosome 12. The breakpoint coordinate was confirmed by FISH analysis and refined by sequencing. Furthermore, transcript and protein expression levels were found to be reduced by >50% in the patient's lymphoblastoid cells. We checked the ortholog expression in the zebrafish and mouse embryo, and observed an enriched expression in the brain and branchyal arches, respectively. This was then supported by similar patterns of expression in the hippocampus and dentate gyrus (BrainAtlas) of mouse and human. Other mutations in the same gene have been reported previously in a patient with congenital heart defects and intellectual disability. In conclusion, *MED13L* might have a potential function that could be the underlying cause of phenotype observed in the patient.

3048F

614kb duplication at chromosome 9q22.32 encompassing PTCH1 gene in a family with reciprocal translocation (9;20)(q22;q13) and intrafamilial phenotypic variability. *C. Vinkler[1], D. Lev[4], A. Singer[3], A. Frumkin[1], M. Davidovich[2], M. Michelson[1], Y. Michaeli-Yossef[2].* 1) Inst Med Gen, Wolfson Med Ctr, Holon, Israel; 2) Child Developmental Center, Maccabi Health Services, Rishon LeZion, Israel; 3) Medical Genetics Unit, Baezilai Med Ctr, Ashkelon, Israel; 4) Department of Genetic Research, The Department of Genetic and Metabolic Diseases, The Hebrew University Medical Center, Jerusalem, Israel.

Several previously unidentified, clinically significant, submicroscopic chromosome abnormalities, have been discovered since the widespread clinical use of comparative genomic hybridization microarray technology. Specifically, there have been reports of clinically significant microduplications found in regions of known microdeletion syndromes. These microduplications have distinct features from those described in the corresponding microdeletion syndromes. Deletions or loss-of-function mutations of the PTCH1 gene, result in basal cell nevus syndrome (Gorlin syndrome), while gain-of-function mutations were proposed to cause holoprosencephaly (HSE). We report on a familial translocation t(9;20)(q22;q13) in three cases with variable phenotypes. A 614kb duplication at chromosome 9q22.32 harboring the PTCH1 gene, was found in one of the cases by array CGH. This is a 2y old boy, who carries a t(9;20)(q22;q13) translocation. He has microcephaly, short stature, failure to thrive, dysmorphic features and developmental delay. His brain MRI is normal. He inherited the translocation from his father who has microcephaly, short stature, dysmorphic features and normal intelligence. Family history discloses a pregnancy that was terminated at 16w gestation due to holoprosencephaly. Fetal karyoytype revealed the paternal t(9;20)(q22;q13)translocation. We postulate that in this family, all three cases may have the same 614kb duplication near the site of the translocation breakpoint which was found in the 2y old boy. The duplicated region harbors two genes FANCC and PTCH1. No clinical significance is known to be associated with FANCC duplication. However, duplication of the region encompassing the PTCH1 gene has been previously described in two families with microcephaly, short stature, and intellectual disability. This is the first report of HSE in a family with PTCH1 duplication. Gain of function of the PTCH1 together with modifier genes involved in the sonic hedgehog signaling pathway, may be responsible for the variable expression in this family causing also an HSE phenotype.

3049W

Case report: A boy with 49, XXXXY Syndrome, diagnosed by karyotype and a characteristic phenotype. *A. Zarante, J.C Prieto, O. Moreno.* Insituto de Genetica, Universidad Pontificia Javeriana, Bogotá, Colombia.

49,XXXXY Syndrome, is a rare sex chromosome aneuploidy ocurr in 1:85,000 - 1:100,000 male births. That conditions are characterized by the growth deficiency with multiple congenital malformations: microcephaly, facial dysmorphism, heart, genitourinary and skeletal defects (limited pronation elbow, radioulnar synostosis, clynodactyly of fifth finger) also present developmental delays, cognitive impairments and behavioural disorders. We present case of a child of 17 months with developmental delay. The positive findings found on the physical examination were: hypertelorism, upward slant to palpebral fissures, inner epicanthal fold, auricular anomaly, micrognathia, radioulnar synostosis, clynodactyly of fifth finger, cryptorchidism and hypoplasic scrotum. The nuclear magnetic resonance brain showed : increased supratentorial ventricles and incipient signs of volume loss with enlarged subarachnoid spaces. The band R karyotype reported 49,XXXXY. We made a a literature review and discussed the case with additional phenotypic manifestations mainly in the brain.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3050F

6p25 Interstitial Deletion in Two Dyzigotic Twins with Gyral Pattern Anomaly and Speech and Language Disorder. *M. Bozza[1], L. Bernardini[2], A. Novelli[2], P. Brovedani[1], E. Moretti[1], R. Canapicchi[1], V. Doccini[1], T. Filippi[1], A. Battaglia[1].* 1) Stella Maris Clinical Research Institute, Calambrone (Pisa), Italy; 2) Mendel Laboratory, IRCCS Casa Sollievo della Sofferenza-Hospital, San Giovanni Rotondo (FG), Italy.

Submicroscopic 6p25 deletion is now recognized as a clinically identifiable syndrome, characterized by intellectual disability, language impairment, hearing deficit, craniofacial, ophthalmologic, cardiac, and varying central nervous system anomalies. We report on two dizygotic twins with a maternal segregating hemizygous interstitial deletion on chromosome 6p25.1, spanning 0.9 Kb; the smallest ever reported. Both had dysmorphic features (prominence of the metopic suture, synophrys, hypertelorism, down-slanting palpebral fissures, tented mouth), and a distinct brain MRI, showing a focal significant increase of the right peri-frontal subarachnoid space, with shallow sulci and a mild anomaly of the gyral pattern. Such brain anomaly has never been reported in association with del 6p25. Both propositi had a borderline-mild intellectual disability, speech and language difficulties, and behavior abnormalities. Their mother, formally tested, had a borderline cognitive impairment. Although none of the genes mapping to the deleted region are apparently related to the phenotype, LYRM4 resulted down-regulated in the cerebellar cortex of schizophrenia patients compared with controls, and Lyrm4 was down-regulated in the prefrontal cortex of mice with microdeletions in the locus syntenic to human 22q11.2 patients affected by schizophrenia. These data are in agreement with the emerging concept that similar CNVs are pathogenic in patients affected by distinct neurological diseases, and that these loci are more general risk factors for different disorders. The resemblance of our patients to those with the more extensive 6p25.1p25.3 terminal deletion suggests that the gene/s responsible for the physical phenotype should reside in the 6p25.1 genomic region.

3051W

Microdeletion of 19p13.3 in a girl with Peutz-Jeghers syndrome, intellectual disability, hypotonia, and dysmorphic features. *Y. Kuroda[1], T. Saito[2], J. Nagai[2], K. Ida[1], T. Naruto[1], M. Masuno[3], K. Kurosawa[1].* 1) Devision of Medical Genetics, Kanagawa Children's Medical Center, Yokohama, Kanagawa, Japan; 2) Department of Clinical Laboratory, Kanagawa Children's Medical Center, Yokohama, Kanagawa, Japan; 3) Genetic Counseling Program, Kawasaki University of Medical Welfare, Kurashiki, Japan.

The Peutz-Jeghers syndrome (PJS) is a rare autosomal-dominant disease characterized by gastrointestinal polyposis and mucocutaneous pigmentation. Germline point mutations in STK11 have been identified in about 70% of patients with PJS. Only a small number of large genomic deletions have been identified. We report a girl with PJS and multiple congenital anomalies. The girl was the second child of non-consanguineous and healthy parents. She was delivered at 35 weeks 2 days. Her birth weight was 1821g, length 42cm, and head circumference 27cm. She had umbilical hernia, bilateral inguinal hernia, scoliosis, strabismus, and malformed ears. She had conductive deafness secondary to middle ear disease. Hypotonia was noted at the first year of age. She started controlling head at 6 months, sitting at 10 months, standing 33 months, and walking at age 43 months. She started speaking several words at age 3 and went to junior school with special class. She developed lip pigments at age 12. Oligonucleotide-based array CGH revealed an approximately 610kb deletion at 19p13.3 including STK11 gene and other 21 RefSeq genes. Most of the PJS patients with partial or whole STK11 deletion were revealed smaller than 200kb deletion at 19p13.3. These cases had a typical PJS and no other phenotype. Le Meur [2004] reported a patient with PJS and intellectual disability associated with 250kb deletion of 19p13.3. The patient had a deleted allele of PJS derived from his mother and also had paternally derived NF-1 mutation allele. Our patient is the second case of smaller than 1Mb deletion with PJS, intellectual disabilities, and other anomalies. Several studies have determined the prevalence of congenital anomalies in children with cancer. Cancer development would share the etiology of congenital malformations. These results indicated that the application of array CGH for MCA/DD is a clue to understand cancer risk and surveillance.

3052F

Maternal mosaicism for deletion of 22q13.3 resulting in siblings with Phelan-McDermid syndrome. *K. Phelan[1], B.R. DuPont[2], R.C. Rogers[2].* 1) Hayward Genetics Ctr and Dept of Pediatrics, Tulane Univ Health Sci Ctr, New Orleans, LA; 2) Greenwood Genetic Center, Greenwood, SC.

We report a phenotypically normal mother with a deleted, dicentric chromosome 22 resulting from a translocation between 22q13 and an unidentified acrocentric chromosome. Two of her three children inherited the deleted 22 and have Phelan-McDermid syndrome. EH was born in 1993 and presented with neonatal hypotonia and failure to thrive. Physical features included ptosis, long eyelashes, full cheeks, flat midface, small jaw, narrow philtrum, and slight syndactyly between toes two and three. Chromosome studies at age 2 were reported as normal. She spoke a few words at age 1 year then lost her speech. Her receptive language skills were better than her expressive skills. She was diagnosed with autism at age 5. Her brother TH was born in 1996 and was diagnosed with autism at age 2. Physical features included dolicocephaly, bitemporal narrowing, flat midface, pointed chin, high arched palate, long eyelashes, and lax ligaments with hyperextensible joints. His hands were slightly large with mild 5th finger clinodactyly. Chromosome studies at age 4 detected a dicentric chromosome designated dic(22;?)(22pter->22q13.3::?cen->?pter). Chromosomes were repeated on EH and identified the same dicentric 22 as seen in her brother. Parental chromosome analyses was performed and revealed that the dicentric was present in 6% of maternal (CH) peripheral blood cells. In 2006, chromosome analysis, FISH, and microarray studies were performed on CH and TH. Custom OGT 22q microarray on TH determined that the deletion size was 3.5 Mb and extended to the terminal region of chromosome 22. Chromosome structure and mosaicism was further characterized in CH with G-banding and FISH. The low level of mosaicism in CH was not detected by microarray analysis. This case demonstrates the challenge of diagnosing parental mosaicism for structural chromosome abnormalities. Typically, a parental rearrangement can be identified by screening a minimal number of metaphase spreads by G-bands or FISH. In individuals with mosaicism, screening a low number of spreads may fail to detect the abnormal cells so the study must be expanded to screen a greater number of cells. The use of microarray analysis is this case was not helpful as it failed to detect the mosaicism in the mother. If not for the birth of a second affected child in this case the study for mosaicism may not have been pursued and genetic counseling based on minimal screening would have provided inaccurate recurrence risks.

3053W

De Novo Interstitial Deletion of 3q22.1-q22.3 in Two Patients- a New Microdeletion Syndrome. *Y. Wilnai[1], W. Wilson[3], C. Jett[3], M. Manning[1,2].* 1) Division of Medical Genetics, Department of Pediatrics, Stanford University Medical Center, Stanford, CA; 2) Department of Pathology, Stanford University Medical Center, Stanford, CA; 3) Division of Pediatric Genetics, Department of Pediatrics, University of Virginia, Charlotteville, VA.

With improved cytogenetic techniques, sub microscopic chromosome deletions and duplications are being identified with increased frequency. We report two unrelated patients who exhibit a distinct behavioral pattern and dysmorphic features that were found to have deletion of chromosome 3 at band q22.1-q22.3. Patient 1-microarray analysis was performed using 135k whole genome oligonucleotide array, UCSC hg 18. Patient 2-microarray analysis was performed using 180k whole genome oligonucleotide array, UCSC hg 19. Patient 1 is a 13-yo male with a 4.5 Mb deletion of 3q22.1q22.3 (43 genes, 2 of which code for OMIM disorders). One gene, SLCO2A1, is associated with hypertrophic osteoarthropathy; the second, PCCB, is associated with propionic acidemia. Patient 2 is a 9-yo male with 4.1 Mb deletion of 3q22.1q22.3 (39 deleted genes, four of which code for OMIM disorders). Two deleted genes are mentioned above; the other two are the TF gene associated with atransferrinemia and BFSP2 associated with cataracts. Both boys' development followed a parallel behavioral pattern. They were diagnosed with autism in toddler years and received intensive behavioral therapy, with progress in speech and socialization. Later they developed attention deficit hyperactive disorder. Both boys had normal birth parameters, however, later developed microcephaly and failure to thrive. Patient 1 developed central obesity during his second decade. The two boys suffered from seizure like episodes with normal EEGs and brain MRI. Both patients have a recognizable pattern of facial features including ptosis, beaked nose, low set posteriorly rotated ears, smooth philtrum and thin upper lip. We report 2 patients with deletion of chromosome 3q22.1-q22.3. The patients presented with a distinct and recognizable behavioral pattern and physical features. After a thorough literature review, it appears that this is the first report describing this microdeletion syndrome. About 40 genes are deleted in this region, less than 5 are associated with recognized disorders. Deletion of the disease associated genes does not seem to explain our patients' features. By delineating the clinical and behavioral features of this new micro deletion syndrome, we can provide information that may lead to a better understanding of the functions of the deleted genes.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3054F

Chromosomal microarray analysis of individuals with autism or learning deficits presenting for genetic services. *J. Roberts[1], J. Hayes[1], N. Dzidic[2], K. Hovanes[2], M. Dasouki[1], M.G. Butler[1].* 1) Department of Psychiatry and Behavioral Sciences, Kansas University Medical Center, Kansas City, KS; 2) Combimatrix Diagnostic, Irvine, CA.

Chromosomal microarray analysis is now in common practice in the clinical setting for identifying deletions and duplications in the human genome. We report our experience with the use of 105K or 180K oligonucleotide microarrays (performed at CMDX laboratory, Irvine, CA) in consecutive patients with autism or autism spectrum disorders (ASD) and/or developmental delay or learning problems presenting for genetic services at Kansas University Medical Center during the past 4 years. Of the 214 patients [139 males and 75 females (male/female ratio=1.85) with average age of 10.3 years and range of 9 months to 52 years; 64 (or 30%) with autism or ASD and 150 (or 70%) with learning deficits], 44 (or 21%) had either a known anomaly detected on the microarray in 30 patients or a non-diagnostic copy number variant (CNV) in 14 patients contributing to the clinical presentation. Six patients with ASD showed a recognized anomaly by microarray analysis while 6 had non-diagnostic CNV findings. The most common chromosome involved in the microarray anomaly in the ASD subgroup was chromosome 15 (3 of 12 microarray findings). For the learning deficit subgroup, chromosomes 2 (5 of 32 findings) and 22 (5 of 32 findings) were most involved. For those with learning deficits, 24 patients showed a recognized microarray anomaly of which 18 had a deletion while 8 patients had a non-diagnostic CNV which was usually a duplication. Twelve out of 64 patients (19%) with ASD had a microarray finding compared with 32 out of 150 patients (21%) with learning deficits. Clinically, no patient with ASD and a microarray finding had seizures compared with 4 patients (12%) with learning deficits. Four patients (33%) with ASD and a microarray finding had dysmorphic features usually craniofacial compared with 18 patients (56%) with learning deficits. No patient with ASD had microcephaly compared with 8 patients (25%) with learning deficits. Three patients (25%) with ASD had macrocephaly compared with 1 patient (3%) with learning deficits. Three patients (25%) with ASD and a microarray finding had a family history of ASD compared with 9 patients (28%) with learning deficits and a family history of learning deficits. In conclusion, our study would further support a significant yield in identifying genetic factors contributing to ASD and learning deficits in patients presenting for genetic services and the usefulness of microarray analysis in the clinical setting.

3055W

A Wolf Hirschhorn - like phenotype in a patient with a *de novo* 6q26-q27 deletion and 20p13-p12.1 duplication. *F. Faucz[1,2], H. Salomão[2], R.B. de Alexandre[1,2], A. Bonalumi[2], J. Souza[2], J.A. Rosenfeld[3], S. Raskin[2,4], V.S. Sotomaior[2].* 1) Section on Endocrinology & Genetics, PDEGEN, NICHD, NIH, Bethesda, MD; 2) Laboratory of Molecular Genetics, NIMA, PPGCS, Pontificia Universidade Catolica do Parana, Curitiba, PR, Brazil; 3) Signature Genomics, Spokane, Washington, DC; 4) 3Genetika - Centro de Aconselhamento e Laboratório de Genética, Curitiba, PR, Brazil.

Wolf-Hirschhorn syndrome (WHS [MIM 194190]) is a chromosome deletion syndrome with a well delineated phenotype. The vast majority of cases are the result of hemizygosity of 4p16.3. Many patients carry a deletion of several megabases that results in severe mental and growth retardation, major malformations and seizures, and a characteristic facial appearance with wide forehead, large and protruding eyes, hypertelorism, prominent glabella, down-turned mouth, and micrognathia. Here we report the case of a seven-year-old girl, child of a non-consanguineous and a healthy 24-year-old mother and 33-year-old father, that was initially referred to medical examination with clinical suspicion of Wolf-Hirschhorn syndrome. Pregnancy was unremarkable, she was born at 40 weeks of gestation with a birth weight of 2760 g (5^{th}-10^{th} percentile), a length of 47 cm (5^{th}-10^{th} percentile). Apgar scores were 1 and 3, and cardiac arrest was reported at birth. She has facial dysmorphia, microcephaly, thin upper lip, epicanthic fold, strabismus, enlarged nasal base, posterior cranial fossa anomaly, recurrent urinary tract infections, right urinary lithiasis, horseshoe kidney, vesicoureteral reflux, intestinal volvulus,hip dislocation, vertebral scoliosis, butterfly vertebrae, syndactyly of 2nd and 3rd toes, right thumb without movement, and generalized hypotonia. Seizures were reported at three years of age. Motor development and language development were delayed. Karyotype was normal and Array CGH analysis of 1543 loci using 4685 BAC clones detected two abnormalities: a single copy loss of 29 BAC clones from the terminal region of the long arm of chromosome 6 at 6q26-q27 (7.5Mb - 18 genes); and a single copy gain of 46 BAC clones from the terminal end of the short arm of chromosome 20 at 20p13-p12.1 (16.3 Mb - 65 genes). FISH analysis confirmed a derivative chromosome 6 that shows deletion of 6q26-6q terminal and the presence of a hybridization signal from the chromosome 20pter-20p12.1 clone. This derivative chromosome 6 results in partial monosomy of 6qter and partial trisomy of 20pter. Both parents had normal aCGH. Thus the derivative chromosome identified in this patient is *de novo*. The clinical manifestations and phenotype-genotype correlation in our patient match no known syndrome described in the literature, therefore, this description is of importance for characterization of a new syndrome.

3056F

Submicroscopic chromosomal rearrangements in patients with an Angelman syndrome-like phenotype. *K. Hosoki[1], T. Ohta[2], N. Niikawa[2], S. Saitoh[3].* 1) Pediatrics, Hokkaido Univ Grad Medicine, Sapporo, Japan; 2) RIPHS, Health Sciences Univ of Hokkaido, Tobetsu, Japan; 3) Pediatrics and Neonatology, Nagoya City Univ Grad School of Medical Sciences, Nogoya, Japan.

Angelman syndrome (AS) is characteristic by severe developmental delay, seizure, ataxic movement and inappropriate laughter. AS is caused by loss-of-function of maternally inherited*UBE3A* located in 15q11-q13. Although clinical features of AS are clinically discernable, AS-like features are seen in other genetic disorders including Rett syndrome and Pitt-Hopkins syndrome. In a series of genetic investigation of patients suspected to have AS, we have collected 33 patients with an AS-like phenotype for whom routine genetic tests for AS (*SNRPN* DNA methylation test, FISH analysis, *UBE3A* mutation screening test) revealed negative results. We performed whole genome microarray analysis using Genome-Wide Human SNP Array 5.0 (Affymetrix) to these patients. We identified three patients with copy number variation (CNV) larger than 300kb (0.3Mb 6q13 deletion, 5.6Mb 10q11 deletion, 0.3Mb 21q22.3 deletion). 10q11 deletion was proven to be *de novo*, but parental samples were not available for 6q13 deletion and 21q22.3 deletion. Recently 10q11 rearrangements were reported in patients with variable clinical features including developmental delay, intellectual disability, epilepsy and ataxia. Our results further confirmed that 10q11 deletion could cause intellectual disability, epilepsy and ataxia resembling AS. Pathogenicity of 6q13 deletion and 21q22.3 deletion are not conclusive but these deletions contain several functional genes. Therefore, these deletions may be associated with AS-like phenotype. In conclusion, we detected three submicroscopic deletions in 33 patients with an AS-like phenotype. Submicroscopic chromosomal rearrangements may be associated with an AS-like phenotype although the frequency appears to be low.

3057W

Twenty-two year follow up of identical twins with discordant phenotype due to a ring 13 chromosomal mosaicism in one of them. *Y. Lacassie.* Dept Ped/Div Clin Gen, LSU Hlth Sci Ctr, and Children's Hospital, New Orleans, LA.

At the 43rd Annual Meeting of the ASHG in 1993 we reported identical twins with discordant phenotype due to a ring 13 chromosomal mosaic syndrome in one of them. Her major manifestations included IUGR, FTT, developmental delay/intellectual disability, hemihypotrophy of the left side of the body with leg length discrepancy, profound hearing loss at left due to inner ear malformations, dental anomalies mainly on the left side, dislocation left hip, distinctive left hand with mbd/clinodactyly of the 5th finger, short left 3rd metatarsal, and telecanthus, congenital torticollis due to Klippel-Feil anomaly, 13 ribs, scoliosis and wider halluces and short 5th toes. At 3 months hypotonia and developmental delay were first noticed. Chromosomes on blood at age 31\2 years were normal and a possible Silver-Russell syndrome was diagnosed. Since our first evaluation and establishment of diagnosis in 1990, at age 4, through chromosomal studies on skin fibroblasts from both sides of the body, we have had the chance to follow up this patient annually for 22 years. Early intervention programs, posterior transpedicular hemi-epiphysiodesis of cervical spine performed at age 61\2, epiphysiodesis of the right distal femur at age 101\2, and Bone Anked Hearing Aid (BAHA) implanted at age 24 for her severe/profound sensory neural hearing loss at left, have allowed her a relative normal life with better prognosis than predicted early in life. She had her menarca at age 138/12 years and stopped growing at age 17. The periodic follow up of her normal identical twin has allowed the physical, mental and dysmorphological comparisons along the years. Her normal twin is 16 cm taller, her OFC is 1 cm larger being in the 50th centile, and has normal superior intelligence. However, streaks of hypopigmentation in the right shoulder and arm, fibular deviation of halluces and 2nd toes and minor clinodactyly of the 5th toes have suggested the possible transfer of abnormal cells during gestation. The detection of a ring 13 chromosomal mosaicism with greater involvement on fibroblasts from the left side of the body, in agreement with the clinical manifestations, and not found on peripheral blood using standard chromosomal analysis and aCGH, stresses the importance of detailed clinical dysmorphological evaluation.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3058F

12p micro RNA expression in Fibroblast Cell Lines from Probands with Pallister-Killian Syndrome. *K. Izumi[1], Z. Zhang[2], M. Kaur[1], I. Krantz[1].* 1) Human Genetics, Children's Hospital of Philadelphia, Philadelphia, PA; 2) Center for Biomedical Informatics, Children's Hospital of Philadelphia, PA.

Pallister-Killian syndrome (PKS) is a multi-system developmental disorder caused by tetrasomy 12p that exhibits tissue-limited mosaicism. The spectrum of clinical manifestations in PKS includes craniofacial dysmorphisms, cleft palate, congenital diaphragmatic hernias, congenital heart defects, involvement of the audiologic, musculoskeletal, gastrointestinal and genitourinary systems, cutaneous anomalies, intellectual disability and seizures. In an attempt to understand the molecular basis of the developmental differences observed in PKS, we performed genome-wide expression analysis on skin fibroblasts from probands and demonstrated a unique gene expression pattern for PKS. As a potential molecular driver of such a globally dysregulated gene expression pattern, we evaluated the potential contribution of micro RNAs. There are 6 micro RNAs (miR-141, miR-200c, miR-613, miR-614, mir-920 and miR-1244) located on 12p. The expression levels of 5 of these micro RNAs (excluding mir-920) were examined using the Nanostring n-Courier system. The gene expression levels of computer-predicted target genes for each of these micro RNAs were measured concurrently. Micro RNA expression was measured using skin fibroblast cell lines obtained from 12 probands with PKS and 5 control individuals. All micro RNAs located on 12p were overexpressed in PKS fibroblasts, although the fold difference of the expression level was 1.07 to 1.29, and is less than that of the genomic copy number difference. Among the 12p micro RNAs, miR-1244 had the highest fold difference of 1.29. One of the computer-predicted targets of miR-1244 is the MEIS2 gene (located on chromosome 15q14), whose haploinsufficiency is associated with congenital heart defects and cleft palate. The expression level of MEIS2 was significantly down regulated in PKS probands. The fold difference of miR-200c was 1.25, and its known targets include ZFPM2 (located on chromosome 8q23.1), whose haploinsufficiency is associated with congenital heart disease and congenital diaphragmatic hernia. The expression level of ZFPM2 was also downregulated in PKS probands. Given that the degree of differential expression of these micro RNAs was rather mild, the contribution of extra copies of micro RNAs may be small, however, overexpression of miR-1244 and miR-200c due to tetrasomy 12p may contribute to the pathogenesis of congenital heart defects, cleft palate and diaphragmatic hernia via modulating the expression of MEIS2 and ZFPM2.

3059W

Genetics of precocious puberty: A proband with Klinefelter syndrome, maternal uniparental disomy 14 and precocious puberty. *K. Reddy, H. Bass, J. Keni.* Kaiser Permanente Southern California, Los Angeles, CA.

Genetics of precocious puberty: A proband with Klinefelter syndrome, maternal uniparental disomy 14 and precocious puberty. Reddy KS1, Bass HN1, Keni J2 Medical Genetics1 and Pediatric Endocrinology2, Kaiser Permanente Southern CA kavita.s.reddy@kp.org A 9-year-old boy with precocious puberty and learning difficulties was found to have a 47,XXY karyotype. Parental peripheral blood chromosome studies identified a Robertsonian (Rob) 13/14 translocation in the father. An oligonucleotide and SNP array CGH study on the proband uncovered mixed maternal uniparental disomy (matUPD) for chromosome 14. Two large blocks of maternal isodisomy in the pericentromeric and terminal regions with intervening heterodisomy were detected and corresponded with a meiosis II non-disjunction error with a double crossover and subsequent trisomy rescue or gamete complementation. Since the father carries a Rob 13/14 translocation, he has an increased risk for producing nullisomy 14 sperm, supporting the possibility of gamete complementation. The clinical work-up of the patient demonstrated a hormone profile for central precocious puberty, a microadenoma of the pituitary by MRI, normal chest CT, normal serum alpha-fetoprotein and hCG, premature growth spurt and obesity. Klinefelter syndrome, upd14mat, obesity, pituitary macro- or microadenoma, and hCG-producing intracranial or mediastinal germ-cell tumors, either independently or in combination, have been associated with precocious puberty. It is probable that these genetic mutations predispose to precocious puberty through tumor formation and obesity. Management of the patient involves treatment with a GnRH analog (leuprolide) and dietary changes to control obesity.

3060F

Molecular genetic testing of recurrent anencephaly in a family without partial trisomy 2p22-pter. *C. Sergi[1], J. Gekas[2], D. Kamnasaran[3].* 1) Department of Laboratory Medicine and Pathology, University of Alberta, Edmonton, Canada; 2) Human Genetics Research Unit, Centre de recherche de l'hôpital Saint-François d'Assise, Québec, Canada; 3) Department of Pediatrics, Laval University, Québec, Canada.

Background: Human neural tube closure involves a complex interplay between genetic and environmental factors. In fact, failed closure of the anterior region by day 28 results in anencephaly. The incidence of anencephaly is estimated to be as high as 1/1000 pregnancies, and about 1.2/10 000 live births in North America. Unfortunately, fetal loss, still births or early neonatal deaths are the prognostic consequences of anencephalic pregnancies. We present a family with recurrent anencephaly and genetic testing of the *VANGL1* and *FOXN1* as candidate genes for anencephaly. **Methods and Results:** The mother, currently 28 years old with 4 gravida, 3 abortus, had a history of recurrent pregnancy losses and with three anencephalic fetuses in separate pregnancies. Both parents are clinically normal and with an unremarkable family history. Standard high G-banding cytogenetic studies performed on the chromosomes of the family members' peripheral blood lymphocytes and/or prenatal amniocytes did not identify any numerically and structurally abnormalities. Genes of the planar cell polarity pathway are implicated in neural tube defects. We examined the coding regions and intron-exon boundaries of *VANGL1* on human chromosome 1p13.1 and *FOXN1* on human chromosome 17q11-q12, genes which have phenotypes within the anencephaly spectrum. However, despite our rigorous testing, no DNA variations or mutations were identified among these genes.**Conclusion:** Familial recurrence of anencephaly is extremely rare and has been previously associated with a structural abnormality of human chromosome 2p. We suggest that other Planar Cell Polarity genes are an intriguing research avenue to further investigate families with recurrent neural tube defects in search of a genetic etiology.

3061W

Urethrocystoscopy, diagnostic laparoscopy, versus testicular descent gonadectomy in a patient with ambiguous genitalia: A case report. *F. Cuellar-López[1], J. M. Aparicio[2,4], I. de la Torre[3].* 1) Pediatric Urology; 2) Genetics; 3) Surgery, Hosp Para el Nino Poblano, Puebla, Puebla; 4) Estomatology, Benemerita UNiversidad Autonoma de Puebla, Mexico.

A pediatric patient with normal male caryotype, where ambiguous genitalia was observed, evaluated and corrected by surgery at the Hospital Para El Nino Poblano in Mexico. Disorders of sexual development (DSD), formerly termed intersex conditions, are considered a rare clinical event and it is important an early genetic and clinical diagnosis in order to counsel parents on therapeutic options. However, the problem of early gender assignment has been challenged by the results of clinical and basic science research, which show that gender identity development likely begins in uterus. Here the techniques for surgical genital reconstruction have been associated to psychological, interfamilial and social implications of gender assignment which in some cases became a real problem for surgical reconstruction. Recently, the Lawson Wilkins Pediatric Endocrine Society, (LWPES, 2008) and the European Society for Pediatric Endocrinology, (ESPE, 2009) have published proposed changes to the nomenclature and definitions of disorders in which the development of chromosomal, gonadal, or phenotypic sex is atypical. They proposed to change the nomenclature to reflect advances in our understanding of the pathophysiology of these sexual disorders in order to the help to the concerns of affected patients. Previous terminology and revised nomenclature of disorders of sexual development (DSD) based on a new nomenclature. Results:The patient with normal chromosomal study, 46 XY, was evaluated in a medical and surgical multidisciplinary manner due to ambiguous genitalia. It was confirmed the presence of vagina and cervix by vaginoscopy and urethrocystoscopy. The vagina size of 7 centimetres of length. by ureteral catheter, the cervix was observed andhysterosalpingography was performed. By laparoscopy process, intra-abdominal testis, left testicular descent effected video-assisted was observed, achieving distal third take up only channel is where ipsilateral. Inguinal testicular biopsy was done. In relation to the right side, a remanent tissue bulk form was found. Resection was performed where histopatological studies reported uterine horn. Resection of vagina, uterus and fallopian tubes by posterior sagittal approach with lifting and moving straight. Moreover, testicular hypotrophy was detected at left order to assess after its resection and testicular prosthesis was used. Make consequential within mixed plastia hypospadias, Snodgrass and Duckett.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3062F

Congenital association VACTERL at the Hospital Para El Nino Poblano, Mexico: Fourteen cases report. *A. Garcia-Guzman[1], J. M. Aparicio-Rodriguez[2,3].* 1) Pediatrics; 2) Genetics, Hospital para el Nino Poblano, Puebla; 3) Estomatology, Benemerita Universidad Autonoma de Puebla, Mexico.

INTRODUCTION. VACTERL association has been described in the literature since 1982. At the beginning only four clinical characteristics were described in comparison with six described nowadays; (V) Vertebral, (A) anal Atresy, (C) cardiac anomalies, (TE) Traqueal-Esophagic fistula, (R) Renal malformations, (L) limbs anomalies. This congenital association has 0.3 to 2.1 cases per 10,000 new born incidences. In México, few clinical cases have been reported, the main goal of this study is to analyze the different clinical patterns among 14 patients with this genetically disease in a third level pediatric Hospital in Mexico. MATERIAL AND METHODS. A Descriptive, retrospective and transversal study was performed among 14 patients diagnosed with VACTERL (with a minimal of three clinical features). RESULTS AND DISCUSSION. The male gender was observed to have an increased incidence if compared to females patients (6;1). 79% was diagnosed at new born stage. The main clinical features founded among the studies population were; 1. Anal atresy (83%) associated to ano-rectal malformation with fistula. 2. Cardiovascular malformations (92%) like Inter Auricular Communication (67%) 3. No VACTERL association as Ambiguous genitalia was also observed (57%) with carytotype in all cases. 4. Annular pancreas (14.28%) The VACTERL association in this study was observed with several different clinical symptoms: A. Three and four clinical symptoms 35.71% B. Five clinical symptoms 21.42% C. Six clinical symptoms 7.14% It is important to have a earlier diagnosis with a medical and surgical treatment to ensure a better quality of life for all patients with these congenital malformations.

3063W

Orthodontic correction in malformated unilateral teeth structure in a patient with pseudoachondroplasia: A case report. *S. Ochoa[1], J. M. Aparicio[2,4], D. D. A. Camarillo[3], S. S. Cabrera[1].* 1) Orthodontics; 2) Genetics; 3) Pediatric Estomatology, Hospital para el Nino Poblano, Puebla, Puebla, Mexico; 4) Estomatology, Benemerita Universidad Autonoma de Puebla.

Pseudoachondroplasia is a genetic disease characterized by a growth factor alteration, is a disorder of bone growth; is characterized by severe growth deficiency and deformations such as bow legs and hyperlordosis. Prevalence is estimated at around 1/60,000. The disorder is usually discovered during the second year of life with the onset of slow growth and walking difficulties. The short stature becomes more prominent with age and the hands and feet appear short and wide. Joint laxity is a general clinical feature, but predominantly affects the hands. Defective epiphyseal growth causes early arthrosis. The limb deformation is caused by metaphyseal lesions. Genetic transmission is autosomal dominant, but most isolated cases are due to de novo mutations. The disorder is caused by small mutations or deletions in the COMP gene (19p13.1) coding for the cartilage oligomeric matrix protein. Diagnosis is made on the basis of epiphyseal and metaphyseal anomalies detected on radiographs during the second year of life. The principle differential diagnosis is achondroplasia, the meaning of the word (a=without; chondro=cartilage and plasia=development), literally means “without cartilage formation,” but the craniofacial anomalies present in this disorder are absent in patients with pseudoachondroplasia and radiographic findings differ significantly. Forms of multiple epiphyseal dysplasia may also be included in the differential diagnosis. Genetic counseling may be proposed and the recurrence risk is 50%. Prenatal diagnosis is feasible if the mutation has been detected in an affected parent. Treatment is based on physiotherapy, management of the spinal deformation and corrective orthopedic surgery. Intensive physical activity should be avoided.

3064F

Hearing loss frequency in a pediatric patient group at a third level hospital. *A. Romero[1], J. M. R. Aparicio[2], L. P. C. Gallegos[1].* 1) Audiology; 2) genetics, Hospital Para el Niño Poblano, Puebla, puebla, Mexico.

The hearing is the usual way to acquire spoken language; one of the most important human attributes and means of communication in all cultures. Language is the main way of learning in children and plays a key role in thinking, knowledge, and intellectual development. The ear is an active sense detected used through our lives, even where sleeping. 75% of the information is received by both ears. Hearing loss is particularly dramatic in childhood, as the intellectual and social development of the child is intimately linked to the cortices of the central nerve system, which allow us to acknowledge of the surrounding word and the correct language acquisition. In modern times, in 1945 Carhart and Canfield coined the word “Audiology”, thus, it was born the modern science of hearing. Just eight years later, in 1953, the National Institute of Audiology was founded. By the fact that the foundation of this institution marked a historic milestone to be advanced to any similar in developed countries. it has been investigated the association between hearing loss and low weight at birth. With the implementation of potential evoked auditory brain stem in infants graduates of neonatal intensive care (NICU) units Hecox and Galambos in the early 1970s offers 100% sensitivity and specificity of 86%. Currently, hearing in infants, pre-school and school sieve is something quite unknown in our country, even among the medical staff, and is limited to very few institutions, some of them belongs to a private sector. The initial experiences in this area date back to 1987, when he placed an cochlear implant at the General Hospital ;Dr. Manuel Gea González; in which these operations continue carrying out. This operational experience are considered important for scientific contributions in Mexico. Also, aware of the realization of single channel and multichannel, Guadalajara and Monterrey cities in Mexico, implants initiated between the years 1996 and 1998. In some hospitals in Mexico, the Neonatal auditory sieve is performed, This causes the vast majority of cases of hearing loss and deafness which is generally detected late; In addition, the staffing of auditory prosthesis covers sparsely to people who need them; and only benefit with cochlear implants a minimum percentage of them. Today among the child population, hearing loss (decrease of the level of below-normal hearing) is the more frequent defect, beating Down syndrome and infantile.

3065W

Diagnostic Exome Sequencing Reveals a de novo Mutation in the DYNC1H1 Gene in a Sporadic Case of Developmental Delay, Seizures, and Polymicrogyria. *K.D. Gonzalez[1], J. Wei[1], X. Li[1], H.M. Lu[1], H. Lu[1], J.S. Cohen[2], R. McClellan[2], S. Naidu[2], W. Zeng[1].* 1) Ambry Genetics, Aliso Viejo, CA; 2) Department of Neurogenetics, Kennedy Krieger Institute, Baltimore, MD.

Clinical diagnostic exome sequencing has rapidly made its way from research to the clinical setting and has been instrumental in successfully providing a molecular diagnosis for families who had previously been unsuccessful in their pursuit of the underlying disease etiology. We sequenced protein-coding regions of the genome (exome) in a 2 1\2 year old female with global developmental delay, microcephaly, seizures, hypotonia, behavioral difficulties, and MRI findings of polymicrogyria. The family history was negative for similar phenotypes. The patient had evaded diagnosis through clinical evaluation and extensive testing over many years including negative *MECP2, TUBB2, WDR62*, and *GPR56* gene sequence analyses, as well as uninformative biochemical and SNP-array results. The patient was heterozygous for a missense mutation (c. 4700G>A; p.R1567Q) in exon 22 of the *DYNC1H1* gene and co-segregation analysis showed that her mother and father did not carry the mutation, indicating a *de novo* mutation occurrence. The amino acid is completely conserved throughout vertebrates. *DYNC1H1*interacts with the LIS1 gene in which mutations cause gross disorganization of cortical neurons leading to lissencephaly. Mutations in *DYNC1H1* lead to neuronal degeneration. *DYNC1H1*mutations have been identified in patients with central and peripheral neuronal defects including mental retardation, speech and motor developmental delay, hypotonia, and microcephaly, Charcot-Marie-Tooth (CMT2), epileptic seizures, brachycephaly, dysmorphic features, and lower extremity spinal muscular atrophy (SMA). The identified mutation is highly consistent with the patient's clinical symptoms and, after years of unsuccessful analyses, led to a molecular diagnosis for the family. Additionally, the findings further delineate the phenotypic spectrum of patients with mutations in the *DYNC1H1* gene.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3066F

Congenital poikiloderma, fatty infiltration of muscles and pulmonary fibrosis: a new syndrome. *S. Mercier[1,2], S. Kury[3], JM. Mussini[2,4], A. Magot[2,5], B. Isidor[1], S. Barbarot[6], A. David[1], S. Bezieau[3].* 1) Service de Génétique Médicale, Unité de Génétique Clinique, CHU de Nantes, Nantes, France; 2) Centre de Référence des Maladies Neuromusculaires Nantes-Angers, CHU de Nantes, Nantes, France; 3) Service de Génétique Médicale, Unité de Génétique Moléculaire, CHU de Nantes, Nantes, France; 4) Laboratoire d'Anatomie Pathologique A, CHU de Nantes, Nantes, France; 5) Laboratoire d'Explorations Fonctionnelles, CHU de Nantes, Nantes, France; 6) Service de Dermatologie, CHU de Nantes, Nantes, France.

Congenital poikiloderma is a key feature in Rothmund-Thomson syndrome (RTS). We report a sporadic case of congenital poikiloderma, hypopigmentation, hypotrichosis and hypohidrosis in a 7 year-old boy, initially diagnosed with RTS-like. At the age of 6 years, he presented with distal lymphoedema of the four limbs. Simultaneously, a rapidly progressive muscle weakness occurs with primary distal lower limb predominance, and secondary clinical extension to proximal upper and lower limbs. Wheel chair use was required at the age of 7 years. Serum creatine kinase level was normal and electromyographic examination showed a myogenic pattern in lower limbs. Muscle MRI showed a very diffuse fatty infiltration of lower limbs, which was confirmed by the quadriceps muscle biopsy. Microscopy of the skin revealed elastic tissue degeneration with alterations of the elastic network in the superficial and deep dermis. Respiratory function tests revealed a mild restrictive impairment. RTS was excluded on clinical evolution and normal *RECQL4* gene analysis. A similar report with poikiloderma, tendon contracture, pulmonary fibrosis and identical histological anomalies has been described in a South African family with autosomal dominant inheritance suggesting a *de novo* mutation in our patient. Array-CGH was normal and trio-based exome sequencing is pending. This second report highly supports that this syndrome should be considered as a distinct entity.

3067W

Prenatal diagnosis of a fetus with a rare association of a multiple heart and renal malformations: case report. *M. Dumitrescu[1,3], D.F. Albu[2,3,4], C.C. Albu[2,3], A. Oncescu[2,3], E. Severin[2].* 1) Marius Nasta Hospital, Bucharest, Romania, Bucharest, Romania; 2) University of Medicine and Pharmacy Carol Davila Bucharest, Romania, Bucharest, Romania; 3) Alco San Clinic, Bucharest, Romania; 4) Panait Sarbu Hospital, Bucharest, Romania.

A 32-year-old pregnant Caucasian woman was reffered for prenatal ultrasound examinations. The sonoghraphic fetal monitoring demonstrated a rare association of a multiple heart and renal malformations: absence of valve of Vieussens, common arterial trunk, unilateral atrial enlargement, severe bilateral hidronephrosis and polyhidramnios. Amniocentesis was performed during the 18th week of pregnacy. Fetal chromosomal analysis showed abnormal karyotype: 46XX,add 12(p13.3). The parents decided to continue the pregnancy. Unfortunately the baby died two weeks later. Postnatal ultrasounds and autopsy confirmed our sonographic findings. Conclusion: Prenatal evaluation both sonographic and genetic was very useful for the pregnancy management. In this family case, any future pregnancy should be be karyotyped for further evaluation. The necessity of routine ultrasound examination for the prenatal diagnosis of chromosomal abnormatities to all pregnancies.

3068F

Complex chromosomal translocation leading to a dual diagnosis of Prader-Willi Syndrome and Cri-du-chat in a 14 year old boy illustrating the importance of re-evaluation in an individual with atypical Prader-Willi phenotype. *J.A. Gold[1, 2], S. Ramanathan[1].* 1) Pediatrics, Loma Linda University Medical Center, Loma Linda, CA. Dept. Pediatrics Division Genetics and Metabolism; 2) University California Irvine University Medical Center Dept Pediatrics Division Genetics and Metabolism.

Introduction: Prader-Willi syndrome (PWS) affects 1/15,000-1/30,000, individuals. It is characterized by prenatal-onset hypotonia, infantile poor feeding and failure to thrive. This is followed by hyperphagia, childhood-onset obesity, short stature, facial dysmorphic features, psychomotor delay, and a distinct behavioral phenotype. Lack of expression from the paternally inherited copy of the PWS/Angelman (AS) critical region on chromosome 15 at 15q11–q13 region causes PWS. Three molecular subtypes of changes include paternal deletion, maternal uniparental disomy (UPD) and imprinting center defects. Cri-du-chat Syndrome or 5p minus syndrome affects 1/20,000-1/50,000 individuals. It is associated with the deletion of short arm of chromosome 5 of varying sizes from extremely small to the complete p arm of chromosome 5. It is characterized by a typical cat like cry at birth, microcephaly, hypertelorism, epicanthic folds, low-set dysplastic ears, hypotonia and severe psychomotor and developmental delay. Case Report: 14 year old boy presented with features of Prader-Willi syndrome due to an unbalanced chromosomal translocation between chromosome 15 and chromosome 5. It has resulted in a deletion of the PWS region on chromosome 15q and a small deletion on chromosome 5p Cri-du-chat region. Microarray result: 8.9 Mb proximal deletion of 15q11.2-q13.1, consistent with Type I deletion for the PWS/AS syndrome region; 6.533 Mb terminal deletion of 5pter-p15.31, consistent with Cri-du-chat syndrome. We will discuss the overlapping features of Prader-Willi and Cri-du-chat syndromes as well as any unusual features for either condition. We do not know how the patient's genotype will affect or modify the expression of the phenotypes for these disorders in him. To our knowledge this is the first patient described with both conditions. Unfortunately this has made it extremely difficult to give accurate information on the natural history and long term out comes with any complications.

3069W

MOLECULAR CHARACTERIZATION OF SHOX GENE AND REGULATORY REGIONS IN PATIENTS WITH IDIOPATHIC SHORT STATURE FROM THREE MEDICAL CENTERS IN OF BOGOTA, COLOMBIA. *T. Vinasco[1], G. Jaimes[2], M. Coll[3], C. Cespedes[4], H. Velasco[5].* 1) Morfology, National University, Bogota, Cundinamarca, Colombia; 2) Central Military Hospital, Bogota, Colombia; 3) Hospital de la Misericordia, Bogota, Colombia; 4) San Ignacio Hospital, Bogota, Colombia; 5) MD, MSc, E. Genetic Institute, National University, Bogota, Cundinamarca, Colombia,.

Short stature is a concept that encompasses normal variants and pathologic variants. One of these variants is Idiophatic Short Stature (ISS) where patients do not have corporal asymmetry, growth defiency, nor chromosomal anomalies or any syndromatic sign. Although several genes have been described in ISS, mutations in SHOX gene has been found in a high frequency in these patients (around 3 to 15%). Our study evaluated 52 patients with ISS from three national referral medical centers and sought any type of deletion / duplication of SHOX gene and conserved non-coding regions (CNE) by MLPA (MRC Holland, SALSA MLPA KIT P018-E1 SHOX y P018-F1 upload version). We found that 9.6% patients had mutations in SHOX, 4 patients had deletions (Exon 1, Exon 2, Exon 6 and Intron 6b) and one had duplication (CNE 9). Two of these mutatios were new (Exon 6b deletion and CNE 9 duplication), one was de novo, 3 mutations were inherited from one of the parents and in the remainder, we could not confirm the source. All patients had height below -2.5 SD and some patients were on growth hormone therapy. This is the second Latin American study reporting SHOX mutations for ISS patients. Sequencing is desirable in negative cases (47) looking for punctual mutation.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3070F

Clinical and molecular characterization of non-syndromic craniosynostosis: an International Consortium Approach. *M.L. Cunningham[1], P.A. Romitti[2], C.M. Justice[3], A.F. Wilson[3], T. Roscioli[4], E. Oláh[5], B. Bessenyei[5], M.R. Passos-Bueno[6], B. Wollnik[7], A.O.M. Wilkie[8], S.A. Boyadjiev Boyd[9], International Craniosynostosis Consortium.* 1) Dept of Pediatrics, Division of Craniofacial Medicine, University of Washington and Seattle Children's Research Institute, Seattle, WA, USA; 2) Department of Epidemiology, College of Public Health, The University of Iowa, Iowa City, IA; 3) Genometrics Section, IDRB, Division of Intramural Research, NHGRI, NIH, Baltimore, MD; 4) School of Women's and Children's Health, Sydney Children's Hospital, University of New South Wales, Sydney, Australia; 5) Department of Genetics, Debrecen, Hungary; 6) Department of Genetics, University of Sao Paolo, Brazil; 7) Institute of Human Genetics, University of Cologne, Germany; 8) Craniofacial Unit, Oxford University Hospitals NHS Trust, John Radcliffe Hospital, Oxford, United Kingdom; 9) Department of Pediatrics, Section of Genetics, University of California Davis, Sacramento, CA.

Non-syndromic craniosynostosis (NSC), the premature closure of one or more of the cranial vault sutures is a common congenital malformation, affecting approximately one out of 2,000 newborns. Rare mutations in the FGFR2, TWIST1, FREM1, LRIT3, EFNA4, ALX4, and RUNX2 duplications have been reported in a minor fraction of NSC cases; however, the etiology of this birth defect is not well understood. In an attempt to better characterize the clinical phenotype and to identify susceptibility loci and contributing environmental factors we have established an International Craniosynostosis Consortium (ICC) coordinating the effort of physicians and researchers at 10 sites in the United States, 4 craniofacial units in the United Kingdom, the Hungarian National Registry for Birth Defects and several major craniofacial centers in Australia, Brazil and Germany. A total of 741 families have been recruited and characterized and clinical databases with sample repository were created. A study website allows collaborators, referring physicians and members of the affected families to contribute to the goals of ICC (https://genetics.ucdmc.ucdavis.edu/icc.cfm). As a result, RUNX2, ALX4, FREM1, and LRIT3 variants predisposing to NSC were identified and reported by our group. The findings of the first genome-wide association study suggested a role for BMP2 and BBS9 in sagittal NSC. The goal of this paper is to present our current approach and to inform the genetics community of possible ways to contribute to the mission of the ICC (contact: simeon.boyd@ucdmc.ucdavis.edu).

3071W

Molecular prenatal diagnosis of a sporadic alobar holoprosencephalic fetus: Genotype-phenotype correlations. *J. Gekas[1], C. Sergi[2], D. Kamnasaran[3].* 1) Human Genetics Research Unit, Centre de recherche de l'hôpital Saint-François d'Assise, Québec, Canada; 2) Department of Laboratory Medicine and Pathology, University of Alberta, Edmonton, Canada; 3) Department of Pediatrics, Laval University, Québec, Canada.

Background: Holoprosencephaly (HPE) is principal defect in the developing prosencephalon with an incidence of 1 in 250 conceptuses and 1 in 10-16,000 live newborn infants. Current prenatal diagnostic protocols on the clinical management of holoprosencephaly involve combined fetal imaging analyses with cytogenetic and molecular genetic testing on suspected cases. To date, mutations are identified in 5–10% of nonsyndromic, non-chromosomal cases in at least 12 genes. We report the molecular prenatal diagnosis of a fetus with alobar holoprosencephaly. **Methods and Results:** CTG band karyotyping and array CGH genome-wide cytogenetic screenings were done, in conjunction with DNA sequence analyses of the *SHH*,*ZIC2*, *SIX3* and *TGIF* genes in search of a molecular etiology. Our proband was conceived by a 37-year-old healthy mother (2 Para, 1 Abortus) and is the second child of non-consanguineous parents with an unremarkable family history. At 18 gestational weeks, transabdominal 2D ultrasonography detected a gestational age appropriate fetus with an alobar holoprosencephalic brain that died in-utero. Post mortem examination disclosed findings of a male alobar holoprosecenphalic fetus with facial findings of cyclopia and absence of the nasal structures. Standard CTG band karyotyping and array CGH did not disclose any plausible findings. However, by using a candidate gene approach, we identified the proband with a sporadic, non-polymorphic, heterozygous c.298 C_T substitution non-sense mutation in the first exon of the *SHH* gene. No mutations were detected in other HPE associated genes: *ZIC2*, *SIX3* and *TGIF*. The proband mutation is predicted to encode a non-functional protein (p.Gln100Ter) which truncates before the cholesterol mediated autoproteolytic cleavage site. A comparison of these findings to the literature identifies three previous nonsyndromic HPE cases with heterozygous substitution mutations in the same codon 100 of the *SHH* gene, but with extreme variability in the holoprosencephaly phentoypes. **Conclusion:** We report a new fetal case manifesting alobar holoprosencephaly with codon 100 of the *SHH* gene having a hotspot for loss-of-function mutations. Mutations in codon 100 of SHH are identified in both sporadic and autosomal dominant inherited cases and with demonstrative evidence of variable expressivity and penetrance of the prosencephalon development deficiency spectrum.

3072F

Sonic Hedgehog (SHH)regulatory region deletion in a patient with Holoprosencephaly. *L. Ribeiro-Bicudo[1], R. Quiezi[1], M. Ansari[2], K. Williamson[2], A. Richieri-Costa[1], D. FitzPatrick[2].* 1) Genetics, Hospital for Rehabilitation of Craniofacial Anomal, Bauru, Sao Paulo, Brazil; 2) MRC Medical Center Research, Edinburgh, UK.

The craniofacial development is a complex and dynamic process which involves diverse stages of embryonic life. Holoprosencephaly (HPE) is one of the most common craniofacial malformations characterized by the incomplete separation of the cerebral hemispheres into distinct right and left halves. Clinical expression is extremely variable ranging from minor facial signs to complexes craniofacial anomalies such as ciclopia. Structural chromosomal anomalies previously compiled from chromosomal HPE data predicted at least 12 different HPE loci that include SHH, SIX3, TGIF, PATCH1, TGF1, FOXH1, ZIC2, SIX3, GAS1 e GLI2, and among these genes, SHH appears to be the major one, accounting for 50% of the identified mutations. We report a Brazilian patient presenting a mild form of lobar HPE associated with eye anomalies who presented microcephaly, neurodevelopmental delay, bilateral microphthalmia, more severe in the right eye, convergent strabismus, pendular nystagmus, nasal septum agenesis, flat face, cleft lip/palate surgical scars, an upper placed pre-maxilla with a single maxillary central incisor, bilateral ungueal hypoplasia of the 5° toe. Ophthalmological evaluation showed right microcornea and left retinal coloboma. G-banded chromosomes in peripheral lymphocytes were normal. Mutation analysis for the genes SOX2, OTX2, PAX6, SHH, SIX3, TGIF, ZIC2 and BMP4 were unremarkable. Array-CGH analysis showed a submicroscopic microdeletion of 540 kb at 7q36.3, the deletion occurred approximately between 150 kb up to 694 kb upstream of the SHH promoter, encompassing the SHH regulatory region and affecting transcription of this gene. This study was approved by the Ethics Committee of the Hospital de Reabilitação de Anomalias Craniofaciais, Bauru/Brazil (HRAC). In mice, the Shh gene has six enhancers that regulate Shh transcription in the embryonic forebrain. The Shh floorplate enhancers (SFPE2) and Shh brain enhancer (SBE1) are localized approximated 200 kb downstream from Shh promoter, the other enhancers, SFPE1, SBE4, SBE2 and SBE3 are approximated 400 kb upstream from Shh promoter, and three of these enhancers are highly conserved in human. The microdeletion found in our study affected at least three of these enhancer elements (SBE2, SBE3 and SBE4) and suggests that the HPE phenotype is probably casual for loss of SHH enhancers. Grants: Fapesp Proc. n° 2011/07012-9.

3073W

Associated malformations among infants with radial ray deficiency. *C. Stoll, B. Dott, Y. Alembik, M. P. Roth.* Faculte de Medecine, Genetique Medicale, Strasbourg, France.

Infants with radial ray deficiencies very often have other associated congenital anomalies. The reported frequency and types of associated malformations vary between different studies. The purpose of this investigation was to assess the frequency and types of associated malformations among infants with radial ray deficiencies in a geographically well defined population from 1979 to 2004 of 346,831 consecutive births. Of the 73 infants with radial ray deficiencies born during this period (prevalence at birth of 2.1 per 10,000), 75 % had associated malformations. Infants with associated malformation were divided into recognizable conditions (16 (22%) infants with chromosomal and 20 (27%) with non chromosomal conditions), and non recognizable conditions (19 (26%) infants with multiple malformations). Trisomies 18 and autosomal deletions were the most frequent chromosomal abnormalities. VACTERL association, thrombocytopenia absent radii syndrome, Fanconi anemia and Holt-Oram syndrome were most often present in recognizable non chromosomal conditions. Malformations in the musculoskeletal, cardiovascular and urogenital systems were the most common other anomalies in infants with multiple malformations and non recognizable conditions. The frequency of associated malformations in infants with radial ray deficiencies emphasizes the need for a thorough investigation of these infants. Routine screening for other malformations especially musculoskeletal, cardiac and urogenital systems anomalies may need to be considered in infants with radial ray deficiencies, and referral of these infants for genetic evaluation and counseling seems warranted.One should be aware that the malformations associated with RRD can be often classified into a recognizable malformation syndrome or pattern (in 49 % of the cases).

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3074F

Orofaciodigital syndrome in a group of patients who attended Operation Smile Foundation in Colombia, between 2005 and 2012. *J. Martinez[1], I. Briceno[1,2], A. Venegas[2], S. Bohorquez[1], M. Montiel[1], A. Patiño[1], L. Arias[1,2].* 1) Universidad de La Sabana, Chia, Colombia; 2) Universidad Javeriana. Bogota Colombia.

Oral-facial-digital syndrome is a group of related conditions that affect the development of the oral cavity (the mouth and teeth), facial features, and digits (fingers and toes).13 potential forms of oral-facial-digital syndrome. The different types are classified by their patterns of signs and symptoms. However, the features of the various types overlap significantly, and some types are not well defined. The classification system for oral-facial-digital syndrome continues to evolve as researchers find more affected individuals and learn more about this disorder. Abnormalities of the oral cavity that occur in many types of oral-facial-digital syndrome include a split (cleft) in the tongue, a tongue with an unusual lobed shape, and the growth of noncancerous tumors or nodules on the tongue. Affected individuals may also have extra, missing, or defective teeth. Another common feature is an opening in the roof of the mouth (a cleft palate). Some people with oral-facial-digital syndrome have bands of extra tissue (called hyperplastic frenula) that abnormally attach the lip to the gums. Distinctive facial features often associated with oral-facial-digital syndrome include a split in the lip (a cleft lip); a wide nose with a broad, flat nasal bridge; and widely spaced eyes (hypertelorism). Abnormalities of the digits can affect both the fingers and the toes in people with oral-facial-digital syndrome. These abnormalities include fusion of certain fingers or toes (syndactyly), digits that are shorter than usual (brachydactyly), or digits that are unusually curved (clinodactyly). The presence of extra digits (polydactyly) is also seen in most forms of oral-facial-digital syndrome. Other features occur in only one or a few types of oral-facial digital syndrome. These features help distinguish the different forms of the disorder. Other forms of oral-facial-digital syndrome are characterized by neurological problems, particular changes in the structure of the brain, bone abnormalities, vision loss, and heart defects. Oral-facial-digital syndrome has an estimated incidence of 1 in 50,000 to 250,000 newborns. We conducted a study among 2500 patients attending “Operation Smile” in Bogota, 8 cases of orofaciodigital syndrome were diagnosed, which corresponds to 0.3% of the population studied. This result indicates a frequency significantly higher than that reported in the literature for general population. The Clinical description is presented.

3075W

A New type of Acrofacial Dysplasia. Prenatal diagnosis and autopsy findings. *K. Millar[1], A. Toi[2], S. Keating[3], P. Shannon[3], S. Unger[4], D. Chitayat[1,5].* 1) Department of Obstetrics and Gynaecology, The Prenatal Diagnosis and Medical Genetics Program, Mount Sinai Hospital, Toronto, Ontario, Canada; 2) Department of Diagnostic Imaging; Mount Sinai Hospital, Toronto, Ontario, Canada; 3) Department of Pathology; Mount Sinai Hospital, Toronto, Ontario, Canada; 4) Service of Medical Genetics, Centre Hospitalier Universitaire Vaudois, Lausanne, Switzerland; 5) Department of Paediatrics, Division of Clinical and Metabolic Genetics, The Hospital for Sick Children; University of Toronto, Toronto, Ontario, Canada.

Acrofacial dysostosis (AFD) is a clinically and etiologically heterogeneous group of disorders characterized by craniofacial and limb anomalies. Various sub types of AFD have been described and the types de Macena Sobreira, Bates and Rodriguez are among the lethal. We report on a new type of lethal AFD. The mother was a 23-year-old G1P0 woman of English/Irish decent and the father was 22 years old and of Hungarian/English decent. The couple was healthy, non-consanguineous and the family history was non-contributory. The fetus presented at 20.2 weeks gestation with ultrasound findings of ventriculomegaly and short femur lengths consistent with 16 weeks gestation. A repeat fetal ultrasound done by us revealed severe micromelia, absent bilateral fibula, bilateral talipes and bilateral abnormal position of both knees. The facial profile was flat with micrognathia and there was moderate cerebral ventriculomegaly. The heart was thought to be anatomically normal. Fetal MRI confirmed severe micromelia, cerebral ventriculomegaly with hypoplasia of the corpus callosum, partially fused fornices and abnormalities of the cerebellum. Chromosome analysis and microarray analysis were normal and female. The couple was counselled and decided to continue the pregnancy. Delivery was at 33.0 weeks gestation and induced for severe maternal gestational hypertension. The baby died two hours following delivery of respiratory failure. The autopsy showed facial dysmorphism including a flat philtrum with protruding lower lip, small cleft in alveolar ridge with tooth like structure and tethered tongue. She had severe micromelia, bilateral axillary pterygia, hypoplastic thumbs, camptodactyly with only two flexion creases on most fingers, possible ventricular septal defect and hydrocephalus with partial rhombencephalosynapsis and fused thalami. There are a number of genetically and phenotypically distinct conditions with mandibulofacial dysplasia with acral anomalies. The anomalies in our case are different and may represent a new AFD syndrome.

3076F

Development of a Visit Assessment Tool to Address Birth Defects and Dysmorphology. *O.A. Abdul-Rahman[1], L. Hayes[2], B. Loyola[2], D.A. Stevenson[2], S. Astley[3], C. Chambers[4], K.L. Jones[4], H.E. Hoyme[5], F. Waffern[6], J.C. Carey[2].* 1) Dept Pediatrics, Univ Mississippi, Jackson, MS; 2) Dept Pediatrics, Univ Utah, Salt Lake City, UT; 3) School Public Health, Univ Washington, Seattle, WA; 4) Dept Pediatrics, UCSD, San Diego, CA; 5) Dept Pediatrics, Univ South Dakota, Sioux Falls, SD; 6) Dept Pediatrics, Univ California Irvine, Irvine, CA.

Introduction: The National Children's Study (NCS) is a longitudinal study of various aspects of children's health including genetic, epigenetic, and environmental effects. However, detailed phenotypic information on these children is not comprehensively collected. The current NCS protocol includes study visits at birth, 6 months, and 12 months of age. Several institutions collaborated to develop, evaluate, and implement an assessment tool to be used by NCS field workers to document physical features by digital photography and physical assessment during these visits. Methods: Clinical geneticists from the participating institutions used the Elements of Morphology definitions and illustrations on over 400 phenotypic features as a guide for variations and anomalies needing documentation to design the dysmorphology assessment instrument (DAI). The DAI is composed of two components. The first component is the Photographic Protocol made up of 15 photographic images (8 views of craniofacies and 7 views of the hands and feet), and three 10-second videos for newborns to capture 332 features. The second component is the Physical Assessment list, which was the result of identifying features that are not captured sufficiently on the Photographic Protocol. Results: As of April 2012, the DAI has been administered to 60 participants. The images and videos from the Photographic Protocol for three participants at each age range were randomly selected for case review by three of the authors (OAR, DAS, JCC). Of the 332 features assessed, the rate of agreement on the presence or absence of a feature was 91.6% for the newborn, 92.2% for the 6 moth-old, and 91.9% for the 12 month-old. There was also general agreement regarding photo quality with some differences on the more subjective features and those difficult to define. Conclusions: The DAI is a comprehensive and valid tool able to be implemented by trained staff to capture dysmorphology in the field. This data will provide the complement to genetic and environmental information so that genotype-phenotype correlations can be determined. Further validation and modification of the DAI is needed prior to implementation with the NCS. The DAI may also be a tool to be used by other population-based studies requiring reliable capture of phenotypic information in the field by staff with limited genetics training.

3077W

A NEW CASE OF CRANIOFACIAL MICROSOMIA AUTOSOMAL DOMINANT ASSOCIATED WITH COMPLETE RIGHT SIDE TESTICULAR AND SCROTAL ECTOPIA. *J. Rojas Martínez[1], J.C. Prieto Rivera[1,2].* 1) Pontificia Universidad Javeriana, Instituto de Genetica Humana, Bogota, Colombia; 2) Hospital La Victoria, Genetica, Bogota. Colombia.

Craniofacial microsomia includes a spectrum of malformations primarily involving structures derived from the first and second branchial arches. The severity of the findings range from subtle facial asymmetry with apparently normal ears, to a typical asymmetry with microtia / anotia, atresia of the ear canals, microphthalmia, and respiratory compromise from severe mandibular hypoplasia. Other associated malformations include cleft lip and/or palate, abnormal vertebrae, cardiac, limbs, central nervous system and genitourinary abnormalities, within the latter has never reported the complete scrotal and testicular ectopia unilateral. We present the case of a nine-month-old boy born in Bogotá (Colombia), product of fifth pregnancy maternal, nonconsanguineous parents, born preterm by cesarean section at 35 weeks due to placenta previa, at 2 days of birth, distal functional colostomy was performed for imperforate anus. Normal neurodevelopment. Background in the family of father, aunt and paternal cousin with microtia. Physical examination found facial asymmetry by right hemifacial microsomia with microtia grade III, left small skin tag of 4 mm of sessile base, bilateral epicanthic folds, infantile male genitalia with complete right scrotal and testicular ectopia and nonfunctioning anal sphincter with imperforate anus. Radiographic findings include decreased height of 3 and 4 dorsal vertebrae that resemble platyspondyly and suspected atlanto-odontoid dislocation. CT of the abdomen with absent right kidney, normal left kidney. Skull radiography shows right side mandibular, maxillary and zygomatic arch hypoplasia. Q Banding karyotype was normal: 46, XY. Brainstem evoked response audiometry: abnormal study consistent with right conductive hearing loss. Most cases of craniofacial microsomia correspond to sporadic cases; there are few reports of autosomal dominant inheritance as in this case, the association with complete scrotal and testicular ectopia unilateral has not been reported in the literature.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3078F

A replication study of genome-wide significant Rheumatoid Arthritis susceptibility loci in the Pakistani population. *S.F. Jalil[1,2], A. Bhatti[1], F.Y. Demirci[2], X. Wang[2], I. Ahmed[1], M. Ahmed[3], M.M. Barmada[2], J.M. Malik[4], P. John[1], M.I. Kamboh[2].* 1) Atta-Ur-Rahman School of Applied Biosciences (ASAB), National University of Sciences and Technology (NUST), Islamabad, Pakistan; 2) Department of Human Genetics, GSPH, University of Pittsburgh, Pittsburgh, PA; 3) Rheumatology Unit Military Hospital Rawalpindi, Pakistan; 4) Rehmat Noor Arthritis Research Center Rawalpindi, Pakistan.

Rheumatoid arthritis (RA) is a common, systemic and multifactorial autoimmune disease. The etiology of RA is very complex and is suggested to be the outcome of various environmental, hormonal and genetics factors. Genome-wide association studies (GWAS) have identified several genome-wide significant SNPs associated with RA. In this study, we sought to replicate the association of 32 reported genome-wide significant SNPs residing outside the MHC locus in a unique Pakistani population. We genotyped 32 SNPs in a sample of 366 Pakistanis that comprised related and unrelated cases and controls. All SNPs were genotyped using TaqMan assays and data were analyzed using family case-control (FamCC) ver 1.0 software. Twelve of the 32 SNPs were replicated in this sample with significant p-values ranging from 1.01E-05 to 4.80E-02, the most significant being the *CD28*/rs1980422 SNP. Suggestive associations were seen with additional 6 SNPs (P - 8.97E-02 to 1.93E-01). These observations suggest that different populations share the genetic background and susceptibility loci for RA.

3079W

A novel missense mutation in *PARD3* is associated with class III malocclusion. *T. Nikopensius[1,2], M. Kals[2], T. Annilo[1], T. Jagomägi[3], M. Saag[3], A. Metspalu[1,2].* 1) Institute of Molecular and Cell Biology, University of Tartu, Estonia; 2) Estonian Genome Center, University of Tartu, Estonia; 3) Dept of Stomatology, Faculty of Medicine, University of Tartu, Estonia.

Class III malocclusion is clinically heterogeneous facial phenotype associated with varying combinations of skeletal and dental components. The familial aggregation rates and recurrence risks provide evidence that genetic factors play an important role in its etiology. Moreover, research in multiplex kindreds has indicated that class III malocclusion may represent a monogenic dominant phenotype. We report a novel mutation that underlies class III malocclusion in an Estonian family. Employing whole-exome sequencing in three affected individuals (grandfather-father-son), we identified a heterozygous missense mutation (c.2677C>T [p.Ala893Thr]) in *PARD3* that segregated with this disorder in three generations. The previously undescribed mutation was validated by Sanger sequencing in all patients whereas two unaffected siblings did not carry this mutation. *PARD3* (par-3 partitioning defective 3 homolog) encodes a protein which similarly to other PARD family members is essential for asymmetrical cell division and direct polarized cell growth. PARD3 interacts with several members of the SMAD family of signal transduction and transcriptional modulator proteins that mediate multiple signaling pathways including transforming growth factor beta (TGF-beta) receptor signaling pathway. The TGF-beta family members have a crucial role in craniofacial morphogenesis, particularly in regulation of epithelial differentiation and epithelial-mesenchymal interactions in extracellular matrix remodeling, and recent evidence suggests that their function during craniofacial development is largely mediated through the SMAD signaling system. SMAD proteins are phosphorylated and activated by transmembrane serine-threonine receptor kinases in response to TGF-beta signaling. TGF-beta has been shown to alter cell polarity via a SMAD-dependent signaling mechanism where SMADs can directly interact with proteins involved in cell polarity, such as PAR proteins. This finding underscores the importance of other yet unknown genes and developmental factors in the etiology of dental anomalies.

3080F

Genotype/Phenotype correlation in Smith-Magenis Syndrome with abnormal 17p deletions. *T. Vilboux[1], A.C.M Smith[2], S. Chandrasekharappa[3], C. Ciccone[1], J. Blancato[4], W.J. Introne[2], W.A. Gahl[1,2], M. Huizing[1].* 1) Medical Genetics Branch, National Human Genome Research Institute, NIH, Bethesda, MD; 2) Office of the Clinical Director, National Human Genome Research Institute, NIH, Bethesda, MD; 3) Cancer Genetics Branch, National Human Genome Research Institute, NIH, Bethesda, MD; 4) Department of Oncology, Georgetown University Medical Center, Washington, DC.

Smith-Magenis syndrome (SMS) is a complex developmental disorder characterized by an interstitial deletion of chromosome 17p11.2. This syndrome includes variable symptoms such as cognitive impairment, craniofacial dysmorphia, height-growth delay, infantile hypotonia, brachydactyly, attention deficit, decreased sensitivity to pain, self-injury, maladaptive behaviors, speech delay, hearing loss, and sleep disturbance. Most SMS patients have the same approximately 3.7Mb interstitial 17p11.2 genomic deletion. Haploinsufficiency of the *RAI1* gene is likely responsible for most of the SMS features, but haploinsufficiency of other genes within the deleted region may also contribute to the phenotype of SMS patients with 17p11.2 deletions. To further investigate the molecular basis of the SMS phenotype we studied phenotype/genotype correlation of 10 SMS patients with abnormal 17p deletions. We identified exact deletion sizes by multiple genomic techniques, including FISH (fluorescent in situ hybridization), MLPA (Multiplex Ligation-dependent Probe Amplification) and quantitative PCR analyses. Furthermore, each patient's DNA was analyzed on a SNP-microarray (Illumina chip) to define the exact 17p breakpoints. With these data we refined the minimal Smith-Magenis syndrome critical region to an approximately 500-kb interval on 17p11.2 that includes RAI1 and 6 other genes. A deeper analysis of the genes in or outside this minimal critical region provided us with candidate genes for some of the unusual phenotypes of certain patients, including, obesity, severe speech delay, hearing-loss or immune defects. Analysis of genotype/phenotype correlations in patients with unusual deletions in complex syndromes like SMS may provide candidate genes for specific phenotypes within the syndrome. These findings may not only benefit the studied syndrome, but may also assist in molecular analysis of more common phenotypes (e.g. for SMS: autism, sleep abnormalities, obesity, speech delay, maladaptive behaviors).

3081W

MiR-133a and MiR-422a in Human Circulating Monocytes are Potential MicroRNA Biomarkers Underlying Postmenopausal Osteoporosis. *Y. Wang[1], L. Li[1], B.T. Moore[1], X.H. Peng[1], X. Fang[2], J.M. Lappe[1], R.R. Recker[1], P. Xiao[1].* 1) Osteoporosis Research Center, School of Medicine, Creighton University, Omaha, NE; 2) Biostatistical Core, Office of Research and Compliance, Creighton University, Omaha, NE.

MicroRNAs (miRNAs) are a class of short non-coding RNAs (~ 22 nt) that involved in posttranscriptional gene regulation, which regulate gene expression by targeting mRNAs. Circulating monocytes play important roles in osteoclastogenesis by acting as osteoclast precursors and secreting osteoclastogenic factors, such as IL-1, IL-6 and TNF-α. This study aimed to find significant miRNA biomarkers in human circulating monocytes underlying postmenopausal osteoporosis. We used ABI TaqMan® miRNA array followed by qRT-PCR validation in human circulating monocytes to identify miRNA biomarkers between 10 high BMD and 10 low BMD postmenopausal Caucasian women. MiR-133a and miR-422a were up-regulated (P = 0.007, P =0.065) in the low compared with the high BMD groups in the array, which were also validated by qRT-PCR (P = 0.044, P =0.029). We performed bioinformatic target gene analysis and found three potential target genes, CXCL11, CXCR3 and SLC39A1 of miR-133a and five potential target genes (CBL, CD226, IGF-1, PAG1, and TOB2) of miR-422a, which are able to inhibit osteoclastogenesis. In addition, we performed Pearson correlation analyses of the expression levels of miR-133a and miR-422a with the expression levels the potential target genes in the 20 postmenopausal women. We did find negative correlations though not significant. Many studies demonstrated that miR-133a is important in the development of muscle, such as skeletal and cardiovascular muscle. In bone, particularly, miR-133a has been found to regulate osteoblastogenesis by targeting and regulating Runx2 expression. A recent study also demonstrated that miR-133a was up-regulated in osteoblast-like periodontal ligament stem cells treated with ibandronate, a nitrogen-containing bisphosphonate that inhibits bone resorption and is widely used to treat osteoporosis. Only one study demonstrated that miR-422a was down-regulated in osteoblast-like cell line (MG-63) treated with peptide-15. However, our study for the first time suggests that in vivo miR-133a and miR-422a in osteoclast precursors, circulating monocytes, are potential miRNA biomarkers underlying postmenopausal osteoporosis.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3082F

Genome-wide Mosaic Paternal Uniparental Isodisomy. *J. Kalish[1,2], A. Wilkens[1], S. Mulchandani[3], E. Zackai[1], N. Spinner[3], M. Bartolomei[2], L. Conlin[3], M. Deardorff[1].* 1) Division of Human Genetics, The Children's Hospital of Philadelphia, Philadelphia, PA; 2) Department of Cell & Developmental Biology, Perlman School of Medicine, The University of Pennsylvania, Philadelphia, PA; 3) Department of Pathology and Laboratory Medicine, The Children's Hospital of Philadelphia, Philadelphia, PA.

The growing number of identified imprinted genes in humans underscores the need for biparental inheritance of the genome; however, 8 cases of live-born genome-wide mosaic paternal uniparental isodisomy (GWpUPiD) have been reported. We describe four additional cases diagnosed by genome-wide SNP array analysis. All four patients presented with persistent neonatal hypoglycemia caused by hyperinsulinism (HI). Three of the four presented with hemihyperplasia as well as placentomegaly. Given the increased risk for hepatoblastoma and Wilms tumor in patients with paternal uniparental disomy (UPD) 11, we initiated screening with alphafetoprotein (AFP) levels and abdominal ultrasounds. Patient 1 developed large cystic liver hamartomas requiring resection. She is currently 26 months old with near normal cognitive development. Patient 2 had severe HI requiring a 95% pancreatectomy. She had persistently elevated AFPs and serial imaging followed by biopsy revealed bladder polyps, adrenal cortical hyperplasia, and hepatic hemangiomas. At 13 months, she has some mild motor delays. Patient 3 was born at 27 weeks gestation with a unilateral multiloculated cystic suprarenal mass and had severe HI. Patient 4 was born at 24 weeks gestation, had severe HI, and had an adrenal mass concerning for neuroblastoma. Patients 3 and 4 died of complications of HI and prematurity. An autopsy of patient 4 revealed adrenal cortical and pancreatic hyperplasia. The clinical features in these patients partially derive from alteration of the 5 known paternal UPD disease loci in humans. These include UPD 6q24 (transient neonatal diabetes), 11p15 (Beckwith-Wiedemann syndrome; BWS), 14q32 (hypotonia and bell-shaped thorax), 15q11.2 (Angelman syndrome), and 20q13 (pseudohypoparathyroidism). The predominant phenotype in all GWpUPiD patients reported to date is BWS. All reported GWpUPiD patients developed adrenal, renal, and/or hepatic tumors. This 100% incidence of tumorigenesis suggests that the UPD11 phenotype is augmented in GWpUPiD. This may be due to additional less well-characterized imprinted regions involved in cell growth. Alternatively, these patients have isodisomy that may uncover recessive loci, which contribute to the increased incidence of tumors. Given the rapidity of tumor growth and the prolonged age of onset seen in GWpUPiD patients, we recommend vigilant ultrasound screening of these patients in the first years of life and consideration of tumor surveillance into adolescence.

3083W

The Catalase: influence of C-262 T polymorphism on Keratoconus. *S. Torabi Dalivandan[1], Z. Salehi[1], S. Saboohi[1], M.J. Mohammadi Fatideh[2, 3], H. Yourdkhani[2].* 1) Biology Department, Science Faculty, University of Guilan, Rasht, Iran; 2) Ophthalmology Department, Amiralmomenin Hospital, Rasht, iran; 3) Guilan University of Medical Sciences, Rasht, iran.

Keratoconus (KC) is an eye disorder in which the cornea assumes a conical shape because of thinning and protrusion, with the possible involvement of both genetic and environmental factors. It has been reported that oxidative damage contributed to pathogenesis of this disease. To inactivate stress oxidative there are several antioxidant enzymes, one of them is catalase which contributes H2O2 to H2O and O2. Given the role of oxidative damage in causing this condition, the aim of this study was to investigate the association of Catalase C-262 T polymorphism with KC. Blood samples were obtained from 30 Iranian patients (mean age of 25.65±2.115 years) with KC and 30 healthy controls (mean age of 23.87±1.125 years). To determine genotype frequencies genomic DNA was extracted from the blood samples and subjected to the allele- specific PCR (AS-PCR) technique. Statistical analyses were performed using MedCalc v.12.1.4.0 software. Results has not emerged significance differences from the comparison of either genotype (P > 0/05). We conclude that the catalase gene C-262 T polymorphism can not be a risk factor in KC.

3084F

Metabolic abnormalities in Williams-Beuren syndrome patients and candidate genes. *M.G. Palacios[1], R. Flores[1], V. Campuzano[1,2], M. del Campo[1,3], M. Segura-Puimedón[1], L.A. Perez-Jurado[1].* 1) Genetics Unit, DCEX, Universitat Pompeu Fabra & Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Barcelona, Spain; 2) Neurosciencies Program, Mar Institute of Medical Research Foundation (IMIM), Barcelona, Spain; 3) Program in Molecular Medicine and Genetics, Hospital Universitari Vall d'Hebron, Barcelona, Spain.

Background: Williams-Beuren syndrome (WBS, MIM 194050) is a neurodevelopmental disorder with multisystemic manifestations caused by a 1.55–1.83Mb deletion at 7q11.23 including 26–28 genes. Reported metabolic abnormalities in WBS include transient hypercalcemia of infancy, subclinical hypothyroidism in ≈30% of children, and impaired glucose tolerance in ≈75% of adult individuals. We have studied several biochemical and hormonal parameters in a cohort of 100 WBS individuals, half of them participating in a pilot clinical trial to evaluate the efficacy and safety of angiotensin II receptor blockers in reducing oxidative stress and blood pressure in WBS. **Results**: 1) Triglyceride plasma levels were significantly decreased in WBS, with a median at the 15th percentile and 24.5% cases with hypotriglyceridemia, while cholesterol levels were above the normal range in 1/3 of cases. Haploinsufficiency at *MLXIPL*, documented by RTPCR expression quantification in patient cell lines and liver samples of Wbs mouse models, is likely related to the lipid abnormalities. 2) Bilirubin levels were significantly increased, with 24.1% of cases showing indirect hyperbilirubinemia, preferentially females (p=0.026) with homozygous hypomorphic *UGT1A1*-promoter alleles. Decreased triglyceride and increased bilirubin levels were also observed in Wbs mouse models. Interestingly, expression of *Ugt1A1* was significantly increased in murine Wbs livers. 3) Additional abnormalities included elevated total protein (30.9%), albumin (26.7%) and transferrin (66.1%) plasma levels. Subclinical hypothyroidism (TSH elevation with normal T3/T4) was found in 44.6% but did not correlate with the biochemical abnormalities. 4) Following therapy with angiotensin II receptor blockers for 12 months (n=16), a significant reduction in protein plasma levels (p=0.015), as well as uric acid (p=0.018) and creatinine levels (p=0.001) was observed. **Conclusions**: Our data indicate that several unreported biochemical alterations are relatively common in WBS and should be studied in greater detail in order to better prevent potential long-term complications. These biochemical alterations can be related to haploinsufficiency for specific genes at 7q11.23. Angiotensin II receptor blockers might contribute to improve some biochemical abnormalities through controlling oxidative stress and blood pressure.

3085W

Bilateral radioulnar synostosis and vertebral anomalies in a child with a 16p13.3 interstitial deletion. *T. Slavin[1, 2], A. Tam[3], K. Lee[3], S. Lee[2], W. Burkhalter[1,4,5].* 1) Department of Pediatrics, University of Hawaii, John A. Burns School of Medicine, Honolulu, Hawaii; 2) Kapiolani Medical Specialists, Hawaii Community Genetics, Honolulu, Hawaii; 3) University of Hawaii, John A. Burns School of Medicine, Honolulu, Hawaii; 4) Kapiolani Medical Specialists, Kapiolani Orthopedic Associates, Honolulu, Hawaii; 5) Department of Surgery, University of Hawaii, John A. Burns School of Medicine, Honolulu, Hawaii.

We describe an 8-year-old boy with developmental delay, bilateral radial ulnar synostosis, Klippel-Feil anomaly and multiple other vertebral deformities who was found to have a de novo, previously unreported, deletion of 114.5kb of 16p13.3. On physical exam he was nondysmorphic, but had myopathic facial features with a small jaw. His mouth displayed dental crowding with ankyloglossia, a large thick tongue and a bifid uvula. He had prominent ear crus bilaterally. He had wide nipple spacing. He displayed limited range of motion in his neck, mild scoliosis, complete loss of forearm rotation on the right and limited pronation and supination on the left. He had a right transverse single palmar crease. He displayed hypoplasia of thenar and hypothenar musculature bilaterally. He had some prominence of the medial aspect to the top of his knees. His gait was mildly uncoordinated and he was not able to run. The remainder of his physical exam was unremarkable. Fanconi anemia testing by chromosome breakage was normal. Renal ultrasound was normal. Full skeletal survey and echocardiogram were recommended but not done at the time of this report. There are only 5 known genes contained in the deletion and one miRNA. Interestingly, none of the genes have any known involvement with skeletal regulation to explain the striking boney defects seen in the patient presented herein. ABCA3 and DCI are involved in surfactant and beta-oxidation of unsaturated fatty acids, respectively. MIR940 is a novel miRNA with little known function that is reported to be expressed in the cervix and renal cell carcinoma. E4F1, DNASEIL2, RNPS1 are expressed in bone marrow and hematologic cell lines and have broad functions and therefore may have a role in bone regulation. Nonetheless, even though we cannot find a specific genotype-phenotype correlation to explain the boney findings in this patient, we believe this deletion is causative, given that it was de novo and that this patient cannot be easily explained as having any other specific recognizable pattern of human malformation. As this is the first case report involving this particular gene deletion, additional case reports and functional studies regarding the genes and regulatory elements involved will be valuable in better understanding the regions role in skeletal regulation.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3086F

ASTN2 deletions in autism spectrum disorder and related neuropsychiatric phenotypes. *M.T. Carter[1], A.C. Lionel[2], A.K. Vaags[2], B.A. Fernandez[3], W. Roberts[4], P. Szatmari[5], C.R. Marshall[2], S.W. Scherer[2].* 1) Division of Clinical and Metabolic Genetics, Hospital for Sick Children, Toronto, ON, Canada; 2) The Centre for Applied Genomics, Hospital for Sick Children, Toronto, ON, Canada; 3) Disciplines of Genetics and Medicine, Memorial University of Newfoundland, St. John's Newfoundland, Canada; 4) Autism Research Unit, Hospital for Sick Children, Hospital for Sick Children, Toronto, ON, Canada; 5) Department of Psychiatry and Behavioural Neurosciences, McMaster University, Hamilton, ON, Canada.

Astrotactin-2 (ASTN2) encodes a cell adhesion molecule widely expressed in the developing brain. The protein, in complex with ASTN1, facilitates glial-guided neuronal migration. Previous studies have implicated disruption of ASTN2 in attention deficit-hyperactivity disorder (ADHD), schizophrenia, and autism spectrum disorder (ASD). We present two unrelated males referred for clinical genetics assessment with autism spectrum disorder and prominent behavioral issues, each with a deletion involving ASTN2. Patient 1 is a 16 year old male with high-functioning autism, ADHD, language-based learning disability, obsessive compulsive disorder (OCD), oppositional defiant disorder (ODD), scoliosis, micrognathia, pectus excavatum, and pre and postnatal symmetric growth retardation. There is a positive family history of Asperger syndrome. Array comparative genomic hybridization (aCGH) showed a 194 kb deletion at 9q33.1, involving three RefSeq genes including ASTN2. The ASTN2 deletion was confirmed by FISH. It was not present in the father, and the mother was unavailable for testing. Patient 2 is a 12 year old male with Asperger syndrome, above average IQ, and a history of escalating behavioral outbursts (swearing, violent threats, sexual inappropriateness). He has dysmorphic features (short and upslanted palpebral fissures, long face, prominent nose, bilateral sandal gap) and both height and head circumference are at the 98th percentile. aCGH showed a 291 kb deletion at 9q33.1, involving exonic sequences of the ASTN2 gene. The deletion was not present in the mother, and the father was unavailable for testing. We found three additional males with deletions within the ASTN2 gene in our autism cohort, as well as one family with an intragenic ASTN2 duplication which segregates with Asperger syndrome diagnoses in a mother and her two sons (and not in the third unaffected son). No copy number variants at this locus are found in 2357 population controls. We present the clinical and molecular features of these six probands, and suggest that ASTN2 disruption is highly penetrant for autism and related neuropsychiatric disorders.

3087W

Atypical deletions of the Williams-Beuren syndrome region implicate genes involved in mild facial phenotype, epilepsy and autistic traits. *G. Merla[1], L. Micale[1], C. Fusco[1], B. Augello[1], P. Alfieri[2], M.C. Digilio[3], D. Menghini[2], S. Vicari[3].* 1) Medical Genetics Unit, IRCCS Casa Sollievo Della Sofferenza, San Giovanni Rotondo, Italy; 2) Child NeuroPsychiatry Unit, Neuroscience Department, IRCCS Children Hospital Bambino Gesù, Rome, Italy; 3) Medical Genetics, IRCCS Children Hospital Bambino Gesù, Rome, Italy.

Interstitial deletions of multiple contiguous genes at chromosome band 7q11.23 are responsible for Williams-Beuren Syndrome (WBS; OMIM # 194050), a genomic disorders with multi-system involvement and variable expressivity, caused by a hemizygous deletion of 1.5 Mb on chromosome 7q11.23 spanning 28 genes. A few patients with larger and smaller deletions have been reported. They show clinical features that vary between isolated SVAS to the full spectrum of WBS phenotype, associated with epilepsy or autism spectrum behavior, infrequently diagnosed in WBS patients with common deletions. In this report, we describe clinical, cognitive, and genetic features of four patients with atypical 7q11.23 deletion. Two of them carry an approximately 3.5 Mb larger deletion towards the telomere that includes also two genes implicated in neuropsychiatric disorders as Huntingtin-interacting protein 1 (HIP1) and tyrosine 3-monooxygenase/tryptophan 5-monooxigenase activation protein gamma (YWHAG). The others two individuals carry two shorter deletions of approximately 1.2 Mb at centromeric side that does not include the BAZ1B and FDZ9 genes. The description of these additional cases further helped us to refine the genotype-phenotype correlations in WBS. Our data suggest, indeed, that HIP1 and YWHAG cause the severe neurological dysfunctions, including epilepsy and autistic traits, seen in the two WBS patients with larger deletion. Mild facial features and moderate neurocognitive profile of patients with the smaller deletions may be related in part to the preservation of BAZ1B and FZD9 gene. This report highlights the importance to identify and characterize further patients with atypical deletions of 7q11.23 region to shed light on the pathogenic role of genes within and outside the WBS region.

3088F

Morphogenesis of the face in Wolf-Hirschhorn syndrome: three elucidative patients. *J. Carey[1], A. Calhoun[1], P. Hammond[2], A. Lortz[3], S. South[1].* 1) Dept Ped/Div Med Gen, Univ Utah Hlth Sci Ctr, Salt Lake City, UT; 2) Molecular Medicine Unit, UCL Institute of Child Health, London, UK; 3) The 4p- Support Group, Sunbury, OH.

Wolf-Hirschhorn syndrome (WHS) is a contiguous gene deletion syndrome caused by deletions of the distal short arm of chromosome 4. Patients with WHS have growth delay, developmental delay, and the distinctive “Greek warrior helmet” craniofacial appearance. Key features include hypertelorism, prominent eyes, arched eyebrows, prominent nasal bridge, downturned corners of the mouth, micrognathia, and a short philtrum. Historically, the typical facial gestalt has been felt to be due to deletion of the gene WHSC1, found within both of the two overlapping putative critical regions for the syndrome, WHSCR1 and WHSCR2. However, recent published reports and assessment of previously unpublished photos reveal 3 informative individuals, whose findings suggest that the facial gestalt of WHS is due to deletion of multiple genes on distal 4p. Summaries of the 3 cases follow: 1: Girl with growth failure, developmental delay, seizures, and facial features including prominent forehead, mild telecanthus, and long palpebral fissures. She had a 1.78MB terminal deletion of 4p, which involves LETM1, but does not include WHSC1 or WHSC2. Dense surface modeling (DSM) fine-grained computerized facial analysis was consistent with a partial WHS phenotype and was right on borderline for WHS face. 2: Boy with growth delay, poor feeding, ASD, and developmental delay. He showed a subtle WHS facial gestalt. He had a 4.3 MB interstitial deletion including LETM1, WHSC1 and WHSC2. DSM showed some difference from control faces, and on DSM was right at the borderline for WHS. 3: Boy with poor growth, feeding difficulties, developmental delays, hypertelorism, mild syndactyly, and prominent eyes. He had a 432kb deletion proximal to the both the putative WHS critical regions which does not include LETM1, WHSC1, or WHSC2. DSM classified his face as more WHS-like than control-like and not borderline like Cases 1 & 2. Although WHSC1 has long been felt to be the key to the gestalt of WHS, these 3 patients point to a more complex picture. By DSM, Patients 1 and 3 have features more consistent with WHS than Patient 2, despite the fact that patient 2 is the only one of the three who carries a deletion of the two overlapping WHS critical regions. Analysis of these patients argues that genes outside of the purported critical regions impact the morphogenesis of the face in WHS.

3089W

Growth hormone receptor gene polymorphism and Prader-Willi syndrome. *M. G. Butler, J. Roberts, J. Hayes, X. Tan, A. Manzardo.* Psychiatry & Behavioral Sci, Univ Kansas Med Ctr, Kansas City, KS.

Prader-Willi syndrome (PWS) is characterized by infantile hypotonia; failure to thrive; a poor suck and feeding difficulties; hypogonadism with genital hypoplasia; growth hormone deficiency with short stature, small hands and feet; hyperphagia leading to early childhood obesity; mental deficiency and behavioral problems. PWS is due to loss of paternally expressed genes from the 15q11-q13 region usually from a de novo paternally derived deletion. When treated with growth hormone (GH), PWS individuals respond favorably in stature and body composition (decreased fat, increased muscle mass). The exon-3 deletion polymorphism (d3) in the growth hormone receptor (*GHR*) gene occurs in about 50% of Caucasians in the general population and has been associated with an increased growth response to GH therapy in non-PWS patients. The aim of our study was to assess the association of the GHR alleles with height, weight, head circumference and body mass index (BMI) at baseline and prior to GH treatment, as well as, gender and PWS genetic subtypes. We examined 69 genetically confirmed individuals with PWS (30 males, 39 females; average age ± SD = 20.1 ± 12.8y). Thirty-nine individuals had the 15q11-q13 deletion and the remaining subjects had maternal disomy 15 or an imprinting defect. The vast majority (i.e., 95%) of our subjects were Caucasians. The distribution of alleles (fl/fl, n=36 or 52%; fl/d3, n=25 or 36%; d3/d3, n=8 or 12%) in our PWS subjects was similar to reported data in Caucasian control subjects. There were no gender or PWS genetic subtype differences identified in the distribution of *GHR* alleles. We found a negative correlation ($p<0.05$) with age for height standard deviation scores and a positive correlation ($p<0.05$) with age for weight and BMI. Adjusting for effects of age and gender, we found that individuals with PWS carrying the d3 allele showed a significant increase in BMI compared with those having the full length allele ($F=3.9$, $p<0.02$). This model explains ~35% of the shared variance in BMI with ~5% of the unique variance attributed to *GHR* alleles. No differences in height, weight or head circumference standard deviation scores were found. In summary, the d3 allele was associated with significantly increased BMI in our cohort of PWS subjects prior to GH treatment but not for height or weight. The presence of the d3 allele and its impact on BMI and the care of individuals with PWS while on GH therapy should be addressed.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3090F

Molecular characterization and genotype/phenotype correlation of ten patients with structural rearrangements in the *NIPBL* gene. *YW. Cheng, C. Tan, K. Arndt, S. Das, D. del Gaudio.* Human Genetics, University of Chicago, Chicago, IL.

Cornelia de Lange syndrome (CdLS) is a genetically heterogeneous disorder characterized by severe growth retardation, intellectual disability, upper-limb abnormalities, hirsutism and characteristic facial features. Heterozygous mutations in the *NIPBL, SMC1A* and *SMC3* cohesin-associated genes account for approximately 60–65% of patients with a diagnosis of CdLS; the majority of them have been found in *NIPBL*. Recent studies have shown that intragenic deletions and duplications in *NIPBL* are present in approximately 2–5% of patient with CdLS. To explore the occurrence of intragenic deletions and duplications in the *NIPBL* gene we analyzed a cohort of 430 patients referred to our laboratory for CdLS clinical testing and that had previously tested negative by standard *NIPBL* sequence analysis. Deletion/duplication analysis was performed by MLPA and custom exon-targeted oligonucleotide array-CGH. *NIPBL* copy number changes were detected in 10 out of 430 (2.3%) patients, including 8 intragenic single- or multi-exon alterations (7 deletions, one duplication), a full *NIPBL* gene deletion and a paternally inherited large duplication involving *NIPBL* exons 32 to 47 and including downstream genes *FLJ13231*, *NUP155* and *WDR70*. Preliminary analysis of the clinical phenotype of these patients suggest that the majority have facial features consistent with CdLS and include synophyrs (6/9), long thick eyelashes (6/9), upturned/anteverted nares (6/9), thin upper lip (6/9) and a long smooth philtrum (7/9). The patient with the large duplication involving partial *NIPBL* and downstream genes was also determined to have a *PTPN11* sequence change and a phenotype consistent with Noonan syndrome. Additional clinical information on all ten patients is being gathered and will be presented. Our experience contribute to the growing amount of data on*NIPBL*-related copy number variations in patients with CdLS, and support the clinical usage of *NIPBL* deletion/duplication analysis in patients with negative *NIPBL* results by sequencing.

3091W

***TSC2-PKD1* Contiguous Deletion Syndrome with Aortic Stenosis and Severe Myopia.** *K. Enomoto[1,2], Y. Sugawara[2], M. Hotate[2], Y. Motoyoshi[2], Y. Hatai[3], S. Mizutani[2], K. Kurosawa[4].* 1) Division of Pediatrics, JA Toride Medical Center, Toride, Ibaraki, Japan; 2) Department of Pediatrics and Developmental Biology, Tokyo Medical and Dental University Graduate School, Tokyo, Japan; 3) Division of Pediatrics, Tokyobay Urayasulchikawa Medical Center, Urayasu, Chiba, Japan; 4) Division of Medical Genetics, Kanagawa Children's Medical Center, Yokohama, Japan.

Tuberous sclerosis complex (TSC) is an autosomal dominant disorder characterized by (1) brain lesions, including subependymal nodules, cortical tubers, and subependymal giant cell astrocytomas, (2) skin lesions, including hypomelanotic macules, facial angiofibromas, and shagreen patches, and (3) renal lesions. Approximately 70% of probands has *TSC2* identifiable causative mutation mostly detected by sequence analysis, rarely exonic-deletion/duplication analysis including multiplex ligation-dependent probe amplification (MLPA) or micro-array. Although the whole-gene deletion of *TSC2* is very rare, it usually involves contiguous gene of *PKD1*, characterized to cause, in case of mutation, autosomal dominant polycystic kidney disease (ADPKD), directly downstream to the *TSC2* separated by only a 60-bp region. We herein present a young case of 16p13.3 microdeletion extending to 700kb involving *TSC2* and *PKD1*identified by micro-array CGH examination. He presents some clinical and dysmorphic features including congenital heart defect, aortic stenosis (AS) and patent ductus arteriosus (PDA), pectus excavatum, severe myopia, and facial dysmorphism which are inconsistent with typical TS, ADPKD or simple addition of them. Therefore, we suggest that the haploinsufficiency of the genes involving the deleted region excepting *TSC2* and *PKD1* contributes to the above mentioned some characteristic features.

3092F

Jacobsen syndrome: Cognitive-behavioral profiles, developmental trajectories, and IQ related to deletion size. *G. Fisch.* NYU College of Dentistry 250 Park Ave. S., 6th fl. New York, NY.

Jacobsen syndrome [JBS] is caused by a subtelomeric deletion at chromosome 11q25. Prevalence is estimated at 1/100,000 births. In addition to its clinical features, JBS causes intellectual disability [ID]. ID severity in patients with JBS is variable. To our knowledge, this is the first comprehensive cognitive-behavioral evaluation of children with JBS. Our purpose is to: (1) ascertain cognitive-behavioral profiles and develop trajectories in children with JBS; and (2) examine IQ related to deletion size. Method Subjects: Ten children with JBS were recruited from various sites in the US and Europe. All were English-speaking. Five were male, 5 were female. Ages ranged from 5–20 years (Mean=11.25 years). Materials: Children were administered a neuropsychological battery that included cognitive assessment with the Stanford-Binet (4th Edition) [SBFE] and adaptive behavior [DQ] with the Vineland Adaptive Behavior Scale [VABS]. Procedure: At initial testing [T1] children were administered the SBFE, and parents interviewed with the VABS. Eight children were available for retesting 2 years later [T2]. Results IQ at T1 ranged from 36–80 [Mean IQ=59.2 ± 16.2]. DQ ranged from 31–68 [Mean DQ=52.1 ± 13.1]. Four children exhibited behaviors consistent with a diagnosis of mild autism. Cognitive-behavioral profiles were relatively flat, with relative strength in quantitative reasoning and significant weakness in short-term memory. Cognitive-behavioral profiles at T2 remained relatively stable. IQ scores ranged from 43–79 [Mean IQ=62 ± 14.2], DQ scores ranged from 40–68 [Mean DQ=54.6 ± 11.8]. The correlation between deletion size and IQ was (r=-0.20; p=ns) and DQ (r=-0.18; p=ns). Discussion The cognitive-behavioral profile is significantly different from children with other microdeletion disorders, e.g. Williams syndrome [WS] or Wolf-Hirschhorn syndrome [WHS], who show relative strengths in verbal and quantitative reasoning. Children with JBS show strengths in quantitative reasoning only. Children with WS or WHS show relative weaknesses in both abstract / visual reasoning and short-term memory. Those with JBS show relative weakness in short-term memory only. Previously, Colden et al [2009] examined 14 individuals with JBS and found a strong negative correlation [r=-0.82] between IQ and deletion size. We find a significantly lower correlation not different from zero. Results will be discussed in the context of sampling and sample size.

3093W

Concurrent Deletion of BMP4 and OTX2 Genes: Clinical Evidence of Synergistic Effect of the Two Master Genes in Ophthalmogenesis. *R. Kosaki[1], T. Takenouchi[2], C. Torii[3], S. Nishina[4], K. Kosaki[3].* 1) Div Med Genet, Natl Ctr Child Hlth & Dev, Tokyo, Japan; 2) Dept Pediatr, Keio Univ, Tokyo, Japan; 3) Center Med Genet, Keio Univ, Tokyo, Japan; 4) Div Ophthalmol, Natl Ctr Child Hlth & Dev, Tokyo, Japan.

BMP4 and OTX2 are master genes in ophthalmogenesis. Mutations of BMP4 and OTX2 often lead to eye defects, including anophthalmia-microphthalmia. A significant degree of variable expressivity has been reported in heterozygous individuals with BMP4 or OTX2 mutation. Interestingly, both BMP4 and OTX2 reside on 14q22, being only 2.8 Mb apart. Previous studies reported that among three patients with 14q22 deletion involving BMP4 and OTX2, all had severe eye defects. The minimal degree of variable expressivity among these individuals who were doubly heterozygous for BMP4 and OTX2 could be attributed to the combinatorial relationship of the two genes observed in animal models. We herein report a patient with a concurrent deletion of BMP4 and OTX2 who exhibited bilateral microphthalmia, more specifically, anterior segment dysgenesis with microcornea. The presence of severe microphthalmia in the propositus supports the notion of functional complementarity between BMP4 and OTX2 in humans. Evolutionarily conserved physical linkage of BMP4 and OTX2 loci may suggest an advantage of the proximal alignment of the two genes. Another striking feature in the propositus was the progressive white matter loss observed by serial neuroimaging. A review of nine previously reported patients with 14q22 microdeletion revealed decreased white matter volume in half of the patients. It remains to be elucidated whether the white matter lesion is age-dependent and progressive. In conclusion, anterior segment defects of the eyes, especially when accompanied by decreased white matter volume on neuroimaging, should raise the clinical suspicion of 14q22 microdeletion.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3094F

Interstitial deletion of 3.8 Mb in chromosome 6q25.2-6q25.3 in a patient with Coffin-Siris syndrome. *S. Mizuno[1], E. Nishi[1], Y. Muramatsu[1], M. Tominaga[2], K. Kurosawa[2].* 1) Dept Pediatrics, Central Hosp, Aichi Human Service Ctr, Kasugai, Aichi, Japan; 2) Dept Clinical Genetics, Kanagawa Childrens Medical Center, Kanagawa, Japan.

We describe the case of a 5-year-old girl with 3.8Mb interstitial deletion of chromosome 6q25.2-5q25.3 identified by oligonucleotid array-CGH. The patient was born to healthy parents at the gestational age of 39 weeks with average height and weight. She presented with psychomotor developmental delay, absent speech, progressive postnatal growth restriction, hypertrichosis,hirsutism of the back, and a distinctive facial appearance with thick lips. The most noteworthy finding was the nails of her fingers and toes were thin, malformed and hypoplastic. These feature of the nails are diagnostic characteristics of Coffin-Siris syndrome, which is also known as “Fifth digit syndrome” Tsurusaki et al. reported germline mutation in one of six SWI/SNF-A chromatin-Remodeling complex subunit genes, *SMARCB1*, *SMARCA4*, *SMARCA2*, *SMARCE1*, *ARID1A* and *ARID1B*, causes Coffin-Siris syndrome. The chromosomal deletion in our patient involved *ARID1B*, the haploinsufficiency of which resulted in the development of Coffin-Siris Syndrome. Although the chromosomal deletion in this case involoved 12 genes other than ARID1B, psychomotor development and physical manifestation of the patient were almost identical to those in other reported cases with Coffin-Siris syndrome caused by mutation in ARID1B. Further studies on interstitial deletions in this region are required to describe the precise phenotype of Coffin-Siris Syndrome caused by chromosomal deletion.

3095W

Detailed Neurodevelopmental Phenotype of a Child with a 22q11.2 Distal Deletion. *S. Mosca[1], L.M. Langevin[2], A.C. Lionel[3,4], C.R. Marshall[3,4], B. Argiropoulos[1,2], A.M. Innes[1,2], S.W. Scherer[3,4], D. Dewey[1,2], J.S. Parboosingh[1,2], F.P. Bernier[1,2].* 1) Medical Genetics, University of Calgary, Calgary, Alberta, Canada; 2) Alberta Children's Hospital Research Institute, Calgary, Alberta, Canada; 3) The Centre for Applied Genomics and Program in Genetics and Genome Biology, Hospital for Sick Children, Toronto, Ontario, Canada; 4) Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada.

Microdeletions mediated by non-allelic homologous recombination (NAHR) between flanking low-copy repeats (LCRs) at 22q11.2 are highly variable, as is the phenotypic presentation in individuals harboring deletions within this region. Eight major LCRs are located in 22q11, and recombination between LCR_{22-3} and LCR_{22-4} results in a common ~3Mb deletion responsible for a range of phenotypes now recognized as 22q11.2 deletion syndrome. Recently, distal deletions resulting from NAHR between LCR_{22-4} and LCR_{22-5} have been associated with a novel disorder which appears to be distinct from the more proximal 22q11.2 deletion syndrome. Common features include developmental delay, microcephaly, pre- and post-natal growth delay, cardiovascular malformations, prematurity, and mild dysmorphic features. While an increased incidence of neurodevelopmental disorders, such as Autism Spectrum Disorder (ASD) and Attention Deficit Hyperactivity Disorder (ADHD), have been described in children with typical 22q11.2 deletions, details on the neurodevelopmental phenotypes seen in children with distal 22q11.2 deletions are lacking. The analysis of copy number variants (CNVs) using Illumina 2.5M genotyping microarrays in a cohort of 91 children recruited as part of a study on the genetics of Developmental Coordination Disorder (DCD) identified an ~1.3Mb distal 22q11.2 deletion flanked by LCR_{22-4} and LCR_{22-5} in an individual with severe DCD and ADHD. This deletion was validated by quantitative real-time PCR and was not inherited from the mother. The participant was born at 37 weeks and was small for gestational age. He went on to have failure to thrive and speech delay. We present detailed neurodevelopmental test results for this participant who had an IQ score at the lower range of normal (IQ of 80), and met diagnostic criteria for both DCD (1st centile on the Movement Assessment Battery for Children) as well as ADHD (Diagnostic Interview for Children and Adolescents). This case demonstrates the variability seen in the clinical presentation of individuals with 22q11.2 distal deletions. It is increasingly recognized that rare CNVs are an important cause of neurodevelopmental disorders with both reports of recurrent as well as rare variants in patients with ASD as well ADHD. We add distal 22q11.2 deletions to the list of CNVs that may contribute to pediatric neurodevelopmental phenotypes although further validation in replication cohorts will be required to confirm this association.

3096F

A *de novo* 163 kb interstitial 1q44 microdeletion in a boy with thin corpus callosum, psychomotor delay and seizures. *K. Selmer[1,2], E. Bryne[3], O. Rødningen[1], M. Fannemel[1].* 1) Dept Medical Genetics, Oslo Univ Hospital, Ullevaal, Oslo, Norway; 2) Institute of Medical Genetics, University of Oslo, Oslo, Norway; 3) Habilitation Center, Vestfold Hospital, Tønsberg, Norway.

The 1q44 deletion syndrome has shown to be a recognizable phenotype with developmental delay, short stature and corpus callosum abnormalities as relatively consistent features. However, the disorder is still clinically heterogeneous and a genotype-phenotype correlation has been challenging to establish. In particular, a delineation of a critical region for the corpus callosum development has turned out to be difficult, and many candidate genes have been proposed. We present here a patient boy with a clinical picture of the 1q44 deletion syndrome, including a thin corpus callosum, and the hitherto smallest 1q44 deletion reported in this syndrome. The deletion spans a maximum of 163 kb and includes only the two genes *FAM36A* and*HNRPU*. This finding supports the previously suggested hypothesis that the *HNRPU* is an essential gene to the development of corpus callosum. However, as patients with deletions outside this interval also have been reported to have corpus callosum abnormalities, other mechanisms are probably also involved. We identified two conserved non-coding regions in the deleted region of the patient, and speculate that also other elements interfere with the complex interplay and spatiotemporal gene expression during embryonic development.

3097W

Genotype phenotype correlation study of 2q23.1 microdeletion syndrome and characterization of three new patients. *L. Zhang[1], H.T. Bjornsson[2], D. Batista[3], G. Wiesner[1], T. Wang[2], A. Parikh[1].* 1) Dept Med/Human Gen, Case Western Univ SOM, Cleveland, OH; 2) IGM, Johns Hopkins Univ, Baltimore, MD; 3) Kennedy Krieger Inst, Baltimore, MD.

2q23.1 microdeletion syndrome is a newly characterized syndrome with 20 patients described in clinical detail to date. Haploinsufficiency of *methyl CpG-binding domain protein 5 (MBD5)* is implicated to account for the core features of this syndrome including intellectual disability, severe speech impairment, epilepsy, behavioral problems, mild craniofacial dysmorphism, small hands and feet, and broad based ataxic gait. We have previously reported a patient with de novo 210 kb deletion at 2q23.1 involving 5'-UTR regions of *MBD5* and *ORC4L (origin recognition complex, subunit 4-like)*. We now report two new patients with de novo deletions at 2q23.1, one with a 110kb deletion involving only the 5'-UTR of *MBD5* and *ORC4L* and another with a 2.7 Mb deletion involving multiple genes including *MBD5*. We compared the genetic and clinical findings in our patients with those reported in the literature and conducted a genotype phenotype correlation study on the genes within these deletions and clinical features of the three patients. Heterozygous deletion of *ORC4L* does not contribute significantly to the clinical phenotype. Haploinsufficiency of *EPC2 (enhancer of polycomb homolog 2)* is likely associated with microcephaly and short stature. Studying the spectrum of clinical phenotype in patients with different sizes of microdeletions provides an opportunity to delineate gene functions at 2q23.1 and understand the pathophysiology of complex neurodevelopmental phenotypes including autism and intellectual disability in 2q23.1 microdeletion syndrome.

3098F

Polycystic kidney disease and a 2.5Mb duplication on chromosome 3p22.1. *K. Dahan[1], M. Madhoun[2], M. Gilliaux[3], V. Benoit[1], B. Grisart[1].* 1) Ctr Human Gen, Institut de Pathologie et de Génétique, Gosselies, Belgium; 2) Nephrology Departement, CHU de Charleroi, Montigny-le-Tilleul, Belgium; 3) Gynecology Departement, CHU de Charleroi, Montigny-le-Tilleul, Belgium.

The use of high-resolution microarray technology for investigation of patients with intellectual disability or congenital anomalies provided the unique possibility to identify new microdeletion/microduplication syndromes and discover the dosage sensitive genes, which are implicated in the manifestation of various genetic conditions. We report 3p22.1 microduplication, 2.5 Mb in size, detected in a 19.5-year-old woman with enlarged cystic kidneys and slight facial dysmorphism, . There are 20 RefSeq genes included in this duplication. In comparaison with transgenic mice models, CTNNB1 gene looks to be a strong candidate gene for the renal phenotype strongly mimicking autosomal dominant polycystic kidney disease. Further cases with similar duplications will contribute to the delineation of a potential new microduplication syndrome of 3p22.1.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3099W

Subtelomeric deletion including the WHSC1 gene located in the critical region of the Wolf-Hirschhorn Syndrome. *A.L.S. Ludmila Serafim, J.M. Pina Neto, I.F. Mazzucatto.*Genetic, Universidade de São Paulo, Ribeirão Preto, SP., Brazil.

We report here a patient with severe intellectual disability, who initially showed a deficit of intrauterine growth. The proband was born premature and hypotonic. After was observed short postnatal stature, dysmorphic facial characteristics suggestive of Pitt-Rogers-Danks Syndrome (PRDS), such as microcephaly, high forehead, frontal hair line high, facies significantly abnormal, triangular face, hypertelorism with telecanthus, eye prominent, high and arched eyebrow, midface hypoplasia, thin lips, high palate, abnormal helix, appendix preauricular, and a small hand. It was also reported that convulsive crisis started at six months of age. After performing classical cytogenetic examination which did not detect any changes, we used the MLPA (Multiplex Ligation Probe Amplification) technique with kit P036-E1 to investigate the presence of submicroscopic chromosomal abnormalities in subtelomeric regions and also the P245-A2 Kit which identifies 21 microdeletion syndromes. The P036 kit detected a deletion in the gene PIGG, located in the chromosomic region 4p16.3. While P245-A2 kit showed a deletion in the gene WHSC1, confirming the alteration in the 4p16.3 region previously detected by the kit P036-E1. The critical region of Wolf-Hirschhorn syndrome (WHSCR) is located at 4p16.3. The WHSC1 gene is located in this. Despite of the deletion of the WHSC1, the proband has phenotypic characteristics slightly different of the WHS. Although previously described as distinct disorders, it is now known that WHS and PRDS represent the clinical spectrum associated to a single syndrome. Some evidences show that PRDS represents a variant of WHS, since the occurrence of deletion in the 4p16.3 region occurs in both sydrome. It has also been found a deletion in the PIGG gene, supposing that the deletion in the region 4p has a minimum size of 1.4 Mb. In this clinical case, besides the use of different sets to confirm the existence of specific alteration, it has been also possible provide a better understanding of the extent of this change. It is believed that the combination of kits with additional probes may eventually offer the best solution for routine diagnostic screening of the submicroscopic chromosomal aberrations, since it allows a precise identification and better delineation of these alterations.

3100F

Atypical Form of Osteopathia Striata with Cranial Sclerosis and Increased Mineral Density in an Adolescent Female with Mosaic Chromosome 2 Rearrangement. *R. Mendoza-Londono[1,2], A. Howard[2,3], E. Sochett[2,4], L. Dupius[1,2], D.J. Stavropoulos[5], R.C.C. Wong[6], S. Robertson[7], A.M. Joseph-George[5,6].* 1) Div Clinical & Metabolic Gen, Hosp Sick Children, Toronto, ON, Canada; 2) The Bone Health Centre, The Hospital for Sick Children and University of Toronto; 3) Department of Orthopaedic Surgery, The Hospital for Sick Children and University of Toronto, Canada; 4) Division of Paediatric Endocrinology, The Hospital for Sick Children and University of Toronto, Canada; 5) Cytogenetics Laboratory, Division of Molecular Genetics, Department of Paediatric Laboratory Medicine, The Hospital for Sick Children and University of Toronto, Toronto, Canada; 6) The Centre for Applied Genomics, The Hospital for Sick Children, Toronto; 7) Department of Women's and children's health, Dunedin school of Medicine, University of Otago, Otago New Zealand.

Osteopathia Striata with Cranial Sclerosis (OSCS-MIM 300373) is a rare disorder of bone mineralization caused by mutations of *WTX*, a protein that regulates WNT signaling. Females with this disorder usually present with macrocephaly, learning disability, sclerosis of the base of the skull, and longitudinal striations of the long bones. We present a 17 year old female with mild learning disabilities, significantly increased bone mineral density (Z score +8.1 at the level of L2–L4), and radiological findings reminiscent of OSCS. Family history was negative for other similarly affected individuals and both parents have normal bone mineral densities. On physical examination, the patient presented with long and narrow face with prominent jaw, marfanoid body habitus (arm span to height ratio of 1.06) and arachnodactyly. Sequencing of *WTX* single coding exon did not reveal any pathologic mutations. ArrayCGH on a 44K oligo platform did not reveal any copy number variations that could explain her phenotype. Chromosome analysis on lymphocytes and fibroblasts revealed a *de novo* mosaic karyotype with two apparently balanced cell lines in both tissues: ~32% of metaphases had an altered chromosome 2, while the remaining cells had a normal karyotype. By G-band analysis the derivative chromosome 2 has two separate intrachromosomal rearrangements that involve five breakpoints. The first is an interarm insertion of material from 2q13-2q21.1 into band 2p11.2 and the second, a paracentric inversion in the long arm with breakpoints estimated at 2q22.2 and 2q33.1. FISH using genomic BAC clones have confirmed this interpretation. Studies are underway to further define the breakpoints and determine which gene(s) are disrupted by the rearrangement and may be responsible for different aspects of the phenotype. Identification of the responsible gene may allow the identification of an autosomal regulator of WNT signaling on chromosome 2.

3101W

Further evidence of the role of HOXA genes in anatomical development. *P.A. Mowery-Rushton[1], R.A. Schultz[1], J.B. Ravnan[1], J.E. Burton[2,4], J. Kussmann[3,5], J.A. Rosenfeld[1], B.C. Ballif[1], L.G. Shaffer[1].* 1) Signature Genomic Laboratories, PerkinElmer Inc, Spokane, WA; 2) Carle Foundation Hospital, Urbana, IL; 3) University of Missouri, Columbia, MO; 4) Current affiliation: University of Illinois College of Medicine at Peoria, Peoria, IL; 5) Current affiliation: Children's Mercy Hospital, Kansas City, MO.

Introduction: The HOXA gene cluster encodes transcription factors that play a role in regulating patterns of anatomical development during embryogenesis. Mutations of specific HOXA genes and deletions of the entire cluster have been described in the literature associated with autosomal recessive Athabaskan brainstem dysgenesis syndrome (HOXA1), autosomal recessive microtia, hearing impairment and cleft palate (HOXA2), autosomal recessive radio-ulnar synostosis with amegakaryocytic thrombocytopenia (HOXA11), and autosomal dominant hand-foot-genital syndrome and Guttmacher syndrome (HOXA13). Here we describe two patients with small overlapping deletions of a portion of the HOXA gene cluster with similar phenotypes in order to further delineate the role of HOXA genes during embryogenesis. Materials and Methods: We analyzed the results of oligonucleotide-based microarrays on over 23,000 patients stored within the Genoglyphix® database. Results: We identified two patients with relatively small deletions (54.8 and 80.7 kb), which include part of the HOXA gene cluster. Both children had hypospadias with kidney abnormalities and mild dysmorphic features. One child had multiple heart defects, fifth finger clinodactyly, tapered digits, and hallux varus. The other child had hypotonia, developmental delay, movement disorder, and laryngomalacia, but not abnormalities of his digits. The smallest region of overlap is 30.2 kb and includes HOXA10, HOXA11, and the 3' end of HOXA13. Conclusion: These results reinforce the role of the HOXA genes in genitourinary tract development and suggest that haploinsufficiency of HOXA10 and/or HOXA13 contributes to renal and genital abnormalities. The additional anomalies noted in each child may be due to deletion of additional genes in the non-overlapping regions or disruption of expression of other HOXA genes within the cluster, due to the deletion.

3102F

4p16.3 deletion limited to WHSCR2. *N. Okamoto[1], K. Shimojima[2], T. Yamamoto[2].* 1) Dept Medical Genetics, Osaka Med Ctr/Res Inst, Osaka, Japan; 2) Tokyo Women's Medical University Institute for Integrated Medical Sciences, Tokyo, Japan.

Wolf-Hirschhorn syndrome (WHS) is a contiguous gene deletion syndrome associated with growth retardation, developmental disabilities, epileptic seizures and distinct facial features resulting from a deletion of the short arm of chromosome 4. Two critical regions for WHS have been narrowed to a 200kb region about 1.9Mb from 4p telomere. Wolf-Hirschhorn Syndrome Critical Region 1 (WHSCR1) is a 165kb region proximal to the *FGFR3* and *LETM1* genes. WHSCR1 includes the 3' part of *WHSC1*and the *WHSC2*. WHSCR2 is distal to WHSCR1 and directly adjacent to it. WHSCR2 includes the *LETM1* and 5' end of the *WHSC1*.*WHSC1* is one of the nuclear receptor binding SET domain protein is a family of three histone-methyltransferase. The haploinsufficiency of *WHSC1* is believed to be responsible for a number of WHS characteristics. *LETM1* has been suggested as a candidate gene for the seizures in WHS. We report on a patient with severe growth retardation, microcephaly and a characteristic facial appearance reminiscent of WHS. The 2-year-old male propositus was the first-born child of healthy and nonconsanguineous parents. Fetal echogram revealed severe intra uterine growth retardation. He fed poorly and his physical growth was severely retarded from early infancy. However, his developmental mile stones were not so delayed. His development quotient was 79 at 2 years of age by Japanese standard method. Physical examination identified dysmorphic features, including a triangular face, prominent glabella, a high nasal bridge, bilateral low set ears, down slanting palpebral fissures, a high arched palate and pointed chin. Hearing and visual acuity were normal. His weight was 5.7 kg (-4.8SD), his length was 70.1 cm (-4.8SD) and head circumference was 41cm (-4.8SD). He was free from seizures. His karyotype by G-banded analysis was 46,XY. By array-CGH analysis, loss of genomic copy numbers was identified in the region 4p16.3, which included *FGFR3*, *LETM1* and *WHSC1*. The size of the deletion was only 101Kb. This is the minimal deletion of 4p16 so far reported. The deletion was compatible with WHSCR2. From these studies, we suspect that haploinsufficiency of *WHSC1* is the most probable cause of growth deficiency, microcephaly and characteristic facial features in our patient. Studies from his clinical manifestations will be a good help to delineate the role of *WHSC1*.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3103W

Array CGH in 190 Korean patients with developmental delay and/or intellectual disability: a single tertiary care university center study. *Y. Sohn[1], C. Lee[1], S. Park[2], J. Yun[1], E. Jung[1], H. Kim[1], S. Yim[3].* 1) Medical Genetics, Ajou University Medical Center, Ajou University School of Medicine, Suwon, South Korea; 2) MG MED, Inc., Seoul, Republic of Korea; 3) Physical Medicine and Rehabilitation, Ajou University Medical Center, Ajou University School of Medicine, Suwon, Republic of Korea.

This study analyzed demographic, clinical, and cytogenetic data (G-banded karyotyping and array CGH) of patients with unexplained DD/ID in Korean single institution. We collected clinical and cytogenetic data based on retrospective chart in Ajou University Medical Center, Suwon, Korea from April 2008 to March 2012 A total of 190 patients identified. The mean age was 5.1 ± 1.87 years. Array CGH yielded abnormal results in 26 of 190 patients (13.7 %). Copy number losses were about 2 folds more frequent than gains. A total 61.5% patient had copy number losses. The common deletion disorders are 22q11.2 deletion syndrome, Angelman syndrome and 18q deletion syndrome. Copy number gains were identified in 34.6% patients and common diseases are Potocki-Lupski syndrome, 15q11-13 duplication syndrome and duplication 22q. Abnormal karyotype with normal array CGH was exhibited in 2.6% patients; balanced translocation (n=2), inversion (n= 1) and low-level mosaicism (n=1). Craniofacial dysmorphic features (P = 3.53 × 10-4) and failure to thrive were also more frequent in group of abnormal CGH. We expected that array CGH complemented by karyotype analysis is useful diagnostic tool, in selected patient groups, especially ID/ DD combined with craniofacial anomaly or failure to thrive.

3104F

Diagnostic investigation in Rwandan patients with MR/MCA. *A. Uwineza[1,2], J. Hitayezu[1], L. Mutesa[1].* 1) Medical genetics, National University of Rwanda, Huye, Rwanda; 2) Center for Human Genetics, CHU Sart-Tilman, University of Liège, Belgium.

Multiple congenital anomalies and mental retardation (MCA/ MR) affect approximately 3% of newborns. There is no available data on genetics of MR and MCA in Central Africa. The aims of this study was to detect chromosomal abnormalities such as submicroscopic microdeletion/microduplication and subtelomeric imbalance in this group of patients. Newborn, children and young adults were recruited from the Teaching University Hospitals of Kigali and Butare, Rwanda. All individuals were clinically examined and karyotype was done in Rwanda and a blood sample was obtained for genetic testing in Belgium (Center for human genetics, Liege). All patients with trisomy 21, 13 and 18 and were excluded from the study. Karyotype analysis of the remaining 52 patients was normal. Fragile-X was excluded in all these patients and we found premutation in 2 male patients. Thus the MLPA was performed and identified (SALSA MPLA P245 Microdeletion and MLPA P0 36 and P070 Human Telomere, MRC Holland) one chromosome 10q duplication, one chromosome 9p duplication, one deletion of the 22q11.2 region and one 7q11.23 microdeletion (William-Beuren syndrome). For the remaining cases array-CGH will be performed. This study will allow to identify the etiology of MR/MCA in Rwandan patient and will show the importance of genetic testing in developing countries.

3105W

Clinical features and genomic characterization of two subjects with a pure duplication of 9q34. *H. Yoshihashi[1], M. Ikegami[2], C. Torii[3], K. Kosaki[3].* 1) Division of Medical Genetics, Tokyo Metropolotan Children's Medical Center, Fuchu-shi, Japan; 2) Department of Pediatrics, Tokai University Hachioji Hospital, Tokyo, Japan; 3) Center for Medical Genetics, Keio University, School of Medicine, Tokyo, Japan.

Introduction: Duplications of 9q34 have been rarely described in the literature. New molecular techniques allow the identification of subtelomeric chromosomal rearrangements, but additional subjects of a duplication of 9q34 have not been increasingly reported. In the previous reports of the duplication 9q34, several clinical features were in common, including a distinctive face, arachnodactyly, hypotonia and developmental delay. However, the additional phenotypic spectrum and genomic critical region remained unclear. We presented two subjects with a pure duplication of 9q34 and compared with the clinical features of previously reported those using array CGH analysis. Materials and methods: Case1. 8y-female. 42w0d, 2294g. She had a distinctive face, laterality of ear malformations, long fingers, clinodactyly, hypotonia during infancy, mild intellectual disability and atlantoaxial rotatory subluxation. The analysis of G-banding, FISH and SKY indicated a distal interstitial duplication of 9q34. The karyotype was 46,XX,dup(9)(q34.1q34.3). Case2. 10mth-male. 40w5d, 3075g. He had a distinctive face, hypospadias, VSD, long fingers and toes, adducted thumb, camptodactyly, hypotonia, hip joint abduction, developmental delay. On the basis of findings of G-banding and FISH, the karyotype was 46,X,der(Y)-t(Y;9)(q11.2;q33). Array CGH analysis was performed using the ISCA 180K microarray platform (Agilent Technologies) to determine the precise breakpoints of 9q34 segments and identify any other genomic abnormalities. Results: The size of segments in the Case1 and Case 2 were 7.7Mb(130.7–138.4 Mb; from the pterminus) and 12.8Mb(127.4–140.2Mb), respectively. No other chromosomal rearrangements were significantly detected. Discussion: The phenotype observed in our subjects with a pure duplication of 9q34 represented a variety of clinical features depending on the size of an individual duplication. These detected segments included the approximately 4.1Mb(131.7–135.8Mb) of 9q34 as the overlapping region documented in the current report, which was highly suggestive of critical region of 9q34 duplication syndrome. Further genetic study may provide insight into the clinical characterization of 9q34 duplication syndrome.

3106F

Osteopathia Striata with Cranial Sclerosis and Developmental Delay in a Male with a Mosaic Deletion in Chromosome region Xq11.2. *S. Chenier[1], A. Noor[1], L. Dupuis[2], D.J. Stavropoulos[1,3], R. Mendoza-Londono[2].* 1) Department of Paediatric Laboratory Medicine, The Hospital for Sick Children, Toronto, Ontario, Canada; 2) The Hospital for Sick Children, Department of Pediatrics, Division of Clinical and Metabolic Genetics, University of Toronto, Toronto, Ontario, Canada; 3) Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Ontario, Canada.

Osteopathia striata with cranial sclerosis (OSCS) is an X-linked disease caused by mutations involving *WTX* (*FAM123B*), a tumor suppressor protein with dual functions. OSCS typically affects females whereas males generally have fetal or neonatal lethality. Surviving affected males have characteristic facial dysmorphisms, skeletal features such as macrocephaly and short stature, neurodevelopmental disabilities and a high prevalence of neuromuscular anomalies. On imaging, hemizygous males display more marked cranial and peripheral skeletal sclerosis without the metaphyseal striations that are seen in women with OSCS. Observations of striation in males may be indicative of a somatic mosaic mutation in *WTX*. To date only two cases of surviving males haves been confirmed with mosaic point mutations in *WTX*. We report the first case of a male with mosaic deletion of the whole *WTX* gene. We show that a mosaic deletion in a hemizygous male patient can cause a mild phenotype of OSCS, including facial and skull base bone striations, nasal stenosis, conductive hearing loss, global developmental delay and mild facial dysmorphology without short stature or macrocephaly.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3107W

Keratoconus associated with Williams-Beuren Syndrome: new case report. *M.C. Frasson[1], M.M. Viana[2], L.L. Leao[2], M. Stofanko[3], H. Goncalves-Dornelas[3], P.S. Cunha[4], M.J.B. Aguiar[2].* 1) Serviço de Retina, Hospital São Geraldo, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil; 2) Serviço Especial de Genética Médica, Hospital das Clínicas, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil; 3) Núcleo de Genética Médica (GENE), Belo Horizonte, Minas Gerais. Brazil; 4) Departamento de Bioquímica e Imunologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil.

Background: Williams-Beuren syndrome is a multisystemic genetic disorder caused by the contiguous gene deletion at 7q11.23. Keratoconus is a complex disease and it is suspected to have a genetic origin, although the specific gene responsible for keratoconus has not been identified. Although there are several ocular features in Williams-Beuren syndrome, keratoconus is not regularly described as part of this syndrome. Purpose: To report a new patient with keratoconus and Williams-Beuren syndrome. Discussion: This is the third case of an association between Williams-Beuren syndrome and keratoconus. The authors believe that the Williams-Beuren Syndrome Chromosome Region can be a possible target for further investigation as the genetic basis of keratoconus.

3108F

Minimal interstitial deletion in a patient with de novo 15q24 microdeletion syndrome presenting with severe scoliosis. *X. Gao[1], C. Johnston[2], S. Sparagana[3], C.A. Wise[1,4].* 1) Sarah M. and Charles E. Seay Center for Musculoskeletal Research, Texas Scottish Rite Hospital for Children, Dallas, TX; 2) Department of Orthopedic Surgery, Texas Scottish Rite Hospital for Children; 3) Department of NeurologyTexas Scottish Rite Hospital for Children; 4) Department of Orthopaedic Surgery and McDermott Center for Human Growth and Development, University of Texas Southwestern Medical Center Dallas.

Early onset “infantile” or juvenile onset idiopathic scoliosis (IS) cases comprise less than 10% of all IS patients, yet deserve special attention because of their high risk of progressive deformity. Respiratory compromise is particularly problematic in this population. We report a Hispanic male patient who presented at our institution at ten years of age with a diagnosis of IS. The family reported onset of deformity at three years, with brace correction from five to ten years of age. Prior genetic studies were reportedly normal. Re-evaluation at age 13 revealed hypotonia, large low set ears, flat nasal bridge, decreased sensitivity to pain, mild sacral dimpling, mild intellectual disability, and a severe left thoracic spinal deformity measuring 75 degrees. Brain MRI findings showed mild cerebral atrophy. Scoliosis was surgically corrected with instrumentation and fusion to avoid further deformity and respiratory compromise. Repeat chromosomal studies using array-CMA-HR+SNP revealed del15q24.1q24.2, spanning approximately 1.6 Mb (74.5–76.1 MB, hg19) in the patient and normal CMA in the father. Normal chromosomes were subsequently confirmed in both parents by Illumina OmniExpress-12 v1.0 beadchip arrays. These findings are consistent with de novo 15q24 microdeletion syndrome and may represent the smallest deletion yet described. 15q24 microdeletion syndrome was first described in 2005 in three patients with hypotonia, dysmorphic facial features, and genital anomalies in males. More recently, a review of thirteen cases in the literature noted three with scoliosis, and no mention of skeletal involvement in the remaining ten. We subsequently reviewed our database of 548 idiopathic scoliosis (IS) cases (66 EOS, 482 adolescent IS) and 744 healthy controls screened by Illumina OmniExpress-12 v1.0 beadchip arrays to assess the frequency of copy number variation in the 15q24 region. We found no other deletions (>10 kb) within the 15q24 region in patients or controls, suggesting that interstitial deletion of this region is not frequently associated with IS. Our patient helps to further delineate genotype/phenotype correlations in 15q24 microdeletion syndrome. We suggest that 15q24 microdeletion syndrome patients should be closely monitored for progressive scoliosis. Our findings also illustrate the potential utility of molecular genetic re-evaluation using updated platforms for patients with aggressive “idiopathic” scoliosis and other non-specific findings.

3109W

Microdeletions in 9q33.3-q34.1 may constitute a contiguous gene deletion syndrome characterized by developmental delay, microcephaly, mild dysmorphisms, strabismus, and seizures of incomplete penetrance. *J.A. Lee[1,2], J.K. Ehret[1], E. Wohlleber[1], S. Vogt[1], A.M. Zink[1], E. Rossier[3,4], M. Bonin[3,5], U. Grasshoff[3], A. Bevot[6], M. Mathieu-Dramard[7], G. Plessis[8], A. de Broca[9], S. Kanafani[7], B. Röthlisberger[10], M. Holder-Espinasse[11], I. Simonic[12], L. Willatt[12], P. Miny[13], I. Filges[13], J. Andrieux[14], H.V. Firth[12], A. Dufke[3], H. Engels[1].* 1) Dept Genomics, Life & Brain Ctr, Inst Human Gen, Univ Bonn, Bonn, Germany; 2) Department of Genomics, Life & Brain Center, University of Bonn, Bonn, Germany; 3) Department of Medical Genetics, University of Tübingen, Tübingen, Germany; 4) Genetikum, Neu-Ulm, Germany; 5) Microarray Facility, University of Tübingen, Tübingen, Germany; 6) Children's Hospital, University of Tübingen, Tübingen, Germany; 7) Centre de Génétique, CHU d'Amiens, France; 8) Service de Génétique CHU Clémenceau, France; 9) Service de Neuropédiatrie, CHU d'Amiens, France; 10) Service de Neuropédiatrie, CHU d'Amiens, France 10 Division of Medical Genetics, Center of Laboratory Medicine, Cantonal Hospital Aarau, Switzerland; 11) Service de Génétique Clinique, Hôpital Jeanne de Flandre, CHRU de Lille, France; 12) Department of Medical Genetics, Addenbrooke's Hospital NHS Trust, Cambridge, UK; 13) Division of Medical Genetics, University Children's Hospital and Department of Biomedicine, Basel, Switzerland; 14) Laboratoire de Génétique Médicale, Hôpital Jeanne de Flandre, CHRU de Lille, France.

We report six patients with overlapping microdeletions of chromosome 9q identified by molecular karyotyping. Their common clinical features include psychomotor developmental delay with delayed or absent speech, strabismus, and microcephaly. *De novo* deletions ranging from 1.23 to 4.13 Mb in 9q33.3-q34.11 were detected, whereas the smallest deletion was detected in patient 4, whose 432-kb deletion was inherited from her mother who is reported to have mild intellectual disability. The smallest region of overlap (SRO) concerning the common clinical features is defined by the deletion in patient 4, and includes only three RefSeq genes. Interestingly, the larger deletions have 29 genes in common including *STXBP1*. *STXBP1* loss-of-function mutations and deletions have been associated with early infantile epileptic encephalopathy (EIEE, or Ohtahara syndrome), and were previously identified in patients with intellectual disability and non-syndromic epilepsy. In our cohort, however, the seizure phenotype appears to be incompletely penetrant (two of six patients). Notably, patient 4 has seizures, although her deletion excludes *STXBP1*. Thus, although one or more of the SRO genes may likely contribute to the phenotype, the involvement of *STXBP1* and/or possibly other genes due to deletion of long-range regulatory elements cannot be ruled out. Compared to patients with *STXBP1* mutations, our microdeletion patients are clinically similar except for distinctive microcephaly and strabismus. Publications by Saitsu and colleagues and Mignot and colleagues support this genotype-phenotype correlation, and may help to further define the SRO. Taken together, this suggests the possibility of a contiguous gene deletion syndrome involving several genes contributing to the phenotype. Sequencing of the three SRO genes in a cohort of 192 unrelated patients with mild to severe idiopathic intellectual disability, as well as gene expression analyses in our patients, is underway to identify the causative gene(s). Thus far, no causative mutations have been identified, arguing against mutations in these genes being a common cause for intellectual disability. We suggest that microdeletions of this region on chromosome 9q cause a clinical spectrum including developmental delay especially concerning speech, microcephaly, mild dysmorphisms, strabismus, and seizures of incomplete penetrance, and may constitute a new contiguous gene deletion syndrome.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3110F

Identification and diagnosis of Williams-Beuren syndrome in a public genetic center in Brazil. *M. M. Viana[1], M. Stofanko[2], P. S. Cunha[3], H. Goncalves-Dornelas[2], M. J. B. Aguiar[1], S. D. J. Pena[2, 3].* 1) Serviço de Genética Médica, Hospital das Clínicas, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil; 2) GENE (Núcleo de Genética Médica), Belo Horizonte, Minas Gerais, Brazil; 3) Departamento de Bioquímica e Imunologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil.

Introduction: Brazil is considered one of the emerging economies in the world. It has a public health system (SUS - Sistema Único de Saúde) that covers medical expenses for 80% of the population and demands cheaper diagnostic approaches when compared to developed countries. For instance, the cost of aCGH techniques is prohibitive for our reality. Thus, genetic diagnosis based on clinical data and testing focused at specific targets is more appropriate than genome wide screening procedures. To test our hypothesis we selected patients with suspected Williams-Beuren Syndrome (WBS), a multisystem disorder with well delineated phenotype, which is caused by microdeletions at 7q11.23. Purpose: To describe our experience in identifying and confirming WBS through a simple and inexpensive targeted diagnostic method. Methods: Between 2005 and 2010, 2,399 patients were evaluated for the first time at a public genetic center in Minas Gerais, Brazil. The medical charts were analyzed and a total of 26 patients were selected because of suggestive clinical features of WBS; 17 accepted to take part of the trial. Three (3) more patients were evaluated in 2011 and accepted to participate. DNA was extracted from oral mucosa cells and submitted to a simple and inexpensive novel technique developed called Microdeletion Fluorescent Quantification by Polymerase Chain Reaction (MFQ-PCR). All results were confirmed by loss of heterozygozity assessed by multiple microsatellites or real-time PCR. Results: A total of 20 samples were collected. Twelve tested positive for WBS. As expected, the main characteristics of positive cases were friendly personality (83%), prominent lips (83%), slanted palpebral fissures (75%), microcephaly (58%), long philtrum (50%), anteverted nares (42%), stellate pattern of the iris (33%). The main cardiac abnormalities were pulmonary stenosis (42%), mild to moderate aortic regurgitation (33%) and supravalvular aortic stenosis (33%), 2 patients had normal echocardiograms (17%). Discussion: In emerging economies, it is necessary to take into consideration the cost of diagnostic tests. Our report shows that it is possible to adequately make precise diagnosis using clinical expertise together with simple, targeted, cost appropriate diagnostic techniques. This experience can be extended to other genetic syndromes.

3111W

Implementation of molecular karyotyping in routine diagnostics provides new insights into the pathophysiology of Silver-Russell syndrome. *S. Spengler[1], M. Begemann[1], N. Ortiz Brüchle[1], P.M. Kroisel[2], B. Oehl-Jaschkowitz[3], G. Raabe-Meyer[4], C. Spaich[5], P. Blümel[6], U. Moog[7], K. Zerres[1], T. Eggermann[1].* 1) Institute of Human Genetics, RWTH Aachen University, Aachen, Germany; 2) Institute of Human Genetics, Graz, Austria; 3) Gemeinschaftspraxis für Humangenetik, Homburg, Germany; 4) Praxis für Humangenetik, Hannover, Germany; 5) Institute of Clinical Genetics, Stuttgart, Germany; 6) G v Preyer Childrens` Hospital, Wien, Austria; 7) Institute of Human Genetics, Heidelberg, Germany.

Silver-Russell syndrome (SRS) is a clinically and genetically heterogeneous disorder characterized by severe growth retardation, relative macrocephaly, a triangular face and asymmetry of body and/or limbs. The clinical spectrum is broad and the diagnosis is rather subjective. Therefore a comprehensive molecular diagnostic workup is needed. Unfortunately, the currently applied diagnostic algorithm leaves about 50 % of patients without molecular diagnosis: (epi)genetic disturbances in 11p15 can be detected in ~40% of patients and maternal uniparental disomy of chromosome 7 in further ~10% (upd(7)mat). By implementation of molecular karyotyping (Affymetrix SNP Array 6.0) in our diagnostic algorithm for SRS we could significantly improve the detection rate for genetic disturbances and thereby contribute to a more personalized clinical management: 19.5% of patients without the characteristic (epi)genetic disturbances on chromosomes 7 and 11 showed pathogenetic copy number variants. As some of them were associated with known microdeletion syndromes, our results illustrate the broad genetic heterogeneity of SRS and the difficulty in clinical diagnosis. Furthermore, the application of molecular karyotyping helps us to get new insights in the molecular aetiology of SRS, either by identifying new genetic regions or by precise characterisation of genetic disturbances affecting the known SRS loci. By analyses of imbalances affecting the imprinting control regions (ICRs) in 11p15 and 7q we obtained relevant data on the complex regulation mechanisms and interactions between the different loci illustrating that the clinical outcome is influenced by the size, the breakpoint positions and the parental origin of the imbalances. In conclusion, implementation of molecular karyotyping in the diagnostic algorithm for patients with SRS features is urgently needed because: a) causal (sub)microscopic imbalances can be found in a significant proportion of patients without the characteristic (epi)genetic disturbances, b) detailed analysis of CNVs in 11p15.5 allows insights in the complex regulation of the imprinted factors localised in this region, c) SRS candidate regions can be discovered, d) based on the molecular findings risk figures can be delineated.

3112F

Clinical, molecular cytogenetic evaluation and genotype-phenotype correlation of a patient with partial Jacobsen syndrome without thrombocytopenia caused by an 13 Mb deletion of del(11)(q24.1). *J. Moon[1], J. Kim[1], J. Kim[2], M. Nam[1], S. Yoon[1].* 1) Laboratory Medicine, Korea Univ. Ansan Hosp., Ansan, Gyeonggi, South Korea; 2) Laboratory Medicine, Korea Univ. Anam Hosp., Seoul, South Korea.

Clinical manifestations of Jacobsen syndrome (JBS) depend on the size of the 11q deletion, which usually varies between ~5 and 20 Mb. Typical JBS features include developmental delay/mental retardation, congenital heart defects, thrombocytopenia, and characteristic dysmorphic craniofacial features. We present a patient with de novo terminal deletion of chromosome 11q24.1q25. The clinical phenotype in our patient, born prematurely, includes mild neonatal thrombocytopenia, neonatal jaundice, short stature, and feeding difficulties. She has mild facial abnormalities, including epicanthal folds mimicing simian crease right hand and both toe deformities and a sacral malformation. The patient did not show any abnormalities in hearing screening tests, brain and abdominal ultrasonography, and heart echocardiogram except for persistent foramen ovale of prematurity, but no substantial congenital heart defect. Cytogenetic analyses revealed the karyotype of 46,XX,del(11)(q23.3). Using fluorescence in situ hybridization (FISH) for *MLL*, located on 11q23 showed normal signals, in contrast with 11q subtelomere FISH, showing one signal deletion. Utilizing whole-genome cytogenetic array analyses, we identified an ~13 Mb deletion of the terminal part of chromosome 11q in the patient. The deletion breakpoint was mapped between 121,560,927 and 134,926,021 bp. This includes multiple genes including LOC. Interestingly, the patient's intact hearing and abscence of heart and kidney abnormalities is suggestive of incomplete penetrance, but relatively mild facial defects support previous data that relatively smaller deletions not involving 11q23 usually have less number of facial dysmorphisms. The patient's parents' karyotypes and 11q subtelomere FISH results were normal, revealing the patient's abnormality as de novo. Notably, neither thrombocytopenia nor congenital anomalies were detected in this family, including the patient's brother. Our data can be compared with previous cases patients with deletion of similar region and multiple anomalies including strabismus and toe abnormalities.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3113W

Triplication of 7q11.23 distal to the critical region for the Williams-Beuren syndrome. *J.G. Pappas[1], R.D Nass[2], E. Ward[1], I.K. Gadi[3].* 1) Dept Pediatrics, Clin Genet Svcs, New York Univ, Sch Med, New York, NY; 2) Dept Neurology, Pediatrics, Child and Adolescent Psychiatry, New York Univ, Sch Med, New York, NY; 3) Cytogenetics, Laboratory Corporation of America, Research Triangle Park, NC.

We present a 14 year old boy with autism spectrum disorder and 1.1 Mb interstitial triplication of 7q11.23 distal to the critical region for the Williams-Beuren syndrome. He presented at age 3 with global developmental delays especially no speech development. Peripheral blood chromosome analysis and fragile-X test by PCR were reported normal. At age 13 he had a neurodevelopmental evaluation which revealed mild intellectual disability, autism with significant apraxia. At age 14 his height was 163 cm, his weight was 52.7 kg and his head circumference 55 cm; all around the 50% centile. He had normal puberty development, sloping forehead and facial asymmetry with right face larger than left, up-slanted eyes with inverted lower eyelids, attached earlobes, down turned corners of the mouth and tapering fingers with fetal pads. Whole genome SNP microarray analysis was performed using the Affymetrix 6.0 platform which uses over 900,000 SNP probes and 900,000 NPCN probes with a median spacing of 0.7 Kb. This microarray revealed triplication 7q11.23 [arr 7q11.23(74,882,566–75,981,763)x4]. This interstitial triplication was confirmed by the region specific BACs RP11-951G4 and RP11-107L23 on 50 interphase and 10 metaphase cells. These BACs were used for parental testing which showed normal hybridization signals on both chromosome 7 homologues in all nuclei of the parental blood samples. These results were consistent with a de novo origin of the triplication. The OMIM annotated genes included in the triplication interval were: *TRIM73*, *HIP1*, *CCL26*, *CCL24*, *POR*, *MDH2*, *HSPB1*, *YWHAG*, *SRCRB4D*, *ZP3* and *DTX2* (NCBI36/hg18). There is no report in the medical literature with the triplication seen in our case. Ramocki et al. (2010) reported cases with 1.2 Mb deletion and two siblings with the reciprocal 1.2 Mb duplication of a segment overlapping the triplicated segment in our case. Common features of the reported cases with the deletion included epilepsy, mild intellectual impairment, neurobehavioral abnormalities including autism. The reported siblings with the duplication had behavioral disorders and no epilepsy. Ramocki et al (2010) studied cases with smaller deletions and concluded that copy number variations of *HIP1* contributed to the neurological abnormalities in their cases. *HIP1* is included in the triplicated segment in our case. Our case shares the neurodevelopmental phenotype attributed to copy number variations of *HIP1* and further supports this association.

3114F

A *de novo* 9.0 Mb Deletion at Chromosome Region 10q21.3-q22.3 associated with severe psychomotor Delay and mild Noonan Phenotype: a Case Report. *S. Baffini[1], G. Scarselli[2], L. Castiglia[3], M. Vinci[3], S. Amata[3], M. Fichera[3,5], J.D. Barp[4], E. Gambineri[4], E. Chiappini[4], M.L. Giovannucci Uzielli[1,4].* 1) Genetic Science, Florence, Italy; 2) Neuroscience Department, University of Florence, Italy; 3) Laboratory of Genetic Diagnosis, IRCCS Maria Santissima, Troina (Enna), Italy; 4) Science of Mother and Child Health Department, University of Florence, Italy; 5) Medical Genetics, University of Catania, Italy.

We report a clinical and molecular study in an 8-year-old child, observed by a periodical follow-up since birth, with severe intellectual disabilities, short stature, mild facial dysmorphisms, reminiscent of Noonan phenotype, thinning and sparse hair, and mild hearing loss. Q- and G- banding karyotyping was normal. PTPN11 gene screening did not reveal any mutations. Utilizing an oligonucleotide array CGH (Agilent 60K) we identified a genomic, interstitial, microdeletion at 10q21.3-q22.3: the deleted segment is approximately 9.0 Mb, and involves a contiguous region of single copy loss of over 23 genes, including NODAL and C10orf11 (http://genome.ucsc.edu). The breakpoints were at 70,288,248 and 79,331,739 (GRCh37). Deletions of the long arm of chromosome 10 are very rare: to our knowledge, the two first cases were only cytogenetically analyzed and associated with abnormal clinical phenotype [46,XX, del (10)(q21.2q22.1)(Glower et al., 1987) and 46,XY,del (10)(q22.1q22.3)(Cook et al. 1999)]. Three further cases [46,XX,del(10)(q21.3q22.2), 46,XY,del(10)(q22.2q22.3) and 46,XX,del(10)(q22.2q22.3)] were also studied at the molecular level (Tzschach et al. 2010). Mutations of NODAL gene, a member of TGF-b gene family, have been associated with heterotaxia and cardiovascular problems. NODAL gene is involved in Activin/Nodal signaling that is interconnected to the Ras-pathway. Moreover, Rap-2, a member of the Ras GTPase family, directs internalized Activin/Nodal receptors into a recycling pathway, thereby preventing their degradation and maintaining their levels on the cell surface (Choi et al. 2008). Tzschach and colleagues (2010) hypothesized a correlation between C10orf11 gene haploinsufficiency and mental retardation. Further study is needed in order to evaluate whether the haploinsufficiency of NODAL and C10orf11 and possibly other genes within the deleted 10q region can be considered responsible for our patient's clinical phenotype.

3115W

A first reported case of a microdeletion in 8q22.22q23 in Colombia. Phenotypic and genotypic correlation. *P. Paez[1,2], S. Perdomo[1], X. Rojas[2].* 1) Dept Cundinamarca, Universidad el Bosque, IINGM Facultad de Medicina,Bogota, Colombia; 2) Hospital Universitario Clinica San Rafael, Bogotá, Colombia.

We present a Colombian seven year old girl referred to a clinical genetics setting because of developmental delay, microcephalia, epilepsy, severe short stature, impaired speech, delayed attention, facial dismorphism, neurosensorial deafness and skeletal abnormalities. She is the second of two born children of healthy non consanguineous parents. She was born at 34 weeks due to intrauterine growth restriction and oligohydramnios. She presented severe short stature (< 3DS), microcephalia, triangular palpebral ptosis, bilateral blefarophimosis, sparse eyebrows, flat philtrum, thin upper lips, low set ears, small hands, short thumbs, bilateral fifth finger clinodactyly. Cognitive impairment, distractibility and poor attention. Normal karyotype: 46,XX[25], Normal brain images. Array CGH (comparative genome hybridization) suggested a minimal/maximal deletion of 5.238 Mb and 5.374Mb respectively at 8q22.2q22.3. Deletion of this region has been associated with a clinical recognizable condition described worldwide in only five patients. This region comprises at least 25 clinical relevant genes including the KCNS2 gen (which could explain epilepsy in this case). Ptosis and blepharofimosis are very frequent signs in all reported patients including our patient. There is no complete molecular explanation for these two signs but there is a critical zone related to blepharofimosis (NMLF) and a candidate locus related to ptosis(ZFHX4) which are out of the zone affected in this patient but that could play a positional effect on the deleted region. This is the first report of this interstitial deletion in Latin-America. This patient shares the majority of the features described in the previous 5 patients except for the neurosensorial deafness not previously described. This report is an example of how high resolution microarray technology facilitates the detection of submicroscopic chromosome alterations in the context of clinical genetics settings in Latin-America.

3116F

A microdeletion of about 831 kb in a young woman causes a loss of more than 30 genes on the long arm of chromosome 17 including a copy of the BRCA1-gene. Implications for genetic counselling and clinical management. *P.M. Kroisel, J.B. Geigl, M. Mach, C. Pischler, E. Vallant, M.R. Speicher, K. Wagner.* Institute of Human Genetics, Medical University of Graz, Graz, Austria.

By CGH analysis using a 60k Agilent oligonucleotide array a de novo microdeletion was identified, leading to a haploinsufficiency of the important tumour suppressor gene BRCA1. These unintended findings in the patient were obtained because cytogenetic results did not provide any clue for presence of her congenital moderate mental impairment and dysmorphic facial features which on the other hand suggested that a small genomic imbalance might be responsible which finally could be confirmed. Besides this dn microdeletion (chr17:40869151–41700815) also a microduplication of about 332 kb at 1q25.1 (chr1:172199642–172532314) inherited from her phenotypically normal mother was found. Because of that latter fact and since the microduplication only affects 4 genes without any known important function, we assume that the microduplication is presumably without any additional clinical relevance. The observations made in this particular case certainly demonstrate again that accidental or coincidental findings in specific subsets of patients should seriously be considered and councelling has to be performed accordingly. The list of well defined inherited cancer predisposition syndromes, which can be attributed to a hereditary susceptibility and having far reaching implications for all family members, is steadily growing. De novo mutations or even microdeletions of such tumour suppressor genes are rather rare events, which for different gene loci obviously will show a remarkable variable probability. We recently identified the first patient with a contiguous gene syndrome due to microdeletion of 17p that also caused a loss of the entire TP53 gene as well as several flanking genes in a mentally retarded person. The current patient is now 28 years of age. Her length is 153.5 cm and weight is 38 kg. A microcephaly with an OFC of 51 cm was noticed and distinct dysmorphic facial features like ear anomalies and bilateral ptosis were recognized. No comparable microdeletions also causing a loss of the entire BRCA1-gene as well as up- and downstream flanking genes is reported thus far or included in databases like ISCA or Decipher. Because of the achieved diagnosis the patient and her family were already subjected to an additional extented genetic councelling with information about results and recommenations and specialists in breast cancer screening and prevention will be included in continued medical surveillance and maintenance as requested.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3117W

A familial atypical Williams-Beuren deletion encopassing ELN and part of LIMK1 gene. *B. Demeer[1], F.M Caron[2], S. Chenichene[2], G. Morin[1], A. Receveur[3], H. Copin[3], A. De Broca[4], O. Godefroy[5], D. Sanlaville[6], M. Mathieu[1], P. Saugier-Veber[7].* 1) Genetics Dept, Hopital nord, CHU, Amiens, France; 2) Pediatrics cardiology- Pôle Pédiatrie - CHU Amiens, France; 3) Cytogenetics - Centre de Biologie de la Reproduction - CHU Amiens, France; 4) Neuropediatrics - Pôle Pédiatrie - CHU Amiens, France; 5) Neurology- CHU Amiens, France; 6) Cytogenetics- Centre de Biologie et Pathologie Est- Hospices civils de Lyon, France; 7) Molecular Genetics- CHU Rouen, France.

Williams-Beuren syndrome (WBS) is a multisystemic developmental disorder caused by hemizygous deletion of a segment on chromosome 7q11.23, spanning 1.55 Mb or 1.8 Mb, and encoding 28 genes. A few individuals have smaller deletions of the region. Their phenotypic features varies from supravalvular aortic stenosis syndrome (OMIM 185500) to classic Williams-Beuren Syndrome. We report a family, a mother and 3 sibs, with history of dominant supravalvular stenosis. The youngest daughter was diagnosed at birth with severe supravalvular aortic and pulmonary stenosis, requiring surgery, and facial dysmorphism. Caryotype and FISH of the 7q11.23 region was negative. Her two brothers present cardiovascular abnormalities, with for the youngest, aged 5 years, supravalvular aortic stenosis, involvement of the pulmonary branch, and mesenteric, coeliac and renal stenosis, and the eldest, aged 10 years, supravalvular pulmonary and aortic stenosis. They both had surgery for bilateral inguinal hernia, no classical WBS facial dysmorphism when seen, and impulsivity and short attention span (ADHD). Mother have supravalvular aortic stenosis, facial dysmorphism, and history of school difficulties, ADHD and still show strong anxiety. MLPA analysis of the 7q11.23 region found an atypical deletion, confirmed by array-CGH (Agilent 244k), of the entire ELN gene and part of the LMK1 gene. Although the loss of an ELN allele produces the cardiovascular pathology of WBS, the phenotypic consequences of losing other alleles within the WBS chromosome region are much less clear. More precise phenotypical data will be shown, and compare to the literature.

3118F

12q21 deletion syndrome with intellectual disability and facial dysmorphism. *A. Matsumoto, T. Yamagata, Y. Nozaki, E. F. Jimbo, M. Y. Momoi.* Jichi Medical University, Shimotsukeshi, Tochigi, Japan.

Interstitial deletion of 12q21 was reported in four patients, who presented intellectual disability (ID) and facial dysmorphisms. We report a patient with 12q21 deletion and tried to nominate these patients as 12q21 deletion syndrome. (Patient report) The patient was a 18-month-old boy. His father had karyotype of mos 46, XY/46, XY, inv(12)(q12q21) of 74:6 cells without abnormal phenotype. The patient's growth parameters were normal. At 18 months of age, the patient had ID (DQ 46) without meaningful words and with mild spastic diplesia. Other phenotypes of the patient included prominent forehead, hypotelorism, short upturned nose, long philtrum, small mandible, high-arched palate, low set ear, sparse hair and internal strabismus. Cranial magnetic resonance imaging (MRI) showed ventriculomegaly and mild hypoplasia of the corpus callosum. (Result) The patient's karyotype was 46,XY, del(12)(q21.2q21.33)inv(12)(q13.1q21.2). On his array CGH analysis using Agilent Human genome CGH 180K, a deletion at 12q21.2-q21.33(77. 2Mb ~91.2Mb)of 14 Mb was detected. (Discussion) Concordant clinical features of 12q21 deletion shared by previously reported four patients and our patient were ID, low set ear, sparse hair, prominent forehead and hyper- or hypotelorism. Other phenotypes observed in three or four patients were short upturned nose, small mandible, and cutaneous findings such as hyperkeratotic papular eruption and atopic dermatitis. On neuroimaging, ventriculomegaly was detected in three patients, and delayed myelinaization and mild hypoplasia of the corpus callosum was detected in one patient each. These findings are considered to be the phenotype of 12q21 deletion syndrome. The deletion in our patient was the shortest in 5 patients, that suggested the presence of responsible genes for the concordant feature should be localized in this region. Several important genes relating to neuronal development and function such as *ZDHHC17*, *NAV3*, *SYT1*, *PAWR*, *LIN7A*, *ALX1*, *DUSP6* and *ATP2B1* are localized at the deleted region. Further analysis for the responsible genes for 12q21 deletion syndrome should be interesting.

3119W

Autism, mild facial dysmorphism, abnormal EEG, sleep problems and language impairment in interstitial duplication 15q11-q13 syndrome. *N. Urraca[1], C. Cleary[2], E. Pivnick[3], K. McVicar[3], V. Brewer[3], R. Thibert[4], N.C. Schanen[5], M.C. Esmer[6], D. Lamport[7], L.T. Reiter[1,3].* 1) Neurology, UTHSC, Memphis, TN; 2) Speech and Language Pathology, University of Memphis, Memphis, TN; 3) Pediatrics, UTHSC, Memphis, TN; 4) Neurology, Mass General Hospital, Boston, MA; 5) Human Genetics Research Laboratory, Nemours Biomedical Research, Wilmington, DE; 6) Genetics, Hospital Central “Dr. Ignacio Morones Prieto”, SLP, MEX; 7) Psychology, University of South Alabama, Mobile, AL.

Chromosomal copy number variants (CNV) are the most common genetic causes of autism. As many as 1–3% of autistic cases are the result of duplications of the 15q11-q13 region. Advances in molecular detection of CNVs have had a significant impact on the detection of 15q duplication cases. Chromosome 15 has a number of low copy repeats that predispose the region to deletion/duplication events. This region has a cluster of genes preferentially expressed from one parental allele in the nervous system that clearly influence phenotypic outcome in deletion cases, but have been difficult to analyze in individuals with interstitial 15q duplications. Most cases that present clinically are de novo and of maternal origin, but paternal duplications can present phenotype as well. We performed an in depth phenotype/genotype analysis of individuals with interstitial 15q duplications to determine if maternal duplication is required for autism diagnosis. We used dysmorphology, neurological examination, neuropsychological, language, sleep, neurophysiology and ASD diagnostic tools for phenotypic analysis. Methylation Sensitive High Resolution Melting analysis of the maternally methylated SNRPN was used to determine the parent of origin of the duplication. 14 int dup(15) subjects were selected for analysis: 10 maternal and 4 paternal cases. Eleven are de novo. Six cases are PWS/AS Class I and 8 are Class II. The most frequent facial features found were long philtrum with a short bulbous nose. 60% also had hypotonia. 8/14 subjects had excessive 18–22 Hz beta activities on EEG throughout the waking record regardless of the parental origin of the duplication or the presence/absence of ASD. All subjects with maternal and 2/4 with paternal duplications scored on ASD for ADOS/ADI-R analysis. Most subjects had a low-moderate adaptive functioning score on the Vineland II test with no differences among groups. All subjects performed below age corrected average for receptive language. On average all subjects had reported-sleep problems, but in subjects with paternal duplications sleep problems were more severe, though they have better sleep habits. In conclusion, the majority of int dup(15) cases found were maternal in origin, most likely due to the coincidence of ASD with maternal duplication, the size of the duplication did not correlate with the severity of the phenotype, and subjects had mild facial anomalies, autism, language impairment, sleep problems and unusual EEG variant.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3120F

Mucolipidosis III gamma, tuberous sclerosis, and polycystic kidney disease caused by a contiguous gene deletion on chromosome 16p13.3. *J. Barea[1], L. Bird[1,2].* 1) Genetics Dept, UCSD, San Diego, CA; 2) Dysmorphology/Genetics, Rady Children's Specialists of San Diego, San Diego, CA.

We present the first known report of a patient with Tuberous Sclerosis 2, Polycystic Kidney Disease 1, and Mucolipidosis III gamma. This patient was noted to have clubfeet and developmental dysplasia of the hips from infancy. At 9 months old, she had a prolonged episode of unresponsiveness with an abnormal EEG and was diagnosed with a seizure disorder. At age 4 the patient was noted to have facial angiofibromas and hypopigmented macules, leading to the diagnosis of Tuberous Sclerosis. The patient later had a brain MRI that showed cortical tubers and subependymal nodules, and a renal ultrasound that showed renal cysts. However, the patient developed a variety of worsening orthopedic problems; including genu valgum, stiffness of finger joints, decreased shoulder mobility, scoliosis, pain on back and upper legs, and decreased exercise tolerance. She also had abnormal radiologic findings including bony deformities involving the wrists and distal forearms; thoracolumbar scoliosis with malformed and irregular appearing vertebral bodies, with spondylolisthesis of L5 on S1; and bilateral hip dysplasia with femoral head irregularities. Microarray analysis was done showing a 1.9 Mb deletion of chromosome 16p13.3. This deletion contains at least 96 genes, including TSC2 and PKD1. Other genes in the region were examined for their possible relationship to the orthopedic symptoms of the patient. The GNPTG gene, which encodes the gamma subunit of UDP-N-acetylglucosamine: lysosomal hydrolase N-acetylglucosamine 1-phosphotransferase enzyme, is within the deleted interval. Mutations in the GNPTG gene cause Mucolipidosis III gamma, and the patient's clinical and radiographic presentation was compatible with this condition. The diagnosis was confirmed by lysosomal hydrolase assay, which showed very high plasma enzyme levels and normal enzyme levels in leukocytes. ML III can be due to deficiency of the alpha/beta subunit encoded by the GNPTAB gene located on chromosome 12q23.3 or due to deficiency of the gamma subunit encoded by GNPTG. Given that ML III gamma is an autosomal recessive condition, a point mutation in GNPTG on the intact chromosome 16 was assumed and GNPTG sequencing revealed Gly126Ser mutation in a highly conserved residue. Further understanding of the differences between ML III gamma and ML III alpha/beta will be enhanced by determining the subclass of ML III patients.

3121W

GH deficiency and epilepsy in a patient with a 19q13.3 duplication encompassing the pth2, kcnc3 and kcna7 genes. *M. Vincent[1], C. Jeandel[2], J. Puechberty[1,3], C. Coubes[1], F. Dallavale[2], F. Rivier[4], A. Schneider[1,3], M. Girard[3], N. Ruiz[5], M. Tournaire[3], I. Touitou[5], G. Lefort[1,3], P. Sarda[1], D. Genevieve[1,6].* 1) Génétique médicale, CHU de Montpellier, Montpellier, Languedoc Roussillon, France; 2) Département de pédiatrie spécialisée endocrinopédiatrie, Hôpital Arnaud de Villeneuve, CHRU Montpellier, France; 3) Plate-forme Puce à ADN, Laboratoire de Génétique Chromosomique, CHRU Montpellier, France; 4) Service de Neuropédiatrie, Hôpital Gui de Chauliac, CHRU Montpellier, France; 5) Unité médicale des maladies autoinflammatoires, Hôpital Arnaud de Villeneuve, CHRU Montpellier, France; 6) Unité Inserm U844, Institut des Neurosciences de Montpellier, France.

GH deficiency of genetic origin can be observed either as an isolated feature due to mutation in the GHRH-R, GH1, GHS-R and BTK genes, or in a malformation context which is sometimes due to a genetic syndrome. The aim of this study is to report a patient with GH deficiency and PTH2 duplication and to discuss the potential role of this gene as a novel genetic etiology of GH deficiency. A young French boy presented with short stature, seizures, intellectual disability and peculiar facial features. Endocrinological tests performed because of moderate short stature revealed GH deficiency without failure in the rest of the pituitary axis, and a good response to GH treatment. Genetic investigations including SNP array showed a 6.8 Mb duplication on the long arm of chromosome 19 (19q13.33). We discuss the co-relation between the features observed in our patient and the genetic content of this chromosomal region. GH deficiency could be due to the PTH2 duplication and seizures due to KCNC3 and KCNA7 duplication. The other patients reported in the literature with distal 19q duplication encompassing the PTH2, KCNC3 and KCNA7 genes also presented with short stature and epilepsy. Based on data from the literature, in this report GH deficiency could be due to increased gene dosage of PTH2, resulting in an excess of PTH2 peptide, which plays a role in the regulation of the hypothalamic-pituitary-adrenal axis by down-regulation of GH secretion. This report suggests that PTH2 is a novel gene involved in GH deficiency in humans.

3122F

A large deletion confined to *COL4A5* causes Alport syndrome and diffuse leiomyomatosis. *M.JN. Sá[1,2], R. Sousa[3], FT. Costa[4], F. Carvalho[5], S. Alves[1], F. Carvalho[1], JP. Oliveira[1,2].* 1) Department of Genetics, Faculty of Medicine, University of Porto, Porto, Portugal; 2) Research and Development Unit of Nephrology (FCT-725), Faculty of Medicine, University of Porto, Porto, Portugal; 3) Department of Gastroenterology, Hospital Garcia de Orta, Almada, Portugal; 4) Department of Nephrology, Hospital Garcia de Orta, Almada, Portugal; 5) Department of Nephrology, Hospital Curry Cabral, Lisboa, Portugal.

Alport syndrome and diffuse leiomyomatosis (MIM#308940; ATS-DL) is a contiguous gene deletion syndrome involving the *COL4A5* and *COL4A6* genes. These genes are localized head-to-head in close proximity on chromosome Xq22.3 and share a common promoter region. Previous studies have suggested that the minimal deletion region associated to ATS-DL includes exon 1 of COL4A5 and exons 1, 1' and 2 of *COL4A6*. Herein, we report the molecular genetic analysis of an affected male with ATS-DL in whom a deletion confined to *COL4A5* was identified. The proband is a 25 years old male with Alport syndrome diagnosed at the age of five years, on the basis of typical electron microscopy findings on a kidney biopsy. Hearing loss became apparent at the age of seven years and anterior lenticonus was diagnosed at the age of 18 years. Renal replacement therapy by hemodialysis was started at the age of 20 years and the patient eventually received a cadaveric kidney transplant six months later. Esophageal leiomyomatosis was diagnosed at the age of 25 years on the investigation of persistent upper gastrointestinal symptoms, including dysphagia. Direct sequencing of *COL4A5* gene showed a large hemizygous deletion, encompassing exons 2 to 51. Direct sequencing of the first three exons of *COL4A6*gene was normal. These results were confirmed by multiplex ligation-dependent probe amplification (MLPA Alport syndrome kit, SALSA P191/P192). His mother and a sister, also affected with ATS-DL, are heterozygous for the same deletion. To our knowledge, this is the first time a family with ATS-DL is shown to have a large deletion confined to *COL4A5*. This observation suggests that deletion of exons 1, 1' and 2 of *COL4A6* is not required for the development of diffuse leiomyomatosis in association with Alport syndrome. Further investigation is needed to confirm the absence of mutations affecting the common promoter region of the two genes and to determine the extent of the deletion towards the telomere, to substantiate this remarkable finding.

3123W

Genetic study of a family segregating Waardenburg-Shah syndrome. *L. Cui[1], H.M. Wong[1], J. Zhu[1], F. de Almeida M[2], P.H. Tam[1], M. Garcia-Barcelo[1].* 1) The University of Hong Kong, Hong Kong, China; 2) Faculdade de Medicina de Ciências Medicas de Minas Gerais, Brazil.

Waardenburg-Shah syndrome (WS4, MIM_277580) is a congenital developmental disorder characterized by pigmentary abnormalities of the skin, eyes and hair, sensorineural deafness and intestinal aganglionosis (HSCR; Hirschsprung disease). Mutations in EDN3, EDNRB, or SOX10 account for approximately 65–85% of the WS4 patients. The disorder is clinically and genetically heterogeneous with either autosomal recessive or autosomal dominant inheritance with incomplete penetrance. EDN3, EDNRB, and SOX10 were sequenced in a three-generation family (14 individuals; 3 members affected with HSCR and 1 with WS4). A novel heterozygous missense mutation in the EDNRB (chromosome 13) was identified in four affected and three unaffected family members. The mutation changes the translation initiation codon Met1 of the EDNRB receptor isoforms 1 and 2 (M1V; predicted damaging) or the residue Met91 of the EDNRB isoform 3 (M91V; predicted benign). Importantly, the three EDNRB isoforms are expressed in the gut of newborns. Functional analyses coupled with confocal microscopy revealed the presence of a shorter EDNRB isoform 1 (M1V) as a result of alternative translation initiation site usage (Met46). Small quantities of this shorter protein were observed in the cellular membrane. Both EDNRB isoform 3 wild-type and the mutated (M91V) counterpart were only found in the cytosol, apparently not being implicated in the receptor-ligand relationship necessary for the differentiation of the neural crest cells from where both enteric neurons and melanocytes derived. Reduced amounts of a shorter form of the EDNRB receptor can indeed account for the disease phenotype. Yet, given the reduced penetrance of M1V, additional loci contributing to this disorder may exist in this family. To identify those, family members were genotyped using Illumina Human Omni2.5-quad BeadChip. After quality control and SNP pruning the Merlin software was used for parametric and non-parametric linkage. A susceptibility locus on chromosome 4q13.3-q24 (28.35 cM) was identified by both nonparametric and parametric linkage analysis. Haplotype analysis refined the linkage region to a 27.76 cM interval (between rs13146209 and rs17033669) encompassing two biologically plausible genes: MAPK10 and PPP3CA which are currently being sequenced in these individuals.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3124F

Founder effect of Spinocerebellar ataxia type 7 (SCA7) in a Mexican population. *Y. Tapia-Guerrero[1], O. Hernández-Hernández[1], M. Maldonado-Rodríguez[1], C. Cerecedo-Zapata[1], N. Leyva[1], L. Velázquez-Pérez[2], B. Cisneros[3], J.J. Magaña[1].* 1) Genetics, National Rehabilitation Institute, Mexico City, Mexico City, Mexico; 2) Center for the Research and Rehabilitation of the Hereditary Ataxias (CIRAH), Holguín, Cuba; 3) Genetics and Molecular Biology, CINVESTAV-IPN, 07360 Mexico City, Mexico.

Spinocerebellar ataxia type 7 (SCA7) is a polyglutamine (polyQ) neurodegenerative disease characterized by progressive cerebellar ataxia and retinal degeneration with ophthalmoplegia, hyperreflexia, sensory loss, dysarthria, and dysphagia. SCA7 is caused by expansion of the CAG trinucleotide repeats within the coding tract of the ATXN7 gene, which encodes ataxin-7, a subunit of protein complex involved in transcriptional regulation. This repeat is polymorphic in normal individuals with alleles ranging from 7 to 36 repeats. Alleles exceeding a threshold of approximately 37 repeats and reaching up to a number of 460 CAGs result in disease. Due to that SCA7 shares practically the same symptomatology with the other ataxias, its accurate diagnosis is based on identification of the genetic mutation that produces it. In this study, we analyzed a group of 62 patients with autosomic dominant cerebellar ataxias characteristics from the central region of Veracruz State, Mexico, applying a multiplex fluorescent PCR method (for SCA1, 2, 3, 6 and 7) in combination with capillary electrophoresis analysis of the amplified products. Furthermore, we analyze the variability at the SCA7 gene polymorphism in 300 unrelated Mexican healthy subjects. We identified SCA7 expanded alleles in 52 out of 61 patients (87%) and 9 patients with SCA2 (13%). We found twenty SCA7 mutant alleles, ranging from 37 to 72 repeats. Preliminary data indicated that in a region of 428.27 miles of Mexico southeast, unlike in other populations studied, SCA7 is by far the most common of all genetically verified dominant ataxias, suggesting the existence of a founder effect for the SCA7. Specifically, we found a preliminar prevalence in a Mexican population (Tlaltetela) of approximately 817 cases per 100,000 people. To investigate this possibility we performed an extensive haplotype study (3145G/A/)/D3S1287/D3S3635/D3S1228) including all SCA7 families available in these populations. In the general Mexican healthy population a total of ten alleles were found, ranging from 7 to 16 repeats. The most common allele had 10 repeats (59.5%) followed by the 12 and 11 repeats alleles (17.2% and 11.94% respectively). Alleles greater than 16 repeats had a frequency of 0.17. We conclude from this study that the relatively high frequency of SCA7 in this region could be the result of a founder effect. Until now this is one of the largest series of cases in specific geographical region around the world.

3125W

Study of GJB2, GJB6 and MT-RNR1m.1555A>G in Mexican hearing loss patients. *M. Arenas-Sordo[1], I. Menendez[2], E. Hernández-Zamora[1], A. Simarci[2], D. Gutierrez[1], P. Murphy[1], X. Leyva[1], F. Huesca[1], J. Dominguez-Aburto[1], M. Tekin[2].* 1) Genetics, INR, Av.Mexico-Xochimilco 289. Mexico DF, Distrito Federal, Mexico, 14389; 2) Hussman Institute of Human Genomics. University of Miami. 1501NW 10th Ave. Miami. Florida 33136.

Objective: The aim of this study was to elucidate the involvement of three common deafness genes in 76 unrelated Mexican probands with prelingual non-syndromic hearing loss. Methods: We screened GJB2, GJB6, and mitochondrial MT-RNR1 genes for mutation analysis in 76 (71 simplex; 5 multiplex) unrelated Mexican probands with prelingual non-syndromic hearing loss from the Department of Genetics at Instituto Nacional de Rehabilitacion in Mexico City. Results: We identified 8 different mutations and two polymorphic variants in GJB2. Eight cases were found to have bi-allelic mutations in GJB2 (10.6% prevalence) and six were found to have only one GJB2 mutation. c.35delG was the most frequent pathological GJB2 allele. The p.V27I polymorphism was frequently detected (allele frequency: 24%). The evaluation of the hearing severity showed that individuals with bi-allelic GJB2 mutations were affected mostly profound. DFNB1 carriers were severely or profoundly affected and non DFNB1 patients were among the three groups (moderate, severe and profoundly affected hearing loss). Asymmetric hearing loss were absent in the DFNB1 group, but it was identified mostly in non DFNB1 patients. Two screened GJB6 deletions and m.1555A>G mutation were not detected. Conclusions: This study showed that GJB2 mutations were an important cause of prelingual deafness in Mexican population.

3126F

A family with an OPA1 mutation and associated hearing loss. *C. Williams[1], H. Stalker[1], S. Schimpf-Linzenbold[2], B. Wissinger[2].* 1) Philips Unit, Dept Pediatrics, Genetics and Metabolism, Univ Florida, Gainesville, FL; 2) Molecular Genetics Laboratory, Centre for Ophthalmology, University of Tuebingen, Germany.

Mutations of OPA1 are the most common cause for familial dominant optic atrophy. This gene encodes a dynamin-like large GTPase that localizes to mitochondria and plays a crucial role in mitochondrial dynamics and cristae junction stabilization. We identified a novel mutation in exon 12 (Pro400Ser) of OPA1 that segregates through at least three generations. The 5-year-old proband had severe visual impairment with optic atrophy (OA) since infancy and a tremulous movement disorder without weakness. She developed sensorineural hearing loss (SNHL) at age 10 years requiring cochlear implants at age 14. An affected sibling with OA is now 19 years old but has no apparent hearing loss. The affected mother has OA with severe vision loss and has severe SNHL with onset at age 15 years (now with cochlear implants). The affected maternal grandmother has mild SNHL with onset at age 40: there is variable expression of SNHL in the kindred. Most individuals with OPA1 mutations (typically truncating or splice type) do not have associated hearing loss. The multisystem features in this family are therefore possibly related to a dominant-negative effect of the missense mutation. Exon 12 is located within the GTPase domain of OPA1 and previously reported families with Arg445His mutations have had OA and SNHL. The hearing loss in those families was attributed to auditory neuropathy (Amati-Bonneau et al., Ann Neurol 2005; 58:958–963).

3127W

Cerebrocostomandibular syndrome: clinical features of three new cases and preliminary analysis of exome sequencing data. *D.C. Lynch[1], E. Lemire[2], B. Chodirker[3], A.M. Innes[1], J.S. Parboosingh[1], F.P. Bernier[1] FORGE Canada Consortium.* 1) Department of Medical Genetics, Alberta Children's Hosp, Calgary, Alberta, Canada; 2) Department of Pediatrics, University of Saskatchewan, Saskatoon, Canada; 3) Departments of Biochemistry and Medical Genetics and Pediatrics, and Child Health, University of Manitoba, Winnipeg, Manitoba, Canada.

CCMS (OMIM 117650) is characterized by posterior rib-gap defects, reduction in rib number, and micrognathia. CCMS is clinically variable, ranging from mildly affected adults who present after the birth of an affected child, to presentation in infancy with absent ribs, respiratory distress, severe microcephaly, and other anomalies. Approximately 40% of affected individuals die within the first year of life due to respiratory insufficiency. Here we describe four new cases of CCMS, including a 13 yo male and his affected brother who died in infancy, and two sporadic cases (a 17 yo male and a 15 yo male). The mode of inheritance is unclear in this cohort, especially given previous reports of both autosomal dominant and autosomal recessive inheritance. All cases had typical features of CCMS include rib anomalies and micrognathia. We have sequenced the exomes of the three surviving patients patients and four unaffected parents or siblings. No variants in a common gene or gene family have been identified, raising the possibility of genetic heterogeneity in CCMS. It is also possible that the causative mutation has been missed in some or all patients due to the technical limitations of exome sequencing, or because CCMS is caused by mutations in non-coding regions. As autosomal recessive mutations in COG1 have been reported in CCMS-like patients, COG1 mutations were ruled out in this cohort through exome data analysis.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3128F

Dandy Walker Malformation in a boy with Ellis-van Creveld dysplasia. *J. Davis[1], E. Carter[2], P. Brill[3], C. Raggio[2].* 1) Dept Pediatrics, NY-PH Hosps - Weill Cornell, New York, NY; 2) Ctr. for Skeletal Dysplasias, Hosp for Special Surgery, New York, NY; 3) Dept Radiology, NY-PH Hosps - Weill Cornell, New York, NY.

Ellis-van Creveld syndrome (EvC; MIM:225500) is an autosomal recessive skeletal dysplasia characterized by short limb dwarfism, narrow chest, post-axial/axial hexadactyly, dysplastic teeth and hypoplastic nails. Approximately 60% of patients have congenital heart disease. Intellectual development is usually normal. Mutation of EVC and/or EVC2 (mapped to chromosome 4p16) has been detected in 2/3 of cases. Three patients with EvC have been reported to also have a Dandy Walker malformation (DVM), characterized by hypoplasia of the cerebellar vermis and cystic dilatation of the 4th ventricle. A 9+8yr-old male was referred to our service for management and evaluation of progressive genu valgum. He is the product of a spontaneous fraternal twin pregnancy discordant for EvC. His parents are 1st cousins of Mexican ancestry. Family history revealed no one else with similar findings. Following NSVD at 36wks, Twin A weighed 4lbs11oz. She has followed a normal course of growth and development. Twin B (our patient) weighed 4lbs10oz. He had natal teeth, a narrow thorax, and a heart murmur. Echocardiogram showed a common atrium with total anomalous pulmonary venous return. He had respiratory difficulties, epispadias, and short limbs with hexadactly of each hand. Congenital hypothyroidism and seizures have resolved. He has GE reflux and is fed through a G-tube. At 9+8yrs his height (106cm) plots at the 5th percentile on EvC growth charts. His head circumference (55cm) plots at the 98th percentile on growth charts developed for males his age. He has dysmorphic facies, posterior bossing, and two hair whorls. His teeth are dysplastic. His arms and legs are short, with short, broad hands and feet and dysplastic nails. There is polydactlyly involving the 2nd and 3rd digits of both hands. He has developmental delay. Molecular genetic studies are pending. Radiographs are consistent with EvC. MRI of the brain at 2mos showed enlarged ventricles with cortical atrophy. Follow-up MRI at 9+8yrs showed a large 4th ventricle, a small/hypoplastic cerebellar vermis and corpus callosum dyspgenesis (DVM). This is the 4th reported patient with EvC and DWM. In some cases, consanguinity may be responsible for two, separate recessive traits. However, it is tempting to speculate that DVM is a rare manifestation of EvC, especially given the demonstration of Evc and Evc2 proteins localizing to the basal bodies of primary cilia in mice, with involvement in moderating hedgehog signaling pathways.

3129W

***NEK1* and *DYNC2H1* are both involved in short rib polydactyly Majewski type but not in Beemer Langer cases.** *J. El Hokayem[1], C. Huber[1], A. Couve[1], J. Aziza[2], G. Baujat[1], R. Bouvier[3], DP. Cavalcanti[4], FA. Collins[5], MP. Cordier[6], AL. Delezoide[7], M. Gonzales[8], D. Johnson[9], M. Le Merrer[1], A. Levy-Mozziconacci[10], P. Loget[11], D. Martin-Coignard[12], J. Martinovic[13], GR. Mortier[14], MJ. Perez[15], J. Roume[16], G. Scarano[17], A. Munnich[1], V. Cormier-Daire[1].* 1) INSERM U781, Hôpital Necker, Paris, Paris, France; 2) Laboratoire d'anatomie et cytologie pathologiques, Hôpital Purpan, Toulouse, France; 3) Centre de pathologie EST, Hôpital Louis Pradel, Hôpital Pierre Wertheimer, Hôpital Femme mère enfant, Lyon, France; 4) Programa de Genética Perinatal, Departamento de Genética Médica, FCM, UNICAMP, Campinas, Sao Paulo, Brazil; 5) Western Sydney Genetics Program, Department of Clinical Genetics, Children's, Hospital at Westmead, Sydney, Australia; 6) Service de Génétique, Groupement Hospitalier Est, HFME, Bron, France; 7) Service de Biologie de Développement, Université Paris Diderot, Hôpital, Robert Debré, AP-HP, Paris, France; 8) Service de Génétique et d'Embryologie Médicales, Hôpital Armand Trousseau, AP-HP, Université Pierre et Marie Curie, Paris VI, France; 9) Sheffield Clinical Genetics Service, Sheffield Children's NHS Foundations Trust, Western Bank, England; 10) Laboratoire de Biochimie et Biologie Moléculaire, Hôpital Nord, Marseille, France; 11) Centre hospitalier universitaire de Rennes, Service d'anatomie et cytologie pathologiques, Rennes, France; 12) Pôle de biopathologie, UF 3162, Centre hospitalier, Le Mans, France; 13) Unit of Fetal Pathology, Cerba Laboratory, Cergy Pontoise; 14) Department of Medical Genetics, Antwerp University Hospital and University of Antwerp, Antwerp, Belgium; 15) Département Génétique Médicale, Arnaud de Villeneuve Hospital, Montpellier, France; 16) Service de génétique médicale, centre hospitalier Poissy-Saint Germain, France; 17) Department of Medical Genetics, Sannio University, Benevento, Italy.

The lethal short rib polydactyly syndromes include four distinct conditions (SRP I-IV), characterised by autosomal recessive inheritance and short ribs, short limbs, polydactyly, and multiple anomalies of major organs. Among them, SRP type II (Majewski; MIM 263520) is characterised by short ovoid tibiae or tibial agenesis and is radiographically closely related to SRP type IV (Beemer-Langer; MIM 269860) which is distinguished by bowed radii and ulnae and relatively well tubulated tibiae. *NEK1* mutations have been recently identified in SRP type II. Double heterozygosity for mutations in both *NEK1* and *DYNC2H1* in one SRP type II case supported possible digenic diallelic inheritance. The aim of our study was to screen *DYNC2H1* and *NEK1* in thirteen SRP type II and seven SRP type IV cases. All cases were either terminated pregnancies (15–28 week gestation) or neonatal death (32–35 week gestation). Among the SRP type II cases, five belonged to consanguineous families from Lebanon, India (with recurrent sibs), Pakistan, and France and the eight remaining cases were from Madagascar, France, Brazil (with recurrent sibs), Vietnam, Haiti, and Belgium. Among the SRP type IV cases, two were consanguineous from Tunisia and Turkey, and five were non-consanguineous cases from France, Algeria, Italy and UK. We identified homozygous *NEK1* mutations in 5/13 SRP type II with 2/5 being missense mutations, 1/5 frameshift mutation and 1/5 nonsense mutation. We also identified nine novel compound heterozygous *DYNC2H1* mutations in 4/12 cases with 7/9 being missense mutations; 1/9 was a frameshift mutation and 1/9 a splicesite mutation. *DYNC2H1*mutations have been previously identified in Asphixyating Jeune dysplasia and SRP type III but no correlation between the nature or the location of the mutation and the severity of the phenotype could be established. Finally, *NEK1* and *DYNC2H1* were excluded in 3/12 SRP type II and in all SRP type IV cases. The main difference between the mutation positive SRP type II group and the mutation negative SRP type II group was the presence of holoprosencephaly and polymycrogyria in the mutation negative group. This study confirms that *NEK1* is one gene causing SRP type II but also reports *DYNC2H1* as another SRP type II disease gene expanding the phenotypic spectrum of *DYNC2H1* mutations. The exclusion of *NEK1* and *DYNC2H1* in 3/12 SRP type II and in all SRP type IV cases further support genetic heterogeneity.

3130F

Refsum Disease: Genotype-phenotype correlation of a novel PHYH mutation in a brazilian patient. *L.A.R. Gabriel[1], J. Chiang[3], R.C.C. Filho[2], L.G. Freitas[2], M.P. Avila[2].* 1) Ocular Genetics Department, UFG, Goiania, GO, Brazil; 2) Retina and Vitreous Department, UFG, GO, Brazil; 3) Director, Casey Eye Institute Molecular Diagnostic Laboratory, Portland, OR, USA.

Herein we present a case report of Refsum Disease (RD) in a male brazilian patient with the p.R329X:c.985C>T mutation in the PHYH gene. The patient entered the Department of Ocular Genetics of the UFG with a previous diagnosis of retinitis pigmentosa. After the clinical examination, metatarsal and metacarpal alterations were noted; the alcohol sniff test revealed a severe anosmia; and retinitis pigmentosa was confirmed. Thus RD was suspected and the phytanic acid high levels corroborated the clinical diagnosis. While waiting for the molecular diagnosis, the patient was referred to the Cardiology and Neurology Department where exams were normal. The genetic testing included PHYH and PEX7 sequencing, where PEX7 showed no mutations and PHYH showed the mutation above. Right now the patient is under a low phytanic acid diet.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3131W

Description of sclerocornea in a patient with Van den Ende Gupta syndrome and homozygous mutation in *SCARF2*. *M. Migliavacca[1], N. Sobreira[2], G. Antonialli[1], M. Moysés[3], M. Melaragno[3], D. Valle[2], D. Brunoni[1], A. Perez[1].* 1) Clinical Genetics, Federal University of São Paulo, São Paulo, São Paulo, Brazil; 2) Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD; 3) Laboratory of Cytogenetics, Genetics, Department of Morphology and Genetics - UNIFESP-São Paulo- S.P. - Brasil.

Van den Ende-Gupta Syndrome (VDEGS) (OMIM - 600920) is a rare autosomal recessive disease characterized by blepharophimosis, beaked nose, hypoplastic maxila, skeletal abnormalities (hooked clavicles, gracile ribs, bowed long bones) and normal psychomotor developmental. Using homozygosity mapping Anastasio et al. (2010) identified homozygous mutations in *SCARF2* in 4 VDEGS patients from three consanguineous Qatari families. Bedeschi at al. (2011) described a VDEGS patient with sclerocornea and cataracts, features not described previously, due to compound heterozygosity for the common 22q11.2 microdeletion and a hemizygous *SCARF2* splice site mutation. They suggested that the ocular abnormalities were caused by the 22q11.2 microdeletion since sclerocornea has been described in patients with DiGeorge syndrome. Here we report a 21 year-old male who presented with mild developmental delay but normal cognitive function. He had microcephaly, high and broad forehead, hypertelorism, short palpebral fissures, bilateral sclerocornea, nystagmus, strabismus, hypoplastic alae nase with prominent columela and depressed nasal tip, malar hypoplasia, prominent everted lips, prominent, low set and posterior angulated ears, pectus escavatum, arachnodactyly, bilateral adduction of his thumbs, camptodactyly, bilateral hallux valgus with long overlapping toes. His skeletal survey showed retrognathia, mild thickening of skull, sclerosis of mastoid cells, thoracic and lumbar scoliosis, gracile ribs, narrow iliac, slender long bones, mild bowing of femur and fibula, shortened ulna not articulated with the radius, long metatarsus, bilateral valgus deformity of the halluces and osteopenia. Brain MRI and echocardiogram were normal. Karyotype and SNP genotyping array were normal. To identify the cause of his malformations, we performed whole exome sequencing in the proband and his unaffected brother. We identified a homozygous 17 bp deletion in exon 4 of *SCARF2* in the proband that was not present in his unaffected brother and led us to the diagnosis of VDEGS. This patient has bilateral severe sclerocornea, a feature previously described in only one VDEGS patient and thought to be secondary to a 22q11.2 microdeletion. Our patient expands the VDEGS phenotype and implicates *SCARF2* as the causative gene of sclerocornea in these patients.

3132F

Two brothers with autosomal-recessive primary hypertrophic osteoarthropathy caused by homozygous deletion in HPGD gene: Neonatal findings and long-term follow-up. *B. Tüysüz[1], S. Yilmaz[2], K. Bilguvar[2], O. Kasapçopur[3], E. Gül[1], M. Günel[2].* 1) Department of Pediatric Genetics, Istanbul University, Cerrahpasa Medical School,, Istanbul, Turkey; 2) Department of Neurosurgery, Program on Neurogenetics, Yale School of Medicine, New Haven, Connecticut, USA; 3) Department of Pediatric Rhematology, Cerrahpaşa Medical School, Istanbul University,Istanbul-Turkey.

A 8-month-old boy who is the first child of first cousin parents was referred to our department due to coarse face. He also had large anterior fontanelle, high-arched palate, hypertrichosis, inguinal hernia and large hands and feet. Radiographs showed osteopenia, wormian bones, subperiosteal ossification and expanded diaphysis and metaphysis. The patient has been followed up to the age of 7.5 years. Re-evaluation of the patient at the age of two revealed pain, enlargement and hyperhidrosis of hands and feet. Methotrexate 7.5 mg once a week intravenously was started, resulting in decrease of swelling and pain. He gained walk at 19 months. At the age of 3.5, swelling of knee joints, digital clubbing and decreased joint mobility were developed. Cold induced swelling of hands and feet was noticed by parents. Mild palmoplantar hyperkeratosis was observed at 6 years of age. IQ was normal. The patient was diagnosed as autosomal recessive primary hypertrophic osteoarthropathy (PHO), is a very rare disorder characterized by delayed closure of the cranial sutures and fontanels, digital clubbing, arthropathy, and periostosis. Mutations in HPGD gene, which encodes the NAD-dependent 15-hydroxyprostaglandin dehydrogenase caused PHO type 1. The patient had 2 bp homozygous deletion in exon 3 (c.310_311delCT) of HPGD gene leading to frameshift, premature truncation (p.L104Afsx2) of the polypeptide. A prenatal diagnosis was rejected during the subsequent pregnancy by parents. The second boy was born with large hands and feet and also typical radiological findings of PHO. This is the first report of detailed neonatal findings and longitudinal observation of a patient with PHO.

3133W

Germline Mosaicism for a 12q24 Deletion Identifies Haploinsufficiency of *MED13L* as a Cause of Hypotonia and Moderate Developmental Delay. *R.C. Gallagher[1], S. Scrivner[1], K. Brown[1], A. Collins[2], M. Saenz[1].* 1) Clinical Genetics and Metabolism, Department of Pediatrics, University of Colorado Anschutz Medical Campus; 2) Neurology, Department of Pediatrics, University of Colorado Anschutz Medical Campus.

Purpose: To report the identification of two siblings with moderate developmental delay, dysmorphic features and hypotonia who share a deletion of the *MED13L* gene, initially reported in association with transposition of the great arteries and developmental delay, as further evidence for the role of haploinsufficiency of *MED13L* in intellectual disability. Background: *MED13L*, previously described as *PROSIT240/THRAP2*, was reported to be deleted or mutated in individuals with transposition of the great arteries (TGA) at a high frequency of 3%. The initial affected individual with TGA had a translocation that disrupted the gene. Other clinical features included intellectual disability, ataxia, and nearly absent speech (Muncke N et al, Circulation, 2003 108:2843). The role of mutations of *MED13L* in intellectual disability without cardiac defects is unknown. Case report: Two maternal half-siblings presented with similar features of moderate developmental delay, hypotonia and strabismus. One is non-verbal; the other has mild to severe, bilateral, high frequency sensorineural hearing loss, and is verbal. Each has mild craniofacial dysmorphisms, though these differ in the two. One had minimal, the other mild, ataxia. There is no known cardiac defect in either. One sibling had a brain MRI that showed mild hypomyelination of the posterior periventricular white matter. Both were found to have a 200 kb deletion at 12q24.21 that encompasses the *MED13L* gene. The deletion was not identified in the peripheral blood of the mother. Conclusion: Haploinsufficiency of *MED13L* is a cause of intellectual disability, and is not always associated with cardiac defects. Elucidation of the phenotypic spectrum of disorders that present as new dominants will be increasingly important as whole exome and whole genome strategies are pursued for individuals with intellectual disability.

3134F

Genetic analysis of the *GBA* gene in Japanese familial Parkinson's disease. *Y. Li[1], M. Funayama[1], T. Sekine[2], L. Li[1], H. Yoshino[1], K. Nishioka[2], H. Tomiyama[2,3], N. Hattori[1,2,3].* 1) Research Institute for Diseases of Old Age, Graduate School of Medicine, Juntendo University, Tokyo, Japan; 2) Department of Neurology, Juntendo University School of Medicine, Tokyo, Japan; 3) Department of Neuroscience for Neurodegenerative Disorders, Juntendo University School of Medicine, Tokyo, Japan.

Background and Objectives: Gaucher's disease is an autosomal recessive lysosomal storage disease caused by glucocerebrosidase (*GBA*) mutation. Recent studies suggest that *GBA*mutation is a strong risk factor of Parkinson's disease (PD). We aim to investigate whether there is an association between familial PD (FPD) and *GBA* mutations. **Methods:** All coding exons and exon-intron boundaries of the *GBA* gene were sequenced in 144 Japanese patients with FPD and 100 Japanese controls. **Results:** We found 27 FPD patients with heterozygous mutations, which were reported in Gaucher's disease. The frequency of *GBA* mutations among FPD was significantly higher than in the controls (27/144=18.8% vs. 1/100=1.0%, $P<0.0001$).**Conclusions:** Our study suggests that heterozygous *GBA* mutation is a risk factor for FPD as well as sporadic PD, in the Japanese population.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3135W

Manganese-related T1 hyperintensities of the basal ganglia in hereditary hemorrhagic telangiectasia (HHT) with iron-deficiency anemia. *M. McKinnon[1], S. Appel-Cresswell[2], B. Jung[3], S. Langlois[1].* 1) Department of Medical Genetics, University of British Columbia, Vancouver, BC, Canada; 2) Pacific Parkinson's Research Centre, University Hospital, University of British Columbia, Vancouver, BC, Canada; 3) Department of Pathology and Laboratory Medicine, Children's and Women's Health Centre of British Columbia, Vancouver, BC, Canada.

Hyperintense T1-weighted MRI lesions secondary to deposition of paramagnetic Manganese (Mn) within the basal ganglia are reported in the setting of chronic hepatic dysfunction and environmental exposure to Mn, and may lead to neurotoxicity and parkinsonism. We present a 58-year old man with hereditary hemorrhagic telangiectasia (HHT) and findings of T1-hyperintensities of the globus pallidus with marked elevation of serum and whole blood Mn levels. Our patient presented with a history of recurrent nosebleeds resistant to therapy, severe transfusion-dependent iron-deficiency anemia, telangiectasias, pulmonary and duodenal arteriovenous malformations (AVM's) and a confirmed mutation in ACVRL1. On MRI screening for cerebral AVM's, incidental findings of T1-hyperintensities were observed within the globus pallidus bilaterally. He complained of mild memory disturbance without motor symptoms. On exam, he had mild retrieval issues on cognitive testing (Montreal Cognitive Assessment 23/30) and subtle bradykinesia and mild rigidity of the upper limbs. Mn levels were 55 nmol/L in serum (ref. range: 2–28 nmol/L)) and 835 nmol/L in whole blood (ref. range: 144–327 nmol/L). He had no history of exposure to Mn and no hepatic dysfunction on liver function tests; however, multiple hepatic arterio-portal shunts were evident on CT exam. Mn-related lesions of the basal ganglia have been rarely described in HHT and our case represents only the 7th reported. It is proposed that abnormal intra-hepatic connections in HHT allow Mn in the diet to bypass normal detoxification in the liver. Those with HHT may be at increased risk as trafficking of Mn relies on transferrin and upregulation of transferrin in the setting of HHT-related iron-deficiency anemia may increase intestinal absorption and accumulation of Mn in the brain. We propose that targeting this patient's chronic anemia and iron-deficiency may be the best approach to management. He is currently being treated aggressively with IV iron-sucrose with further efforts to control epistaxis. Planned directions at our centre include reanalysis of MRI's and retrospective chart review in order to further characterize the frequency of these lesions in HHT. We highlight this case in order to draw attention to a rarely reported, potentially treatable complication of HHT, so that appropriate testing and therapeutic intervention might be initiated to prevent the development of parkinsonism in this patient population.

3136F

***CASK* aberrations in males with Ohtahara syndrome and cerebellar hypoplasia.** *H. Saitsu[1], M. Kato[2], H. Osaka[3], N. Moriyama[4], H. Horita[5], K. Nishiyama[1], T. Yoshinori[1], H. Doi[1], N. Miyake[1], K. Hayasaka[2], N. Matsumoto[1].* 1) Human Genetics, Yokohama City University, Yokohama, Kanagawa, Japan; 2) Pediatrics, Yamagata University Faculty of Medicine, Yamagata, Japan; 3) Neurology, Kanagawa Children's Medical Center, Yokohama, Japan; 4) Pediatrics, Hitachi, Ltd. Hitachinaka General Hospital, Japan; 5) Pediatrics, Ibaraki Disabled Children's Hospital, Mito, Japan.

Purpose: Ohtahara syndrome (OS) is one of the most severe and earliest forms of epilepsy. *STXBP1* and *ARX* mutations have been reported in patients with OS. In this study, we aimed to identify new genes involved in OS by copy number analysis and whole exome sequencing. Methods: Copy number analysis and whole exome sequencing was performed in 34 and 12 patients with OS, respectively. Fluorescence *in situ* hybridization, quantitative PCR, breakpoint-specific and reverse transcriptase PCR analyses were performed to characterize a deletion. Key findings: Genomic microarray analysis revealed a 111-kb deletion involving exon 2 of *CASK* at Xp11.4 in a male patient. The deletion was inherited from his mother, who was somatic mosaic for the deletion. Sequencing of the mutant transcript expressed in lymphoblastoid cell lines derived from the patient confirmed the deletion of exon 2 in the mutant transcript with a premature stop codon. Whole exome sequencing identified another male patient harboring a c.1A>G mutation in *CASK*, which occurred de novo. Both patients showed severe cerebellar hypoplasia along with other congenital anomalies such as micrognathia, a high arched palate and finger anomalies. Significance: The detected mutations are likely to cause the loss of function of the CASK protein in males. *CASK* mutations have been reported in patients with intellectual disability with microcephaly and pontocerebellar hypoplasia or congenital nystagmus, and those with FG syndrome. Our data expand the clinical spectrum of *CASK* mutations to include OS with cerebellar hypoplasia and congenital anomalies at the most severe end.

3137W

Unusual presentation of combined saggital-metopic synostosis is caused by mutations in the MSX2 gene: expanding the phenotype of the Boston-type craniosynostosis syndrome. *OM. Vanakker[1], A. Janssen[2], M.J. Hosen[1], Ph. Jeannin[2], PJ. Coucke[1], A. De Paepe[1].* 1) Center for Medical Genetics, Ghent University Hospital, Ghent, Belgium; 2) Department of Paediatrics, Jan Palfijn Hospital, Ghent, Belgium.

Background. Craniosynostosis, caused by early fusion of one or more cranial sutures, can include premature fusion of the saggital (scaphocephaly) or metopic suture (trigonocephaly). Though often occurring as isolated findings, their co- existence in a craniosynostosis syndrome is infrequent and mostly sporadic. Case description. The male proband, born at term, presented at birth with premature fusion of the saggital and metopic suture. Antropometric measurements were normal. Besides the scaphocephaly, a prominent boney mass on the forehead was noted, next to proptosis. Imaging suggested the frontal mass to originate partially from the fused metopic suture. Multiple endocranial hypoplastic areas of the skull were noted. Radiographs revealed bilateral agenesis of the middle phalanges in the feet. Family history revealed the father, his sister and half-sister, to have an isolated scaphocephaly with variable cutaneous syndactyly. The paternal grandfather did not have craniofacial anomalies nor syndactyly, though the great grandfather was said to have isolated syndactyly of both hands. Molecular analysis in this family revealed a missense mutation (p.P148L) in the MSX2 gene, associated with the Boston-type of craniosynostosis in a single family. Segregation analysis of the variant, which affects the same highly conserved residue as in the original report, confirmed non-penetrance in the paternal grandfather. Conclusion. This four-generation family shows various expression of a craniosynostosis phenotype with scaphocephaly and a particularly severe and unusual form of trigonocephaly. A saggital-metopic synostosis together with agenesis of the middle phalanges has to our knowledge not been reported. Molecular analysis revealed a MSX2 mutation, previously described in the Boston craniosynostosis family. Besides several unique features such as incomplete penetrance, the limb abnormalities and the cranial sutures involved, our patients share with the original family the autosomal dominant inheritance with anticipation and the observation of multiple endocranial hypoplastic areas on 3D CT imaging of the skull. Though these consistent findings prove to be important diagnostic clues for MSX2-related craniosynostosis, it must be noted that the generation initially affected in this family presented merely with an isolated scaphocephaly and syndactyly. Molecular analysis of the MSX2 gene should therefore also be considered in patients with scaphocephaly and syndactyly.

3138F

Automated Preparation of Samples for Benchtop Sequencing Applications. *I. Meek.* Caliper Life Sciences, Inc., Hopkinton, MA.

Next generation sequencing technologies continue to result in decreases in the cost per basepair of DNA sequencing data. This has enabled sequencing to be a tractable readout for a number of applications, and has made feasible large scale projects, looking across hundreds of samples, that continue to provide information to help answer questions that are fundamental to our understanding of the factors that govern human health. In the past 12–18 months, scaled down versions of next generation sequencing technologies have become commercially available. Dubbed “benchtop sequencers”, these instruments are finding a wide variety of applications that are specific for the throughput and decreased sample-to-answer time that they offer. The Janus NGS Express is an automated sample preparation instrument that was developed specifically as an upfront to benchtop sequencers. The platform comes with a variety of pre-installed, validated applications including several commercially available amplicon panels and kits, small RNA sequencing, and preparation of libraries for de novo assembly of bacterial genomes. Key advantages of the system include reduction in manual processing, increased reproducibility and sample traceability, and an increase in the number of sample a single technician can prepare.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3139W

Neurofibromatosis Type 1 (NF1) and Infantile Myofibromatosis: A shared genetic basis or independently co-existent entities? *S. Krishnamurthi[1], V. Cox[1], H. Guo[2], K. Rauen[1].* 1) Department of Pediatrics - Division of Medical Genetics, UCSF, San Francisco, CA; 2) Department of Pathology, Children's Hospital of Oakland, CA.

Neurofibromatosis Type 1 is a common RASopathy with variable expressivity that affects virtually every organ system of the body. It has been extensively studied and is caused by heterozygous mutations of the NF1 gene on chromosome 17q11.2. Infantile Myofibromatosis (IM) is a mesenchymal disorder presenting in infancy or early childhood with soft-tissue tumors involving skin, muscle, bones and occasionally viscera. Although relatively rare, IM is the most common fibrous tumor of infancy. In contrast to NF1, the etiology is unknown with reported congenital cases having autosomal dominant or autosomal recessive inheritance. We present a 2 year old male who originally presented at birth with multiple tumors in his calf, scalp and back. These progressed over time to involve his face, trunk, retroperitoneum and extremities. Excisional biopsy of the lesion showed focal spindle shaped cells arranged in fascicles. Immunohistochemical staining revealed revealed reactivity for smooth muscle actin and muscle specific actin; negative staining for S100 protein and Desmin. Based on this, a diagnosis of infantile myofibromatosis was made. At 11 months of age, he was noted with multiple café au lait macules. A brain MRI was done at a year of age in follow-up for an interhemispheric lesion that was previously noted during a screen for IM; the MRI revealed bilateral optic gliomas. NF1 mutation analysis demonstrated a causative heterozygous splicing mutation, c.6579+31T>G, providing a confirmatory molecular diagnosis of NF1. On our exam, in addition, he was noted to be macrocephalic and had a tongue lesion consistent with a neurofibroma. A concomitant diagnosis of NF1 and IM has previously been described by De Schepper et al (Dermatology 2004; 209:223–227). To our knowledge, this is the second case reported of a child with a histopathological diagnosis of IM and a confirmation of NF1 by molecular testing. Given that these two unique entities with common embryonal tissue origins presents in our patient, we raise the question and discuss whether these disorders are coincidental or disorders with a common genetic basis.

3140F

Homozygosity Mapping and Candidate Gene cloning identified CLDN10 variation to be possibly responsible for Congenital ichthyosis - renal hypokalemia in an Algerian family. *S. Hadj-Rabia[1], Y. Al-Sarraj[2], M. Kambouris[2,3], C. Bodemer[1], H. El-Shanti[2,4].* 1) Department of Dermatology - MAGEC Reference Center, Hôpital Universitaire Necker - Enfants Malades, INSERM U781, Paris V-Descartes University, Paris, France; 2) Shafallah Medical Genetics Center, Doha, Qatar; 3) Yale University School of Medicine, Genetics, New Haven CT, USA; 4) University of Iowa, Pediatrics, Iowa City, IA, USA.

Ichthyosis is one of the most frequent single-gene disorders in humans, comprising a large group of scaling disorders with diverse etiology. To date, more than 25 genes have been identified that encode a wide spectrum of epidermal proteins, including enzymes of lipid metabolism and of peptide cross-linking, proteases and their inhibitors, epidermal structural proteins, and proteins involved in cellular communication, signaling and gene transcription. Claudins are epithelial cell tight junction protein with Claudin1 involved in neonatal ichthyosis - sclerosis cholangitis. Claudin16 and 19, are involved in renal hypomagnesemia. We identified an Algerian family with two siblings, a boy and a girl, affected with ichthyosis, keratosis pilaris and palmoplantar keratoderma. The two siblings had xerostomia, abnormal shape of the teeth but normal nails and hair. Hypokalemia is a consistent finding in both. A skin biopsy showed hyperkeratosis and thick stratum corneum. The parents are first cousins once removed and there were three unaffected girls in the same sibship. We performed genome-wide SNP genotyping using 200K chips from Illumina for all family members. Data filtration and allele calling were performed by GenomeStudio2011v1. Two regions of homozygosity were identified by the web-based program Homozygosity Mapper, both shared by the affected siblings and were heterozygous in obligate carriers and the unaffected siblings were either heterozygous or homozygous for the other allele. The first is a 13.5 MBregion on the long arm of chromosome 13 (13q13.3-13q33.2) and containing about 44 genes. The second is a 15.3 MB in the pericentromeric region of chromosome 20 (20p11.21-20q12), containing about 135 genes. The two regions did not contain any of genes known to be implicated in ichthyoses or any other skin disorder. We identified 7 candidate genes from the two homozygosity regions and prioritized them according to their function and their plausible involvement of in a skin and possibly renal disease. We identified a missense mutation, p.S129Lin CLDN10, that segregates with the phenotype in the family and is not found in the 1000 genome. The mutation is predicted to be probably damaging by Polyphen. We are examining the allele frequency of the mutation in an ethnically matched control cohort. We will perform in silico protein modeling to anticipate the effect of this variant on the structure of the protein, and possibly perform functional studies.

3141W

Homozygous mutation of SIX1 associated with a severe Branchio-Oto-Renal syndrome phenotype. *S.L. Sawyer[1], M.A. Thomas[1,2], R. Lamont[1], L. Dimnik[3], P. Gordon[2], X.C. Wei[4], F.P. Bernier[1,2], J.S. Parboosingh[1,2], A.M. Innes[1,2].* 1) Medical Genetics, Alberta Children's Hospital, Calgary, Alberta, Canada; 2) Alberta Children's Hospital Research Institute for Child and Maternal Health, Calgary, Alberta, Canada; 3) Molecular diagnostic Lab, Alberta Children's Hospital, Calgary, Alberta, Canada; 4) Department of Pediatric Radiology, Alberta Children's Hospital, Calgary, Alberta, Canada.

Branchio-Oto-Renal syndrome (BOR, OMIM 602588) is an autosomal dominant syndrome caused frequently by mutations in EYA1 and less often SIX5 and SIX1 (1–3% of cases). A clinical diagnosis can be made based on the presence of 3 major features (second branchial arch anomalies, deafness, preauricular pits, auricular deformity and renal anomalies) or two major and two minor criteria (external auditory canal anomalies, middle or inner ear anomalies, preauricular tags, facial asymmetry or palate abnormalities). We report a child born to consanguineous parents who clinically has many features in keeping with BOR as well as other congenital and developmental anomalies. She was delivered at 29 weeks secondary to premature rupture of membranes and presented with multiple dysmorphic features including: a dysplastic left ear with absent external auditory canal on the left She also has bilateral branchial cysts, bilateral preauricular pits and a bilateral cleft lip and palate. She has hearing loss. Pyelectasis was noted on prenatal ultrasound and a follow up cystourethrogram demonstrated grade 2 reflux. Additional features not previously reported to our knowledge in the BOR spectrum include growth parameters <3rd percentile, hypotonia and severe gastro-esophageal reflux disease requiring fundoplication. On brain MRI she has absent olfactory bulbs, an empty sella with compressed anterior pituitary, partial absence of the anterior falx and persistent Blake's pouch cyst. She is grossly delayed developmentally. At 20 months of age she brings objects to her mouth and is babbling. She is unable to sit or roll. Given her phenotypic complexity and parental consanguinity we did not originally suspect a diagnosis of typical BOR. Array CGH was normal. A SNP homozygosity mapping array identified multiple regions of homozygosity. Follow-up whole exome sequencing revealed a homozygous c.560 +1 G>C mutation in SIX1. A mutation presumably affecting this splice site has been previously reported in the heterozygote state in an atypical BOR patient with ear pits and branchial cysts (c.560 +3 A>T). In our kindred, the obligate carrier parents have minor findings. Her father has an ear pit and her mother has hearing loss. In summary we demonstrate that homozygous mutations in SIX1 confer a more severe BOR-like phenotype, with structural brain anomalies, facial clefting and developmental delay.

3142F

Cardiac Malformations Associated With Germline WT1 Mutations in Children Presenting With Wilms' Tumour. *S. Bowdin[1,2], C. Owens[3], D. Malkin[2,3], N. Parkinson[1], M. Friedburg[2,4], R. Grant[2,3].* 1) Clinical & Metabolic Gen, Hosp Sick Children, Toronto, ON, Canada; 2) Department of Pediatrics, University of Toronto, Toronto, ON, Canada; 3) Division of Haematology/Oncology, Hosp Sick Children, Toronto, ON, Canada; 4) Division of Cardiology, Hosp Sick Children, Toronto, ON, Canada.

Background The WT1 gene was identified as a result of the cytogenetic investigation of individuals with WAGR (Wilms' tumor, Aniridia, Genitourinary anomalies and Retardation). Point mutations in WT1 have been shown to cause Denys-Drash, Frasier and Meacham syndromes, isolated Wilms' tumor and nephrotic syndrome type 4. WT1 encodes a zinc-finger protein which acts as a transcriptional activator or repressor depending on the cellular or chromosomal context. It is required for the normal formation of the genitourinary system and mesothelial tissues, and in the mouse is required for cardiovascular progenitor cell formation. Cardiac malformations are rarely reported in each of the known WT1 phenotypes in humans, and in certain case reports are hypothesised to be coincidental findings. We previously found that 18% of patients with Wilms' tumor at our institution had cardiac findings unrelated to treatment. Method Utilizing oncology and genetic databases at The Hospital for Sick Children, we identified patients with known WT1 mutations or 11p13 deletions and reviewed their cardiac evaluations (including clinical examination, echocardiograms, ECG). Results We identified 14 patients with Wilms' tumor and mutations in WT1 or 11p13 genomic abnormalities. Of two patients with WAGR and genetic testing concordance, 1 had insufficient cardiac information, the other normal findings. 5 of the remaining 12 patients had cardiac findings of note, including complete transposition of the great arteries, VSD, ASD, atrial tachycardia and right-sided aortic arch with aberrant left subclavian artery. Discussion This retrospective review provides evidence of an increased incidence of congenital heart defects in children with WT1 mutations presenting with Wilms' tumour. In conjunction with reports of abnormal cardiac morphogenesis in WT1 conditional knockout mice, evidence suggests the phenotypic spectrum of WT1-associated disorders might include cardiac malformations as well as Wilms' tumour and urogenital malformations.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3143W

False positive diagnosis of Marfan syndrome in adult patient with Homocystinuria who fulfilled Ghent Criteria. *S. Dyack[1,2], H. MacDonald[2], K. Schindeler[2], A. Rideout[2].* 1) Div Med Gen, Dept Pediatrics and Medicine, Dalhousie University, Halifax, NS, Canada; 2) IWK Health Centre, Halifax, NS, Canada.

We present an adult patient with a longstanding clinical diagnosis of Marfan syndrome who was discovered to have Cystathionine Beta Synthase (CBS) deficiency at 48 years of age. He was diagnosed with Marfan syndrome after he developed ectopia lentis (EL) in childhood. He developed a thrombus in his leg in his mid 20's that required anticoagulation therapy. He fulfilled the revised Ghent criteria for Marfan syndrome by having EL, dilatation of his aortic root (AoD) of >4.5 cm, and involvement of the skeletal system, including a reduced upper to lower segment ratio of 0.81 and increased armspan to height ratio of 1.08. He fulfilled the 2010 Ghent criteria by having EL and AoD with Z score >2. He also exhibited joint hypermobility of hips, crowding of his teeth, striae of shoulder, mild scoliosis and arachnodactyly. He had normal intelligence, was married, and had full time employment. His total plasma homocyst(e)ine was markedly elevated at 265.5 µmoles/L with a reference range of 7.1 to 17.3 µmoles/L; consistent with classic homocystinuria. His methionine was elevated at 691 with a normal range of 10–50. Bone density was normal. Three mutations of the CBS gene were found: P78R, K102N and C109R. All have been previously reported in individuals affected with CBS deficiency. He was responsive to B6 therapy. Currently with B6 200mg BID, Betaine 3 grams BID, and attention to diet, his total plasma homocyst(e)ine is in the 20–30 µmoles/L range. Methionine is normal or near normal. After treatment, obsessive compulsive symptoms, that had been attributed to “his personality”, were significantly reduced, by report of his astonished wife. CBS deficiency can be associated with normal intelligence in up to 10 percent of untreated individuals. As illustrated by this case, patients with CBS deficiency may fulfill Ghent Diagnostic criteria and be misdiagnosed with Marfan syndrome. It is important to distinguish between the two conditions as the natural history and treatment differ significantly. It is not known if this patient's AoD is related to CBS deficiency. In any adult patient with ectopia lentis, strong consideration should be given to assessing total plasma homocyst(e)ine even if the patient fulfills Ghent criteria for Marfan syndrome and has normal intelligence. Testing for total plasma homocyst(e)ine should always be done in this context if there is a history of thrombus formation.

3144F

VCP Disease is Associated with Cytokine Imbalances in Patient Plasma. *E. Dec[1], F. Zaldivar[2], M. Wencel[1], M. Khare[1], J. Vesa[1,3], V.E. Kimonis[1].* 1) Division Genetics, Dept. of Pediatrics, UC Irvine, Irvine, CA; 2) ICTS, Dept. of Pediatrics, UC Irvine, Irvine, CA; 3) UC San Diego, San Diego, CA.

VCP disease is caused by mutations in the gene for Valosin Containing Protein. VCP disease classically affects muscle, bone and brain however, new data suggest a common shared pathogenesis with proteinopathies such as amyotrophic lateral sclerosis (ALS) and Parkinson's disease consistent with the fundamental role VCP plays in protein homeostasis. The muscle wasting seen in VCP disease is suggestive of cytokine protein imbalances and was investigated in an effort to determine if VCP mutations alter cytokine profiles. We hypothesized that cytokine protein imbalances are apart of the dysfunctional protein homeostasis caused by VCP disease and contributes to the muscle wasting phenotype. Cytokine analysis was carried out on patient plasma drawn from 16 patients with IBMPFD, 2 pre-symptomatic mutation carriers and 16 controls. Plasma was analyzed for cytokine protein content via ELISA. Four cytokines (TNF alpha, IL-6, EGF, IL-4) historically implicated in muscle wasting and myogenesis were analyzed. Statistical analyses (A 2-sample T-test adjusted for age and gender) revealed the TNFalpha and Epidermal Growth Factor (EGF) were significantly altered in VCP disease as compared to control. TNFalpha was elevated, consistent with a cachexia phenotype, and EGF was low which is consistent with poor myogenesis. IL-4 and IL-6 were not significantly altered when taking in to account age and gender. Cytokine imbalances may be associated with VCP disease and may play a contributory role in VCP myopathy. Further understanding of how VCP dysfunction leads to aberrant protein homeostasis and subsequent cytokine imbalances may also aid in the understanding of other proteinopathies.

3145W

Further expansion of the Chediak-Higashi phenotype: three adult siblings with neurodegenerative disease and homozygous for a novel *LYST* deletion. *L. Mehta[1], C. Cho[2], F.R. Dembitzer[3], A. Szporn[3], M.C. Chicka[4], J.D. Wesfeld-Adams[1].* 1) Department of Genetics & Genomic Sciences, Mount Sinai School of Medicine, New York, NY; 2) Department of Neurology, Mount Sinai School of Medicine, New York, NY; 3) Department of Pathology, Mount Sinai School of Medicine, New York, NY; 4) Prevention Genetics, Marshfield, WI.

Mutations in *LYST*, a ubiquitously expressed gene concerned with lysosomal trafficking, cause Chediak-Higashi syndrome (CHS), an autosomal recessive disorder typically characterized by immune dysfunction, pigmentary abnormalities, and neurologic dysfunction. Rare, atypical CHS patients present with absent or subclinical immunologic abnormalities, and develop neurodegenerative disease in early adulthood. We present a 42-year-old South Asian woman with progressive spastic paraparesis, cognitive impairment, and parkinsonism. As an infant she developed diffuse hyperpigmentation, then mottled hypopigmentation. She had mild learning disability (LD) and developed widespread periodontitis in early adulthood. Examination revealed the pigmentary abnormalities described, and age-appropriate hair graying. MRI of the brain showed cortical atrophy; EMG revealed subacute axonal neuropathy. Skin biopsy demonstrated amphophilic globules in the papillary dermis. A 38 year old brother had mild LD and hypopigmentation in childhood, and lower extremity weakness and parkinsonism from the fourth decade. A 32 year old brother presented age 30 with neuropathic bladder and paraparesis initially attributed to a lumbar arteriovenous malformation. He reported hair graying from adolescence, and adult-onset fine skin hypopigmentation; skin appearances were similar to those of the siblings. Neither brother reported any infections or dental disease. The parents denied consanguinity. Further testing revealed, in all three affected siblings, lymphocytic inclusions, and pigmentary clumping in hair shafts visualized with light microscopy. Patients had abnormally low NK cell activity. Sequencing of the entire coding region of *LYST* in the female proband identified homozygosity for a novel 6 bp deletion in exon 43 (c.9827_9832ATACAA), predicting the loss of two amino acids from the protein product (p.Asn3276_Thr3277del). Targeted sequencing of *LYST* confirmed homozygosity for the deletion in both affected brothers. This family expands the phenotype of CHS to include milder cases with subtle pigmentary abnormalities and neurodegenerative disease, without classic immunodeficiency. The natural history of such cases remains to be defined. Milder, late onset lysosomal storage diseases are well described, and are likely to be under-ascertained. Variant CHS is phenotypically heterogenous and should be considered when young adults develop neurodegenerative disease and have pigmentary abnormalities.

3146F

Familial acanthosis nigricans: Phenotypic features of aberrant area cutanea and hyperpigmented dots identified by dermoscopy. *N. Oiso[1], M. Miyake[1], K. Fukai[2], A. Kawada[1].* 1) Department of Dermatology, Kinki University Faculty of Medicine, Osaka-Sayama, Osaka, Japan; 2) Department of Dermatology, Osaka City University Graduate School of Medicine, Osaka, Japan.

Acanthosis nigricans can be present in the following forms: inherited, obesity-associated, autoimmune-related, drug-induced, and malignancy-associated. Inherited acanthosis nigricans is sub-classified into non-syndromic and syndromic forms. It is characterized by papillomatous brownish lesions mainly in intertriginous areas. We used dermoscopy to examine such lesions in a family with non-syndromic acanthosis nigricans. The dermoscopic images from a 12-year-old Japanese female proband and an 11-year-old younger brother showed the same aberrant skin structure of linear crista cutis and sulcus cutis, and hyperpigmented dots in crista cutis. The hyperpigmented dots, which could not be seen with the naked eye, may contribute to the color of the pigmented skin. Recent studies show that syndromic acanthosis nigricans is caused by a germline mutation in FGFR2 encoding a receptor for the fibroblast growth factors, and epidermal nevi are caused by a somatic mutation in FGFR2. Epidermal nevi are additionally caused by a somatic mutation in the RAS (HRAS, NRAS, and KRAS) genes which are associated with neuro-cardio-facial-cutaneous syndromes including Leopard syndrome, Costello syndrome, and neurofibromatosis type 1. We hypothesize that the tiny dots develop along a pathway similar to lentiginosis in Leopard syndrome or café-au-lait macules in neurofibromatosis type 1, and that the hyperpigmented dots may contribute to the pigmented skin color.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3147W

Clinical clues to differentiate among severe forms of osteogenesis imperfecta. *P. Prasun, A. Jay, D. Stockton.* PEDIATRICS, Children's Hospital of Michigan, Wayne State University, Detroit, MI., USA.

INTRODUCTION Osteogenesis imperfecta (OI) is a hereditary collagen disorder with autosomal dominant and recessive forms. Autosomal recessive OI is rare. Recently several genes responsible for autosomal recessive OI have been identified. Identification of a genetic defect is helpful for accurate genetic counseling. Autosomal recessive OI is usually severe and clinically resembles severe forms of autosomal dominant OI caused by COL1A1 or COL1A2 mutations. However, certain clinical and radiological features may suggest autosomal recessive inheritance. Awareness of these features is helpful in formulating a time and cost effective diagnostic strategy. We describe three patients with autosomal recessive OI. CASE REPORTS Patient 1: Patient 1 was found to have fractures from the antenatal period. He had wide open cranial sutures, large fontanels, white sclerae, narrow chest, bowed lower extremities, rhizomelia, and postaxial polydactyly of hands. His parents are African American and non consanguineous. The patient was found to have compound heterozygous mutations in the CRTAP gene and no detectable COLIA1 or COL1A2 mutations. Patient 2: Patient 2 was found to have fractures from the antenatal period. She had white sclerae, bowed lower extremities, and rhizomelia. Her parents have Lebanese ancestry and are first cousins through their mothers and second cousins through their fathers. A maternal first cousin of the patient's father was reportedly diagnosed with OI in the 1990s. Patient 2 was found to have homozygous mutations (c.1345 G>A) in the LEPRE1 gene. Testing for COL1A1 and COL1A2 was not done. Patient 3: Patient 3 was found to have fractures and shortening of long bones from the antenatal period. He had wide sutures, large fontanels, white sclerae, and short bowed extremities. His parents are African American and non consanguineous. The patient was found to have compound heterozygous mutations in the LEPRE1 gene and no detectable COLIA1 or COL1A2 mutations. DISCUSSION White sclerae and rhizomelia suggest autosomal recessive inheritance in severe forms of OI. Testing for LEPRE1 and CRTAP may be offered first in presence of these features with consanguinity or African American ethnicity. Autosomal recessive OI may be more common in African Americans than realized before. We have developed an evaluation process based upon clinical and radiological features which may be helpful when investigating OI patients.

3148F

Description of a fetal syndrome associated with HNF1B mutation and a wide intrafamilial disease variability. *M. Rasmussen[1], I. Vogel[1], OB. Petersen[2], M. Ramsing[3], L. Sunde[1].* 1) Department of Clinical Genetics, Aarhus University Hospital, Skejby, Aarhus, Denmark; 2) Department of Obstetrics and Gynecology, Aarhus University Hospital, Skejby, Aarhus, Denmark; 3) Institute of Pathology, Aarhus University Hospital, Noerrebrogade, Aarhus, Denmark.

Introduction We present two fetuses and their father with a HNF1B mutation in heterozygous form. Clinical and fetopathological reports Fetus 1. The pregnancy was terminated at week 21+3 due to bilateral enlarged hyperecogenic kidneys with normal amniotic fluid volume. Post mortem examination: Male fetus with mild dysmorphic features. Severe pancreatic hypoplasia was observed and the pancreatic tail could not be identified. Scrotum was hypoplastic and histological examinations showed that the epididymis contained more connective tissue than normal. The kidneys were symmetrically enlarged and contained 1–3 mm cysts, separated by thin strands of parenchyma. The total weight of the kidneys was 11,8 g (expected weight 3 g). Histology showed a segmentally disorganized corticomedular differentiation, glomerular dysplasia, cysts derived from all nephron segments, and primitive ducts. Between the ductal structures regions of interstitial fibrosis were seen. Fetus 2. The pregnancy was terminated at week 22+0 due to bilateral enlarged hyperecogenic kidneys with normal amniotic fluid volume. Post mortem examination: Male fetus with mild dysmorphic features. Abdomen prominated due to the large kidneys and scrotum was hypoplastic. Internal investigation revealed massively enlarged kidneys with a total weight of 20,4 g (expected weight 4 g). The pancreas was hypoplastic and annular. The histological picture of the kidneys and epididymis were identical to fetus 1. The father. He was diagnosed with bilateral hypoplastic kidneys and slowly progressive kidney failure. Genetics. The HNF1B mutation p.Arg165His (c. 494G>A) was identified in heterozygous form in both fetuses and in the father. Discussion The molecular mechanism that causes this wide disease spectrum within the family remains unknown. It can be speculated that the father is a mosaic or the fetuses had inherited one or more modifying variants from the mother. Alternatively, the father could harbor one or more protective variants, which was not passed on to the fetuses. The fact that these two fetuses are siblings suggests a common genetic modifying background. In the severely affected fetuses previously described in the literature several different HNF1B mutations seem to cause a common phenotypic pattern of multicystic dysplasia, pancreas hypoplasia and abnormal genital tract. The repetition of very specific histological findings in two siblings supports the conclusion that a fetal HNF1B syndrome exists.

3149W

Novel Molecular Changes and Their Associated Clinical Characteristics in Saudi Patients with Familial Hemophagocytic Lymphohistiocytosis. *A. Al-Ahmari[1], O. Alsmadi[2], L. Elbaik[2], T. Elamin[2], B. Al-Saud[3], S. Al-Shambri[1], M. Al-Awwami[4], I. Al-Fawaz[1], M. Ayas[1], K. Siddiqui[1], M. Viqaruddin[1], A. Hawwari[2].* 1) Pediatric Hematology/Oncology, King Faisal Specialist Hospital and Research Centre, Riyadh, Saudi Arabia; 2) Section of Immunogenetics, Genetics, Research Center, King Faisal Specialist Hospital and Research Centre, Riyadh Saudi Arabia; 3) Pediatric Allergy/Immunology, King Faisal Specialist Hospital and Research Centre, Riyadh Saudi Arabia; 4) Pathology and Laboratory Medicine, King Faisal Specialist Hospital and Research Centre, Riyadh Saudi Arabia.

Familial Hemophagocytic Lymphohistiocytosis (FHL) is a rare autosomal recessive disorder which is characterized by uncontrolled activation of lymphocytes leading to massive lymphohistiocytic infiltration of body organs including central nervous system (CNS).The hallmark of the disease is impaired or absent function of cytotoxic lymphocytes and natural killer cells as a result of several genetic defects. Up to date, four FHL-causing genes have been identified: PRF1, MUNC13-4, and STX11 and STXBP2. Higher incidence of inherited diseases has been reported in Saudi Arabia due to high consanguinity. Saudi patients with the clinical diagnosis of FHL in the period of 1995 through 2011 and available DNA samples were identified. As of now, analysis for PRF1, MUNC13-4, and STX11 and STXBP2 mutations by direct gene sequencing from 40 patients were completed. Twenty-eight patients with ten different novel mutations were identified and their clinical and biochemical profiles were retrospectively captured. One novel STX11 mutation was identified in 7 patients of the same tribe whereas another mutation was identified in 2 related patients with PRFI genetic defect. Additional two novel mutations were identified in 10 patients of 9 unrelated families with STXBP2 gene defect. The remaining 6 novel mutations were identified in 9 patients of 7 unrelated families with Munc13-4 molecular defect. In addition, three previously reported mutations were detected in patients with Munc13-4 (two patients) and PRF1 (one patient) genes. No molecular defects were identified in the remaining 9 (23%) patients. The presenting clinical and biochemical features were similar in the various genetic groups. The disease presented in the first year of life in all but 2 patients with STX11 mutations where they presented at 5 and 3 years of age. Parents Consanguinity was observed in 70% of our patient's cohort. Two thirds of the Unc13-4 and all PRF1 patients died of disease of progression prior to Stem Cell Transplantation. Four STXBP2 patients with P477L nonsense mutation died either of disease progression or transplantation-related complications whereas two patients continued to have multiple episodes of disease reactivation after Stem cell transplantation with 9 and 12 years survival.

3150F

Hip Pathology in Majewski Osteodysplastic Primordial Dwarfism Type II. *M.B. Bober[1], A.F. Karatas[2], A.L. Duker[1], K. Rogers[2], C. Ditro[2], W.G. Mackenzie[2].* 1) Gen Div, DuPont Hosp Children, Wilmington, DE; 2) Orthopedics Department, DuPont Hosp Children, Wilmington, DE.

Introduction Majewski osteodysplastic primordial dwarfism type II (MOPDII; OMIM #210720) is characterized by severe pre- and postnatal growth failure alongside microcephaly, characteristic skeletal dysplasia, abnormal dentition, an increased risk for cerebrovascular disease and insulin resistance. MOPD II is caused by mutations in the pericentrin (PCNT) gene and is inherited in an autosomal recessive manner. Hip pathology in MOPD II patients includes: hip dysplasia, coxa vara, coxa valga and hip subluxation. The purpose of this study was to determine the incidence of hip pathology in molecularly confirmed MOPD II patients and describe the functional outcomes of surgical treatment. **Methods** An IRB approved Primordial Dwarfism Registry has been established at the Alfred I. duPont Hospital for Children. 33 enrolled patients have a clinical diagnosis of MOPDII. Biallelic PCNT mutations or absent pericentrin protein was confirmed in 25 of these patients. 13 patients (7 female and 6 male) had appropriate clinical and radiographic records and were included in this study. The data collected included: age at presentation, age at surgery, sex, body weight and height, weight-bearing status at diagnosis, and the clinical examination. Clinical evaluation included pain and functional complaints, hip, spine, and lower extremity ranges of motion. **Results** Four patients (31%) had coxa vara: 3 unilateral and one bilateral. Three unilateral patients had in-situ pinning at mean age 4 years. The bilateral coxa vara patient was operated at 5 years of age by valgus osteotomy. Two children had bilateral hip dysplasia and subluxation with no surgery. Two patients had developmental hip dislocations. One was unilateral, treated by closed reduction. One was bilateral, treated open reduction, and two years after surgery, coxa valga was noted. Another patient was diagnosed with bilateral AVN at 12 years old. Four patients did not have hip pathology. **Conclusions** Coxa vara seems to be a common hip pathology among MOPD II patients. Routine clinical and radiographic evaluation of the hips is important. The capital femoral epiphysis appears to slip down along the shaft giving the appearance of a proximal femoral epiphysiolysis. This pathology of the hip, diagnosed as slipped capital femoral epiphysis in early years of life, may progress to severe coxa vara on the follow up visits, often requiring surgery. Strong consideration for pinning is recommended to avoid this progression.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3151W

Boyadjiev-Jabs syndrome: clinical and molecular characterization of three new patients. *S.A. Boyadjiev Boyd[1], S. Kim[1], L. Bivina[1], E. Zackai[2], P.L. Crotwell[3], K. Õunap[4], J. Kim[1].* 1) Department of Pediatrics, Section of Genetics, University of California Davis, Sacramento, CA; 2) Children's Hospital of Philadelphia, Philadelphia, PA; 3) Sanford Clinic and Genetics Laboratory, University of South Dakota, SD; 4) Department of Genetics, United Laboratories,Tartu University Hospital, Tartu, Estonia.

Boyadjiev-Jabs syndrome (Cranio-lenticulo-sutural dysplasia, OMIM 607812) is a dysmorphic syndrome manifesting with late-closing fontanels and calvarial defects, typical gestalt due to facial dysmorphisms, hypertelorism and eye anomalies, generalized osteopenia and developmental delays. We identified a homozygous mutation F382L in SEC23A, an integral member of the COPII-mediated early secretory pathway in the original family. Here we present the clinical features of three new unrelated probands and expand the phenotypic characterization of this syndrome. We report a heterozygous M702V mutation in SEC23A causing a defect of collagen export from the endoplasmic reticulum (ER) and a heterozygous deletion of SEC23A in another proband. No second mutant allele was identified in these two probands raising the possibility of more complex inheritance pattern. We will present the effect of the mutant alleles on the export of procollagen from the ER as analyzed by cell-based assays and electron microscopy. We propose that Boyadjiev-Jabs syndrome is more common than previously thought and should be considered in patients with skull ossification defects, hypertelorism, eye and skeletal anomalies.

3152F

Schinzel-Gideon syndrome in two Brazilian patients: report of a novel mutation in SETBP1. *E.D.F. Carvalho[1], M. Lazar[2], T.F. Almeida[1], C.R.D.C. Quaio[1], G.L. Yamamoto[1], K.M Rocha[2], C.A. Kim[1], M.R. Passos-Bueno[2], D.R. Bertola[1,2].* 1) Unidade de Genética, Instituto da Criança do Hospital das Clínicas, Universidade de São Paulo, São Paulo, SP., Brazil; 2) Instituto de Biociências, Universidade de São Paulo, São Paulo, SP, Brazil.

Schinzel-Giedion syndrome (SGS) is an autosomal-dominant disorder recognizable by the presence of typical facies, growth failure, skeletal malformations, hydronephrosis, cardiac and urogenital defects, developmental delay and a propensity to develop neuroectodermal tumors. Most affected individuals die before the age of ten. SGS was recently regarded as caused by de novo mutations in SETBP1. In 2008, when SGS pathogenesis had not yet been determined, Lehman et al. proposed clinical diagnostic criteria (LDC) that yielded 100% sensitivity for the previously reported cases. The LDC considers the obligatory presence of both developmental delay and facial phenotype (prominent forehead, midface reatraction and short, upturned nose) plus one of the following features: hydronephrosis or typical skeletal abnormalities to establish the diagnosis. Only 15 patients with molecularly proved SGS have been described. We report two unrelated Brazilian patients with SGS confirmed by molecular analysis. Case 1 is a 1-month-old male, who fulfilled the LDC for SGS; he had microcephaly with large anterior fontanelle, typical facies, transverse palmar creases, typical skeletal malformations, bilateral ventriculomegaly, hydronephrosis and hypospadias. He presented seizures within the first month of life, profound developmental delay and hypertonia. Molecular analysis detected the previously described heterozygous mutation c.2602G>A. Case 2 is a one-year-old female that did not fulfill the LDC. She had typical facial phenotype, developmental delay, alacrimia and seizures. She did not have hydronephrosis or the typical skeletal malformations. The molecular analysis revealed a novel mutation (c.2601C>A), assumed as pathogenic by prediction analysis by Polyphen and SIFT, as it affects the protein function and by the fact that it is a de novo event. This is the first report of molecularly-confirmed patients with SGS in Brazil, one among whom presenting a new mutation. We emphasize the importance of the presence of the facial gestalt and developmental delay for the clinical suspicion of SGS. Once the molecular background of this syndrome is now recognized, it is possible that atypical cases or cases with a milder presentation (not fulfilling the LDC, such as the case here described) will be reported, broadening the phenotypic spectrum within SETBP1 gene mutations.

3153W

HOXA10 and HOXA13 sequence variations in human female genital malformations including congenital absence of the Uterus and Vagina. *A. Ekici[1], C. Büttner[1], P. Strissel[2], P. Oppelt[2], S. Renner[2], S. Brucker[3], M. Beckmann[2], R. Strick[2].* 1) Institute of Human Genetics, University of Erlangen-Nürnberg, Erlangen, Germany; 2) Department of Gynecology and Obstetrics, Laboratory for Molecular Medicine, University-Clinic Erlangen, Germany; 3) University-Clinic, Department of Obstetrics and Gynecology, Tübingen, Germany.

Background: Congenital genital malformations occurring in the female population are estimated to be 5 per 1,000 and associate with infertility, abortion, stillbirth, preterm delivery and other organ abnormalities. Complete aplasia of the uterus, cervix and upper vagina [Mayer-Rokitansky-Küster-Hauser (MRKH) syndrome] has an incidence of 1 per 4,000 female live births. The genes responsible for the etiology of congenital genital malformations including MRKH are unknown. The homeobox (HOX) genes HOXA10 and HOXA13 located on chromosome 7 are involved in the development of human genitalia. **Methods:** HOXA10 and HOXA13 genes of 27 patients with the MRKH syndrome, 10 non-MRKH patients with genital malformations and 53 control women were sequenced. A novel exonic nucleotide deletion discovered in HOXA10 was analysed by real time PCR in a total of 103 patients with MRKH and 109 non-MRKH patients with genital malformations. **Results:** A total of 15 DNA sequence variations (11 novel and 4 known) within exonic and untranslated regions (UTRs) were detected in HOXA10 and HOXA13 among our cohorts. Five HOXA10 and two HOXA13 DNA sequence variations were found solely in patients with genital malformations. In addition to mutations resulting in synonymous amino acid substitutions, missense mutations in the HOXA10 gene were found in one MRKH patient and one non-MRKH patient with a bicornate uterus, with the latter predicted by computer analysis as probably damaging to protein function. An exonic HOXA10 cytosine deletion was identified in a non-MRKH patient with a septate uterus and renal malformations resulting in a premature stop codon and loss of the homeodomain helix 3/4. This cytosine deletion in HOXA10 demonstrated an incidence of one in a total of 212 MRKH and non-MRKH patients with genital malformations. Lastly, in the 5'UTR of HOXA10 in a non-MRKH patient with genital malformations, a cytosine to an adenosine base change was found. **Conclusions:** Rare DNA sequence variations in the HOXA10 gene were detected in MRKH patients but especially in non-MRKH patients with genital malformations. Consequences of non-synonymous mutations, a cytosine deletion or a 5'UTR mutation, could result in altered protein function or abnormal gene expression of HOXA10.

3154F

Establishing a paternal age effect for Crouzon syndrome with acanthosis nigricans. *T. Greer[1], C. LaDana[2], P. Barros-Nunez[3], F. di Rocco[4], C. Collet[5], E.W. Jabs[6], R.L. Glaser[1].* 1) Department of Biology, Stevenson University, Stevenson, MD; 2) Department of Cell Biology and Molecular Genetics, University of Maryland, College Park, Maryland; 3) Centro de Investigación Biomédica de Occidente, IMSS, Guadalajara, Mexico; 4) Unité de Chirurgie Craniofaciale, Pediatric Neurosurgery, Assistance Publique-Hôpitaux de Paris, Hôpital Necker Enfants Malades, Paris, France; 5) Service de Biochimie et de Biologie Moléculaire Assistance Publique-Hôpitaux de Paris, Hôpital Lariboisière, Paris, France; 6) Department of Genetics and Genomic Sciences, Mount Sinai School of Medicine, New York, New York.

Crouzon syndrome with acanthosis nigricans, also known as Crouzonodermoskeletal syndrome (CDSS), is an extremely rare autosomal dominant disorder with an estimated incidence of 1/ 1,000,000. CDSS is characterized by craniosynostosis, acanthosis nigricans, vertebral anomalies, along with choanal atresia, and hydrocephalus. Molecular diagnosis of CDSS can be established by confirming the presence of the Ala391Glu mutation in the fibroblast growth factor receptor 3 (FGFR3) gene. This is the only mutation which has been found in cases of CDSS. Parental age data was obtained for ten families. Parental ages at the time of the affected child's birth were matched to parental age data in the general population for each birth year. The average paternal age for fathers of affected children with CDSS was 38.47 (+9.62) years compared with 30.23 (+1.34) years in the general population (p=0.013). The average maternal age for mothers of affected children with CDSS was 31.20 (+4.86) years compared with 27.15 (+1.25) years in the general population (p=0.04). Fathers of affected children were, on average, 7.27 years older than mothers of affected children. DNA from parent-child trios was obtained from five families with sporadic CDSS. The DNA samples were PCR amplified and subjected to restriction enzyme digests to determine the informativeness of the families. Two of the five families were shown to be informative. For these two informative families, allele specific PCR was performed on the parents and affected child in order to determine which parent contributed the allele with the mutation. In both cases, it was determined that the mutations were paternal in origin. Our findings of a paternal age effect for CDSS and two paternally derived mutations in CDSS are consistent with findings in other autosomal dominant disorders such as Crouzon syndrome without acanthosis nigricans with FGFR2 mutations and achondroplasia and Muenke syndrome with FGFR3 mutations. All of these disorders are strongly associated with advanced paternal age and have an exclusive paternal origin of mutation.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3155W

Characterization of Hyper-IgM Syndrome Due to CD40 Deficiency in 11 Patients. *A. Hawwari[1], H. Alassiri[1,2], Z. Al-Sum[3,5], A. Al-Ghonaium[3], S. Al-Muhsen[3,5], H. Al-Dhekri[3], R. Arnaout[3], O. Alsmadi[1], E. Borrero[1], A. Abu-staiteh[1], H. Al-Mousa[1,4], B.K. Al-Saud[3,4].* 1) Section of Immunogenetics, Department of Genetics, Research Center, MBC03, King Faisal Specialist Hospital & Research Center, Riyadh, Saudi Arabia; 2) College of Science, King Saud University, Riyadh, Saudi Arabia; 3) Section of Allergy and Immunology, Department of Pediatrics, King Faisal Specialist Hospital & Research Center, Riyadh, Saudi Arabia; 4) College of Medicine, Alfaisal University, Riyadh, Saudi Arabia; 5) College of Medicine, King Saud University, Riyadh, Saudi Arabia.

Hyper-IgM syndrome due to CD 40 deficiency (HIGM3) is a rare form of primary immunodeficiency, where only few reported cases in the literature, therefore we review in this report our experience in the management of 11 patients. In this study, we sought to identify the Clinical, Immunological, and Molecular Characterization of Hyper-IgM Syndrome due to CD40 Deficiency. For this purpose, a clinical review and a molecular genetic analysis of patients diagnosed with HIGM3 at a single center from 1994–2011 was performed. Eleven patients from seven families were studied. The median age was 8 years (1–21). All 11 patients had recurrent chest infection at presentation. Pneumocystis jiroveci pneumonia was confirmed in three patients. Four patients had Sclerosing cholangitis and four patients had cryptosporidium isolated from their stool. Six patients had nasal and sinus infection, two with destructive nasal fungal infections. Eight patients had neutropenia, all had low IgG and normal or high IgM. IgA was undetectable in all patients except for three patients. Two novel mutations were identified in seven patients; a splice site mutation was found in 5 patients at position +2 of the invariant donor splice site of interon 3 (c. 256+2T>C) and a missense mutation was found in two patients in the coding region of exon 3 (c.170C>T). Two patients underwent successful stem cell transplantation from a matched donor. Five patients doing well on prophylaxis, two are very sick one with protracted diarrhea and persistent cryptosporidium and the other patient had neurological complications. Three patients died early in life because of severe sepsis. To our knowledge this is the largest and longest fellow up case series reported in this disease so far. Our series showed theat CD40 deficiency is s heterogeneous disorder with a spectrum of disease severity. In summary: Hyper-IgM syndrome due to CD40 deficiency is a rare primary immunodeficiency disorder, only few cases published. Here in this report and for the first time a study present 11 patients long term clinical outcome, immunological and molecular Characterization.

3156F

New insights into the renal aspects of hypotrichosis-lymphedema-telangiectasia syndrome caused by a mutation in the *SOX18* gene. *S. Moalem[1], M. Vikkula[2], E. Harvey[3], D. Chitayat[1,4].* 1) Department of Pediatrics, Division of Clinical and Metabolic Genetics, The Hospital for Sick Children, The University of Toronto, Toronto, Canada; 2) Laboratory of Human Molecular Genetics, de Duve Institute Université catholique de Louvain, Walloon Excellence in Life Sciences and Biotechnology (WELBIO), Belgium; 3) Division of Nephrology, The Hospital for Sick Children, University of Toronto, Toronto, Canada; 4) Department of Obstetrics and Gynecology, The Prenatal Diagnosis and Medical Genetics Program, Mount Sinai Hospital; University of Toronto, Toronto, Canada.

SOX18 - The SRY related HMG box-containing gene 18 transcription factors belongs to the highly conserved SOXF group of genes which plays a pivotal role in the development and/or maintenance of the developing heart, blood and lymphatic vessels by activating a transactivation domain adjacent to the HMG domain. SOX18 mutations in humans are associated with both recessive and dominant HLTS. We report a family with two affected children with SOX18 mutations, with the living child developing renal failure. The couple, who was healthy and non-consanguinous, had their first pregnancy complicated with nonimmune hydrops fetalis that resulted in death in utero at 30 wks gestation. The autopsy showed chylous effusions in the pleural and peritoneal cavities, generalized vascular congestion of the lungs with dilation of the lymphatic vessels. The pregnancy with the proband was uncomplicated. After delivery he developed RDS and CXR showed bilateral pleural effusions, which resolved spontaneously. He had large bilateral hydroceles that required surgical repair at 3 months of age and scrotal telangiectasia. At 6 months of age he experienced progressive hair loss resulting in alopecia universalis by the age of two. As an infant and toddler he experienced multiple episodes of facial, peripheral and mild pulmonary oedema. His IgG level was persistently low but he did not experience unusual or particularly severe infections. He was diagnosed with HLTS and DNA analysis for the SOX18 gene showed a novel heterozygous C865A mutation. DNA analysis done on his late brother showed the same mutation and parental DNA analysis using DNA extracted from peripheral blood showed no detectable mutation. At 5 years of age he presented with renal failure with severe hypertension. Renal biopsy demonstrated a chronic microangiopathy involving glomerular and extraglomerular vasculature. He required almost a year of chronic peritoneal dialysis but recovered sufficient renal function to discontinue dialysis for 9 years, during which time he had a slow progressive deterioration in renal function culminating in the need for transplantation at age 14. The SOX18 gene has a role in the formation of blood and lymphatic vessels and a mutation in this gene, as seen in our patient, results in HLTS. This is the first report of renal failure associated with a heterozygous SOX18 gene mutation, most probably due to abnormal angiogenesis. This is also the first reported case of gonadal mosaicism in HLTS.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3157W

Missense and synonymous *TCF4* mutations are responsible for splicing defects in Pitt-Hopkins Syndrome. *M. Nasser[1,2], L. Drévillon[1,2], A. Briand-Suleau[1,2], J. Ghoumid[1,2], T. Gaillon[1], V. Bodereau[1], L. Pasquier[3], M. Goossens[1,2], J. Amiel[4], D. Héron[5], I. Giurgea[1,2].* 1) AP-HP, Groupe hospitalier Henri-Mondor, Service de Biochimie et Génétique, Créteil, France; 2) INSERM U955, Equipe 11, and Université Paris-Est, Créteil, France; 3) Service de Génétique Clinique, CHU Rennes, France; 4) AP-HP, Service de Génétique, Hôpital Necker-Enfants Malades, Paris, France; 5) AP-HP, Groupe Hospitalier Pitié Salpêtrière, Unité Fonctionnelle de Génétique Clinique, Paris, France.

Pitt-Hopkins syndrome (PTHS, MIM #610954) is a neurodevelopmental disorder associating severe intellectual disability, hyperventilation episodes, and characteristic facial gestalt. *TCF4*, encoding a basic helix-loop-helix (bHLH) transcription factor which belongs to the family of E-proteins, was identified as the disease causing gene. PTHS patients usually carry *de novo*deletions or truncating mutations. So far, without two exceptions (c.1073G>T, p.Gly358Val and c.1604A>G, p.Asp535Gly), missense mutations are always localized in the bHLH domain. Recent data demonstrate impairment in the protein function for the bHLH missense mutations, but not for those located outside this domain. In this study, we investigated the impact on splicing of both missense mutations situated outside of the bHLH domain, and of a synonymous mutation (c.1071A>G, p.Ala357Ala) identified *de novo* in an unreported patient with typical PTHS phenotype. No splicing defects were predicted for these mutations by use of bioinformatics tools. Using a functional splicing assay based on a minigene construct, we showed the exon 14 skipping for the c.1071A>G and the c.1073G>T mutations and the skipping of 46 bp of the exon 17 for the c.1604A>G mutation. This results were confirmed by reverse transcription-PCR analyses from peripheral blood lymphocytes mRNA in two patients (c.1071A>G and c.1604A>G). In all cases, the mutations were responsible for frameshift and premature stop codons (c.1071A>G, p.Gly358LysfsX4; c.1073G>T, p.Gly358LysfsX4 and c.1604A>G, p.Asp535AlafsX8). In conclusion, we first demonstrate that the deleterious effect of the missense mutations located outside the b-HLH domain was due to *TCF4* splicing anomalies. Second, we describe the first synonymous mutation in PTHS, and proved its splicing deleterious effect. Finally, for the molecular diagnosis of PTHS we suggest to test the segregation in parents DNA, for the missense mutations outside b-HLH, and for the synonymous, or intronic not reported sequence variants. *De novo* variants in *TCF4* gene have to be cautiously interpreted and splicing tests have to be performed even in absence of *in silico* prediction of a deleterious effect.

3158F

Clinical genetics in silico: Use of an electronic medical record in an integrated healthcare system to identify individuals with undiagnosed Noonan syndrome. *M. E. Nunes[1], R. E. Barber[2], S. K. Kwok[2], R. R. Wilson[2], D. J. Levy[3].* 1) Dept Genetics/Pediatrics, Kaiser Permanente, San Diego, CA; 2) Consulting Services, Kaiser Permanente, San Diego, CA; 3) Dept Pediatrics, Cardiology, Kaiser Permanente, San Diego, CA.

Noonan syndrome (NS) is an autosomal dominant syndrome characterized by short stature, dysmorphology, pulmonic stenosis (PS) and hypertrophic cardiomyopathy (HCM). These and several other NS features can be coded in an electronic medical record (EMR) under the International Classification of Diseases (ICD-9). Left axis deviation (LAD) on electrocardiogram (ECG), independent of the presence of structural heart disease, is found in 50% of those with NS. LAD is an unusual ECG finding, present in 1% of the healthy Pediatric population. We used the EMR and ECG database to see if previously undiagnosed patients with NS could be identified. Using the EMR of an integrated healthcare system serving 954,650 individuals ≤ 19 years old, we identified those over a 5 year period with a combination of common NS features (short stature, PS, etc.) coded in their problem list. We also identified 73 patients with the ICD-9 diagnosis NS (759.89) already in their EMR (Pediatric prevalence 7.6/100,000). Using the ECG electronic database, we identified those in the same age range over the same period with ECG LAD (-30° to -90°). Lists were cross referenced. 65 individuals with at least 2 NS features and LAD on ECG were identified (study group). The EMRs of the study group were reviewed by both a Pediatric Cardiologist and Clinical Geneticist to identify those with additional documented features of NS for further evaluation. 7 (9.6%) of the 73 individuals with known NS coded by ICD-9, would have been identified by the outlined *in silico* process. 60 (82%) of these 73 with existing NS had an ECG, 32 (53%) ECGs revealed LAD. Of the 65 study group individuals identified with the combination of short stature, PS or HCM and LAD on ECG, 4 carried a diagnosis of NS, 1 cardiofaciocutaneous syndrome (CFC), and 1 Kabuki syndrome. 18 of the remaining 59 were felt to have a combination of additional NS associated features (e.g. hearing loss, ptosis) to warrant further evaluation. 8 of these 18 thus far have sufficient documented features to establish a clinical diagnosis of NS. Using an EMR and ECG database, undiagnosed patients with NS can be identified. This is the first demonstration of an integrated healthcare system's EMR being used to identify previously undiagnosed rare genetic disorders. ECG may be of utility in evaluating individuals with short stature and other NS associated clinical features. This method can be reproduced using unique ICD-9 codable features in other syndromes.

3159W

Detailed clinical and orofacial phenotype of three families with compound heterozygous *WNT10A* mutations using three-dimensional imaging methods. *C.W. Ockeloen[1], C. Vink[2], S. ten Kate[1], C. van Heumen[3], T. Kleefstra[1], C. Carels[2].* 1) Department of Human Genetics, Radboud University Nijmegen Medical Center, Nijmegen, the Netherlands; 2) Department of Orthodontics and Craniofacial Biology, Radboud University Nijmegen Medical Center, Nijmegen, the Netherlands; 3) Centre for Special Dental Care, Radboud University Nijmegen Medical Center, Nijmegen, the Netherlands.

Introduction: Odonto-Onycho-Dermal Dysplasia (OODD) and Schöpf-Schulz-Passarge Syndrome (SSPS) are autosomal recessive ectodermal dysplasia syndromes caused by homozygous or compound heterozygous mutations in the *WNT10A* gene. Mutations in *WNT10A* are also responsible for anhydrotic or hypohydrotic ectodermal dysplasia (ED). Moreover, heterozygous mutations account for almost half of isolated hypodontia cases. We present clinical data of three Dutch families with *WNT10A* mutations and a detailed description of the orofacial and dental phenotype using three-dimensional imaging techniques. The first family was diagnosed with OODD, the second family with SSPS and the third family showed hypohydrotic ED. **Methods:** Detailed medical history and examination was performed within our multidisciplinary Craniofacial Genetics clinic. Sequence analysis of *WNT10A* was performed using Sanger sequencing. The orofacial phenotype was documented using an orthopantomogram, a lateral headplate and intraoral and facial pictures. Threedimensional imaging was performed with facial stereophotogrammetry and dental models. **Results:** In the OODD and SSPS family, a similar compound heterozygous mutation in *WNT10A* was identified in the probands: c.321C>A (p.Cys107X) and c.682T>A (p.Phe228Ile). In the SSPS family, these mutations segregated with the phenotype in two affected siblings. The ED patient harboured a homozygous p.Cys107X mutation. Some of the heterozygous carriers displayed a milder phenotype. In the SSPS patients large diastemata were observed intraorally, due to absence of permanent teeth after loss of deciduous teeth. Facial profile analysis points to a retrognathic mandible in the OODD patient and one patient with SSPS, and a midfacial dished-in aspect in the SSPS and ED patients. Common to all affected individuals are the hypoplastic tooth bearing alveolar processes in maxilla and mandible. **Discussion:** These three families clearly demonstrate the broad phenotypic spectrum caused by identical *WNT10A* mutations. Until now, seven different mutations have been reported in *WNT10A* patients. p.Cys107X and p.Phe228Ile are the two most frequently identified mutations. To our knowledge, this is the first report that shows an extensive descriptive overview of the orofacial and dental characteristics of *WNT10A* patients. Part of the heterozygous carriers show a mild phenotype. Taken together, our data on these three families contribute to the expanding phenotype of *WNT10A* patients.

3160F

Tegumentary manifestations in RASopathies are common and deserve special attention. *C.R.D.C. Quaio[1], A.S. Brasil[1,2], A.C. Pereira[2], C.A. Kim[1], D.R. Bertola[1].* 1) Instituto da Criança, Genetics Unit, Faculdade de Medicina da Universidade de São Paulo, São Paulo, SP, Brazil; 2) Incor, Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, São Paulo, SP, Brazil.

RASopathy is a class of common autosomal dominant syndromes with neurocardiofaciocutaneous involvement that includes Noonan syndrome and Noonan related disorders. These syndromes are caused by germline mutations in genes that participate in the RAS/MAPK pathway. Proteins belonging to the RAS/MAPK pathway play important roles in cell proliferation, differentiation, survival, and death. Some components of this pathway, like Gab1 and SHP-2, are related to the promotion of epidermal proliferation and they may interfere in the epidermal cell state and the tegumentary function. The objective of this article is to describe novel tegumentary findings in a cohort of 44 patients with molecularly-proven RASopathy. We found a considerable prevalence of patients complaining of tegumentary problems: 66% (29/44), including xeroderma (20/29), photosensitivity (13/29), hair loss (10/29) and recurrent oral ulcers (12/29). These patients presented mutations in PTPN11 (69%), SOS1 (17%), BRAF (7%), KRAS (3%) and SHOC2 (3%). We were not able to establish a genotype-phenotype correlation or a statistically-relevant ($p<0.05$) difference between the groups of patients presenting tegumentary disorders because the number of patients, though considerable for a genetic disorder, was small to reach a statistical significance. Although the role of the RAS/MAPK pathway in the development and functioning of the skin are not completely understood, several skin manifestations in individuals with RASopathies have been described, like pigmented lesions (café-au-lait spots, lentigines, melanocytic lesions), ectodermal lesions (ichthyosiform manifestations, follicular hyperkeratosis and short, curly, thin hair) and hyperplasia (redundant skin and papillomatous growths). Additionally, somatic mutations in genes of this pathway are associated with skin tumors. Several of the tegumantary manifestations reported in this article have not been previously described in the RASopathic literature. The high prevalence of tegumentary involvement in RASopathy deserves special attention and must be managed properly. Therefore, dermatologists must be part of the multidisciplinary team in the follow-up of RASopathic patients.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3161W

Molecular analysis and expression studies in a novel candidate gene for syndromic coloboma. *N.K. Ragge[1, 2], A. Wyatt[3], D. Robinson[4], D. Bunyan[4], D. Wong[5], I. Ragoussis[5].* 1) Oxford Brookes University, Oxford, UK; 2) Wessex Clinical Genetics Service, Princess Anne Hospital, Southampton, UK; 3) Vancouver Prostate Centre and Department of Urologic Sciences, University of British Columbia, Vancouver, British Colombia, Canada; 4) Wessex Regional Genetics Laboratory, Salisbury, UK; 5) Wellcome Trust Centre for Human Genetics, Oxford, UK.

Developmental eye anomalies affect around 1 in 3,500–4,000 of the population, and are responsible for up to 25% of childhood visual impairment. They range from rare severe forms of bilateral anophthalmia or severe microphthalmia, to milder forms of microphthalmia and coloboma. They form a highly heterogeneous group of disorders, and currently it is possible to identify genetic causes in around 25% of individuals. The commonest causes are genetic mutations or deletions in SOX2, which is responsible for around 10–15% of cases, and OTX2 (around 3–4%), but also include BMP4, BMP7, STRA6, GDF6, GDF3, FOXE3, VSX2, TFAP2A, SHH, HCCS, RABGAP1, RABGAP2, BCOR, SMOC1, ABCB6, FRAS1, FREM1, GRIP1, RAX, PAX6, and LRP2. Clues to genetic identification are provided by the combination of eye and other anomalies. We screened a cohort of 279 individuals with anophthalmia, microphthalmia and coloboma for mutations and deletions in a known transcription factor, not previously associated with a human condition, but known to interact with a known human eye gene. We identified a *de novo*missense mutation in a critical domain of this transcription factor in a girl with bilateral chorioretinal coloboma. She also had a heart defect, abnormal left kidney shape, hearing deficit, mild learning difficulties, and partial toe syndactyly and her growth was on the 2nd centile for height, 9th centile for weight, and <3rd centile for head circumference. We detected 4 heterozygous neutral variants in the gene in the rest of the cohort. This particular base is highly conserved across the phyla and in other members of the group of proteins. The missense alteration causes a change from polar positively charged arginine to non-polar hydrophobic leucine, and is predicted to be “probably damaging" by Polyphen and to “affect protein function" (p=0.01) by SIFT. It was not present in a screen of 280 members of the ECACC control panel, and it is not a reported SNP. We performed *in vitro* studies that demonstrated that the mutant protein had reduced DNA binding ability compared to wild type. Furthermore, we performed *in situ* expression studies in developing mouse and human eye tissue, and detected expression in the lens and retina. Our data implicates this gene in early eye development, and we suggest that it should be included when developing panels of genes to investigate the causes of ocular malformations.

3162F

Variable expressivity of FREM1 related anomalies in a family with novel mutation. *A. Singer[1], A. Slavotnek[2], H. Leiba[3], S. Josefsberg[4], C. Vinkler[5].* 1) Dept Clinical Gen Unit, Barzilai Med Ctr, Tel Aviv, Israel; 2) Department of Pediatrics, Division of Genetics, University of California, San Francisco USA; 3) Department of Ophthalmology, Kaplan Medical Center, Rehovot, Israel; 4) Institute of Medical Genetics, Kaplan Medical Center, Rehovot, Israel; 5) Institute of Clinical Genetics Wolfson Medical Center Holon, Israel.

Manitoba-oculo-tricho-anal (MOTA) syndrome (MIM 248450) is a rare autosomal recessive disorder. It was first described in 1992 by Marels SL among the Oji-Cree community of the Island Lake region of northern Manitoba, Canada. Since the first publication, a few patients were described originating in other parts of the world. The syndrome is characterized by an aberrant hairline, several eyes anomalies (e.g. eye coloboma), a bifid or broad nasal tip, and gastrointestinal anomalies . In 2011 it was found that MOTA syndrome is caused by mutations in FREM1. We present two sibs born to a Jewish couple originating from Uzbekistan, diagnosed with MOTA syndrome and carrying a novel mutation in the FREM1 gene. This is the first case described in Israel. The proband is a 3 month old male born at term to healthy parents who are second degree cousin. At birth, ocular hypertelorism, corneopalpebral synechiae and upper eyelid colobomas were noticed. The anterior hairline extended from the temporale regions bilaterally and the lateral margins of the eyebrows were sparse on both sides. He has a wide nose with a bifid nasal tip and pilonidal dimple. His younger sister had a similar bifid nasal tip and a depressed and hypoplastic columella. Both children were otherwise healthy and had normal development. FREM1 sequence analysis found both the proband and his sister, homozygous for the deletion of a single G nucleotide from a pair of two G nucleotides situated at the last base pair of intron 35 and the first base pair of exon 36. The deletion was predicted to produce a frameshift and to abolish the splice acceptor site. Both parents were heterozygous for the same mutation. MOTA syndrome has some clinical overlap with Fraser syndrome (FS) (MIM 219000), a complex multiple malformation with more severe phenotype. Fraser syndrome is genetically heterogeneous, with mutations in FRAS1 and FREM2 accounting for approximately half of cases. The FRAS1 and FREM2 proteins form a macromolecular ternary complex with a third protein, FREM1. Mutations in the FREM1 gene cause bifid nose sometimes with anorectal and renal anomalies, BNAR syndrome (MIM 608980). Our family emphasize the variable expressivity of FREM1 related anomalies.

3163W

Identification of a *KRIT1* p.Gln201Glu mutation in a Persian family with multiple cerebral, spinal and skin cavernous malformations. *S.M. Sperber[1], D. Fathi[2], M. Shahbazi[3], M.M. Motahari[4], B. Friedman[5], A. Haghighi[6].* 1) DNA Diagnostic Laboratory, Department of Pediatrics, University of Colorado, Anschutz Medical Campus, Aurora, CO, USA; 2) Department of Neurology, Golestan University of Medical Sciences, Gorgan, Iran; 3) Medical Cellular & Molecular Research Center, Golestan University of Medical Sciences, Gorgan, Iran; 4) Department of Ophthalmology, Golestan University of Medical Sciences, Gorgan, Iran; 5) GeneDx, Gaithersburg, MD, USA; 6) Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK.

Cerebral cavernous malformations (CCM) are congenital vascular anomalies predominantly of the central nervous system but may include lesions in other tissues such as the retina, skin, and liver. These hamartous dysplasias, generally occurring sporadically, consist of dynamic clustered tortuous capillary cavities without intervening brain parenchyma that may lead to headaches, seizures, paresis, cerebral hemorrhages and focal neurological deficits. Familial forms of CCM, inherited in an autosomal dominant manner with incomplete penetrance, are attributed to mutations in three genes, KRIT1, CCM2 and PDCD10. Here we report a pedigree of Persian descent clinically evaluated and exhibiting a range of clinical symptoms that include seizures, multiple lesions of the brain, spinal cord and severe hyperkeratotic cutaneous capillary-venous malformations. Clinical sequencing and deletion-duplication testing of the three genes in the proband revealed a KRIT1 p.Gln201Glu mutation. These findings illustrate the phenotypic heterogeneity observed in such pedigrees, and highlight the importance of clinical testing for early diagnosis in familial CCM.

3164F

Novel 3q26 EVI1/MECOM deletion syndrome in a newborn with multiple severe congenital abnormalities and bone marrow failure. *L.T. van der Veken[1], M.B. Bierings[2], M.C. Maiburg[1], F. Groenendaal[3], A.C. Bloem[4], N.V. Knoers[1], A. Buijs[1].* 1) Dept. of Medical Genetics, University Medical Center Utrecht, Utrecht, Netherlands; 2) Dept. of Pediatric Oncology, University Medical Center Utrecht, The Netherlands; 3) Dept. of Neonatology, University Medical Center Utrecht, The Netherlands; 4) Dept. of Immunology, University Medical Center Utrecht, The Netherlands.

We report on a patient born with severe congenital abnormalities including macrocephaly, brain MRI abnormalities, bilateral clubfeet, and bone marrow failure with pancytopenia presenting with none but T-lymphocytes in the peripheral blood. A 824 kb to 868 kb heterozygous deletion on chromosome 3q26.2 was identified by array-CGH on peripheral blood using an Agilent 180k oligo array (Amadid 023363). The deletion truncated the 3' part of *EVI1/MECOM* and extended into microRNA 551b (*MIR551B*). The 3' *MECOM* deletion was confirmed by FISH using a*MECOM* specific probe on uncultured peripheral blood cells in 10 investigated metaphases and 200 interphase nuclei. To assess whether the deletion was constitutional, FISH analysis was performed on buccal swaps and cultured skin fibroblasts. Loss of one *MECOM* signal was observed in 200 interphases, and 10 metaphases of cells derived from buccal mucosa and skin fibroblasts, respectively, demonstrating the constitutional nature of the 3' *MECOM* deletion. Karyotyping and FISH analysis of both parents did not reveal a structural or numerical abnormality of 3q26.2 *MECOM*, indicating the deletion to be *de novo*. EVI1/MECOM is a transcriptional regulator essential for maintaining embryonic and adult hematopoietic stem cells by directly regulating transcription of *GATA2*. In mouse models, homozygous disruption of *MECOM* results in embryonic lethality, with hypocellular bone marrow, reduced body size, small or absent limb buds, abnormal development of the nervous system and heart and massive hemorrhaging. Furthermore, *MECOM* heterozygosity leads to a marked impairment of the self-renewal capacity of hematopoietic stem cells. In our case the heterozygous deletion results in 3' terminal truncated *MECOM* lacking the C-terminal acidic amino acid cluster domain (AD) encoding sequences. The AD of MECOM is important for activation of *GATA2* transcription *in vitro*. This supports the notion that haploinsufficiency of MECOM causes bone marrow failure due to loss of critical*MECOM* transcriptional AD encoding sequences in this case, supposedly by deregulating GATA2 mediated control of hematopoietic stem cell homeostasis. We will discuss our data in relation to one seemingly similar case deposited recently in the DECIPHER database. In conclusion, we report on a novel 3q26 *EVI1/MECOM* deletion syndrome with multiple severe congenital abnormalities and bone marrow failure due to *EVI1/ MECOM* haploinsufficiency.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3165W

KIAA 2022-related X-linked intellectual deficiency: confirmation of a discrete entity in five patients. *L. Van Maldergem[1], V.M. Kalscheuer[2], M. Doco-Fenzy[3], A. Medeira[4], A. de Brouwer[5], E. Landais[3], L. Villard[6], J. Dupont[4].* 1) Centre de Génétique Humaine, Univ Franche-Comté, Besançon, France; 2) Dept Hum Mol Genet, Max-Planck Instit for Molec Genet, Berlin, Germany; 3) Dept Genetics, CHU, Reims, France; 4) Dept Med Genet, Hosp Santa Maria, Lisbon, Portugal; 5) Dept Mol Genet, NUMC, Nijmegen, The Netherlands; 6) INSERM UMRS910, Univ de la Méditerranée, %Marseilles, France.

We describe five boys aged 7 to 22 years, belonging to three families, with a distinctive X-linked intellectual deficiency syndrome. Three of them were diagnosed with infantile autism and two of them suffered from severe epileptic encephalopathy. All but one affected males have hand stereotypies, bruxism, drooling, pes equinovarus with mild spasticity and a distinctive facial appearance that includes round face, strabismus and coma-shaped eyebrows. In two of the families where this phenotype is observed, the molecular basis is a mutation leading to the absence of KIAA2022 expression while in the singleton from the third family, a small interstitial duplication reduces the expression of KIAA2022 and gives rise to a milder phenotype, lacking severe autism and facial dysmorphia, although language delay and repetitive behaviors are present.

3166F

Extending the phenotype spectrum of IQSEC2 mutations : report of a patient with IQSEC2 intragenic duplication and atypical variant of Rett syndrome. *M. Willems[1], F. Tran Mau-Them[1], J. Puechberty[1], G. Lefort[2], A. Schneider[2], I. Touitou[1], M. Girard[2], M. Tournaire[2], N. Ruiz-Pallares[2], F. Rivier[3], S. Drunat[4], P. Sarda[1], D. Genevieve[1].* 1) Genetics Dept, INSERM U844, Hopital Arnaud de Villeneuve, Montpellier, France; 2) Cytogenetics Dept, Hopital Arnaud de Villeneuve, Montpellier, France; 3) Neuropediatrics Dept, Hôpital Gui de Chauliac, Montpellier, France; 4) Cytogenetics Dept, Hôpital Robert Debré, Paris, France.

Missense mutations in the IQSEC2 gene have recently been shown to cause a non-syndromic X-linked intellectual disability (XLID), characterized by substantial limitations in intellectual functioning and adaptative behavior. We report a male patient born from healthy unrelated parents, with normal birth parameters, severe intellectual disability (ID), non inherited postnatal onset microcephaly, hypotonia, hyperkinesia, strabism, Rett-like movements, seizures, and self-injury. Brain MRI, EEG, metabolic screening, and molecular analysis of the MECP2, CDKL5, MEF2C, and FOXG1 genes were normal. Using array-CGH (Agilent 180 k), we identified a de novo Xp11.2 intragenic duplication, interrupting the sequence of the IQSEC2 gene, encompassing exon2. Atypical variants of Rett syndrome correspond to a clinically and genetically heterogeneous group of diseases, characterized by stereotypical hand movements, postnatal microcephaly and severe neurodevelopment disorder. Mutations in the MeCP2, CDKL5, and FOXG1 genes has been identified in these entities. IQSEC2 is a guanine nucleotide exchange factor (GEF) for the ADP-ribosylation factor (ARF) family of GTP-binding proteins. Previous functional studies of the reported missense mutations in the IQSEC2 gene showed that partial loss of the GEF activity of IQSEC2 is the likely underlying molecular mechanism of IQSEC2-XLID. Such missense mutations lead to a reduction of the ARF6 substrate activation that influences the regulation of the actin cytoskeleton organization. The intragenic duplication described here may lead to an interruption of the IQSEC2 gene and generate a more severe loss-of-function of IQSEC2. It has recently been showed that MecP2 stabilizes microtubule dynamics and that MecP2 deficiency could lead to impaired microtubule stability. It is tempting to speculate that the partial loss of actin cytoskeleton organization by either mutation of the MecP2 or IQSEC2 genes could explain phenotypic similarities between our patient and Rett syndrome patients. RNA studies are performed in an attempt to explain the severe phenotype observed in our patient. In addition, molecular analysis of the IQSEC2 gene is performed in patients presenting with atypical variant of Rett syndrome without MecP2, FOXG1 and MEF2C mutations in an attempt to investigate if IQSEC2 mutations could be responsible for atypical variants of Rett syndrome.

3167W

Mutations in TGF-beta binding domains in the *FBN1* gene result in diverse connective tissue dysplasia syndromes with short stature. *K.H.C. Wu[1,2,3], A. Baxter[2], J.R. Pinner[4], D. Mowat[5], T. Dudding[6], B. Bennetts[7,3], K. Holman[7], E. Ormshaw[7], T. Gayagay[7], L.C. Adès[2,8,9], C.L. Goff[10], V. Cormier-Daire[10], D.O. Sillence[1,2,3].* 1) Genetic Medicine, Westmead Hospital, Westmead, NSW, Australia; 2) Department of Clinical Genetics, Sydney Children's Hospital Westmead, NSW, Australia; 3) Discipline of Genetic Medicine, Sydney Medical School at Sydney Children's Hospital Westmead, NSW, Australia; 4) Department of Molecular and Clinical Genetics, Royal Prince Alfred Hospital, Camperdown, NSW, Australia; 5) Department of Clinical Genetics, Sydney Children's Hospital Randwick, NSW, Australia; 6) Department of Clinical Genetics, Hunter Genetics, Waratah, NSW, Australia; 7) Department of Molecular Genetics, Sydney Children's Hospital Westmead, NSW, Australia; 8) Marfan Research Group, The Children's Hospital at Westmead, NSW, Australia; 9) Discipline of Paediatrics and Child Health, University of Sydney, NSW, Australia; 10) Département de génétique médicale, INSERM U781, Université Paris Descartes-Sorbonne Paris Cité Institut Imagine, Hôpital Necker-Enfants Malades, Paris, France.

Purpose: Heterozygous mutations in the 4th (exons 37 & 38) and the 5th (exons 41 & 42) TGF-beta binding protein-like domains in *FBN1* have recently been shown to cause stiff skin syndrome (SSS) and acromelic dysplasias, respectively. We report the clinical, radiological and molecular genetic findings in five patients with heterozygous mutations in *FBN1* from New South Wales, Australia and compare these with previously published reports. **Methods:** Retrospective chart review and molecular genetic testing. **Results:** Three patients had geleophysic dysplasia (GD), one had acromicric dysplasia (AD), and one had SSS. All cases were sporadic. All four patients with acromelic dysplasias had severe short stature (-3 to -4 SD) with onset in infancy, short hands/feet, variable joint contractures, characteristic facial features, and a heterozygous mutation in exon 41/42. The three patients with GD all had hepatomegaly and cardiac valvular involvement; the latter was progressive in one patient, in whom surgical intervention was unsuccessful at age 4 years, and stable in the other two (one required oxygen therapy at age 20 years and the other was an athletic male without any intervention at age 33 years). One GD patient was on medical therapy for primary pulmonary hypertension. The AD patient had laryngeal stenosis and borderline megalocornea. Skeletal features included delayed carpal bone age, predominantly acral changes, short long tubular bones, and a J-shaped sella. Notching of the 2nd & 5th metacarpals and femoral heads was an additional feature in the AD patient. The patient with SSS had generalized stiff skin from infancy, diffuse joint contractures, relative short stature (adult height <10th percentile) with hand length measuring between 3rd-25th percentile, and multiple interphalangeal nodules, as reported recently. This SSS patient had a mutation in exon 37 and had ectopia lentis, representing the second hybrid case following a recently reported similar case with a mutation in exon 38. He also had restrictive lung disease in the absence of significant kyphoscoliosis, normal cardiac status and no peripheral neuropathy at age 36 years. None of the five patients had aortic root pathology or known aneurysmal vascular disease. **Conclusion:** Our findings support the fact that mutations in the 4th and 5th (of seven) TGF-beta binding domains in *FBN1* result in diverse connective tissue dysplasia syndromes with short stature.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3168F

Clinical and Genetic Analysis in Chinese Patients with Megalencephalic Leukoencephalotathy with Subcortical Cysts. *M. Guo[1], Y. Jiang[1], H. Xie[1], Y. Wu[1], J. Shang[2], Q. Gu[1], X. Wu[1], J. Wang[1].* 1) Department of Pediatrics, Peking University First Hospital, Beijing, Beijing, China; 2) Department of Neurology, Shanxi Big Hospital, Shanxi, Taiyuan, China.

Objective Megalencephalic leukoencephalotathy with subcortical cysts(MLC,OMIM 604004)is a rare leukodystrophy caused by mutations in MLC1 (NM_015166)or HEPACAM(NM_152722.4).Up to now no case with HEPACAM mutation is identified in Chinese MLC patients. In our study,the analysis of clinical features and MLC1 and HEPACAM mutations screening were performed in 20 Chinese patients. Methods Clinical data were collected carefully for our 20 Chinese patients. Genomic DNA from 20 Chinese patients and their parents were screened for mutations in the entire coding region, including the exon-intron boundaries and the surrounding intronic regions,first of the MLC1 gene. For patients without MLC1 mutation, we performed HEPACAM mutation screening. Results All 20 enrolled subjects were diagnosed MLC clinically. We detected MLC1 mutation in these 20 patients and a total of 15 MLC1 mutations were identified, including three novel missense mutations(c.803C>G (p.T268R); c.881C>T(p.P294L); c.858C>G(p.I286M)), one novel small deletion (c.596delCAgt(p.S199Cfs220X)), eight known missense mutations(c.65G>A(p.R22Q);c.95C>T(p.A32V);c.206C>T(p.S69L); c.218G>A(p.G73E); c.274C>T(p.R92W); c.353G>T(p.T118M); c.823G>A(p.A275T); c.832T>C(p.Y278H)),one known splicing mutation(c.772-1G>C in IVS9-1), one known small deletion(c.907_930del(p.V303_L310del)), one known nonsense mutation (c.594delCTCA(p.Y198X)). HEPACAM screening were performed in 6 patients without MLC1 mutations. In one patient, we found two HEPACAM mutations (c.203A>T(p.K68M); c.395C>A(p.T132N)) and one mutation is from her father, while the other one from her mother. A reported HEPACAM missense mutation c.274C>T(p.R92W) was found in two patients from two families, respectively. One was inherited from his mother, another was de novo. Conclusions 17 out of 20 Chinese MLC patients diagnosis clinically were identified by gene mutations related MLC1 or HEPACAM. 14 MLC1-related MLC patients were diagnosed clinically and genetically. Mutation c.772-1G>C in IVS9-1, accounting for 21.4%(3/14) of the total number of MLC1 patients in this study, may be a putative hot-spot mutation in this present study group. Three HEPACAM-related MLC patients were diagnosed as one presented phenotype MLC2A, while the other two were MLC2B. The missense mutation c.274C>T(p.R92W) found in Chinese MLC2B patients maybe common mutations in phenotype MLC2B patients. A systemic study to assess mutation spectrum in different population needs to be carried out.

3169W

Skeletal Clinical Characteristics of Osteogenesis Imperfecta Caused by Haploinsufficiency Mutations in COL1A1. *I.M. Ben Amor, P. Roughley, F.H. Glorieux, F. Rauch.* Medical Genetics, Shriners Hospital, Montreal, QC, Canada.

COL1A1 haploinsufficiency mutations lead to a decreased production of collagen type I by osteoblasts and other cell types and consistently result in the mildest form of osteogenesis imperfecta (OI), OI type I. The skeletal clinical characteristics resulting from such mutations have not been characterized in detail. In this study we assessed 86 patients (36 male, 50 female; mean age 13.3 years; range 0.6 to 54 years) with COL1A1 haploinsufficiency mutations. Birth history was positive for fracture or long-bone deformity in 12% of patients. The average rate of long-bone fracture (femur, tibia/fibula, humerus, radius/ulna) was 0.57 fractures per year, half of which affected the tibia/fibula. Long-bone fracture rate was negatively associated with age and lumbar spine areal bone mineral density. Vertebral compression fractures were observed in 71% of the 58 patients who had lateral spine radiographs. The median number of vertebral fractures was higher for females (median 4; range 0 to 14) than for males (median 1, range 0 to 8) (P = 0.03). Lumbar spine areal bone mineral density was negatively associated with the severity of vertebral compression fractures, as reflected in the spine deformity index. Scoliosis was present in about 30% of patients but the Cobb angle was <30 degrees in all cases. The average final height z-score was -1.1, representing a deficit of 8 to 10 cm compared to the general population. In summary, OI patients with COL1A1 haploinsufficiency mutations have high rates of significant skeletal involvement. Systematic follow up of such patients including radiographic screening for vertebral compression fractures and scoliosis is warranted.

3170F

A new dominant frontonasal dysplasia with major posterior cranial defect. *S. Odent[1], S. Mercier[1], M. de Tayrac[2], J. Mosser[2], P. Loget[3], C. Rozel[4], S. Jaillard[5], J. Milon[4], L. Riffaud[6], G. Le Bouar[7], P. Poulain[7], E. Martin[8], C. Dubourg[9], V. David[9].* 1) Genetique Clinique CLAD-Ouest, CHU de Rennes Hopital SUD, Rennes Cedex 2, France; 2) Plateforme Biogenouest, CHU, Hopital Pontchaillou, Rennes, France; 3) Service d'anatomopathologie, CHU, Hopital Pontchaillou, Rennes, France; 4) Service de radiologie, CHU, Hopital Sud, Rennes, France; 5) Service de cytogenetique, CHU, Hopital Pontchaillou, Rennes, France; 6) Service de neurochirurgie, CHU, Hopital Pontchaillou, Rennes, France; 7) Service de gynecologie-obstetrique, Centre de diagnostic prenatal, CHU, Hopital Sud, Rennes, France; 8) Integragen Society, Genopole Campus1, Evry,, France; 9) Laboratoire de genetique moleculaire, CHU, Hopital Pontchaillou, Rennes, France.

Frontonasal dysplasia (FND) is characterized by distinctive facial appearance including medial facial cleft affecting the nose and/or upper lip and palate. Additional features may be present as upper eyelid ptosis and anterior encephalocele. Inheritance is heterogeneous. The ALX3-linked FND is of autosomal recessive inheritance but most FND cases are sporadic. We report on a recurrence of an atypical form of frontonasal dysplasia in a mother and a more severe phenotype in her male foetus. The 30 year-old mother was born at term with a cranio-facial malformation described as an “occipital meningocele” associated with a posterior cranial defect of 8 cm in diameter. No familial history was reported and her parents were not consanguineous. She had hypertelorism, ptosis, bifid nasal tip, high-arched palate and dysplastic ears. Surgical interventions were performed to place prostheses in occipital and frontal regions. She also had cosmetic surgery on nasal tip, maxillary osteotomy and genioplasty. Moreover she had a bilateral mixed deafness, but no mental retardation. During her first pregnancy, a large parietal defect with a voluminous encephalocele was noticed in a male foetus at 15 GW+6 by US scan. Foetal brain MRI at 17 GW found a cranial defect of occipital and parietal regions associated with large brain hernia and ventricular dilatation. Medical termination of pregnancy was performed at 18 GW+3. In addition to the major encephalocele, anatomopathological examination showed a facial dysmorphy resembling to his mother's and a sacral vertebral anomaly. Karyotype, array CGH, Ephrin B and ALX3 gene analyses were normal in the mother and the foetus. Whole exome sequencing led to the identification of a significant heterozygous missense mutation in two new genes. Conclusion: this atypical frontonasal dysplasia is concordant with dominant inheritance and phenotypic variability with increased severity in the male foetus. The occurrence of a de novo variant was confirmed in the mother.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3171W

Cerebrovascular changes in a patient with Noonan syndrome and a RAF1 mutation. *Y. Zarate[1], A. Lichty[1], G. Matheus[2], K. Champion[3], K. Clarkson[1], K. Holden[4].* 1) Greenwood Genetics Ctr, Columbia, SC; 2) Department of Radiology and Radiological Science. Medical University of South Carolina, Charleston, SC; 3) Molecular diagnostics laboratory. Greenwood Genetics Ctr, Greenwood, SC; 4) Department of Neurosciences. Medical University of South Carolina, Charleston, SC.

Noonan syndrome (NS) is a common autosomal dominant developmental disorder characterized by facial dysmorphism, skeletal abnormalities, short stature, and heart defects. Gain-of-function germline mutations affecting components of the Ras-MAPK pathway are involved in the development of NS. We present the case of a 6-year-old male with NS, a RAF1 mutation and unique cerebrovascular changes. The proband was noted to have macrocephaly at 10 months of age (HC: 49 cm, >97th centile) and a head CT showed hydrocephalus, requiring ventriculoperitoneal (VP) shunt placement. By 1 year of age he was diagnosed with hypertrophic cardiomyopathy. At the time of his most recent evaluation at 6 years of age, his height was below the 3rd centile, he had typical NS dysmorphic features, mild pectus excavatum, and a short broad neck. RAF1 gene sequencing revealed a de novo c.770C>T heterozygous change (p.S257L). His most recent brain MRI at 5 years of age showed stable decompressed ventricles with a VP shunt, mild parenchymal volume loss, and Chiari I malformation. However, there were numerous diffusely distributed abnormal small-caliber tortuous vascular structures over the subarachnoid space along the cerebral hemispheres and midbrain suggesting underlying vasculopathic changes. CT angiogram confirmed these prominent abnormal vascular structures predominantly in the region of the sylvian fissures as well as the frontal lobes within the prominent subarachnoid spaces. There was no evidence of an AVM or acute major sinus thromboses. Few cases of Chiari malformation/hydrocephalus have been reported in NS. The association between NS and structural cerebrovascular abnormalities is equally unusual and can rarely include: cerebral AVM, intracerebral occlusive artery disease, aneurysms, hypoplasia of the posterior cerebral vessels, cavernous hemangiomas, and moyamoya disease. The cerebrovascular changes seen in this patient do not fit well into these previously noted categories and their origin and potential clinical implications are currently unknown. While there is clear evidence of the important role of RAF1 in cardiac hypertrophy, its role in vasculopathic changes is more obscure. The findings reported here indicate that it is prudent for patients with NS, particularly in the context of a RAF1 mutation and neurological symptoms, to undergo a complete cerebral neurovascular evaluation.

3172F

***EFTUD2* haploinsufficiency leads to syndromic esophageal atresia.** *J. Amiel[1, 2], C. Gordon[1], M. Oufadem[1], C. Decaestecker[3], AS. Jourdain[4], J. Andrieux[10], V. Malan[2], JL. Alessandri[5], C. Baumann[6], O. Boute-Benejean[3], B. Delobel[7], D. Lacombe[8], S. Mehta[9], I. Simonic[9], F. Escande[4], N. Porchet[4], S. Manouvrier-Hanu[3], F. Petit[3,4], A. Munnich[1,2], M. Vekemans[1,2], S. Lyonnet[1,2], L. de Pontual[1, 11], M. Holder-Espinasse[3].* 1) INSERM U781 and Université Paris Descartes - Sorbonne Paris Cités, Institut Imagine, Paris, France; 2) Genetic department, Neker Hospital, AP-HP, Paris, France; 3) Service de Génétique Clinique, Hôpital Jeanne de Flandre, CHRU Lille, France; 4) Laboratoire de Biologie Moléculaire, Centre de Biologie Pathologie, CHRU Lille, France c Plateforme de génomique, Centre de biologie pathologie, CHRU Lille, France; 5) Service de Pédiatrie, CHD Felix Guyon 97405 St Denis la Réunion, France; 6) Service de Génétique, Hôpital Robert Debré, AP-HP, Paris, France; 7) Laboratoire de génétique chromosomique, Hôpital St Vincent de Paul, Lille, France; 8) Service de Génétique, Hôpital Pellegrin-Enfants, Bordeaux, France; 9) East Anglian Medical Genetics Service, Addenbrookes Hospital, Hills Read, Cambridge, UK; 10) Plateforme de génomique, Centre de biologie pathologie, CHRU Lille, France; 11) Service de Pédiatrie, Hôpital Jean Verdier, Bondy, France.

Esophageal atresia (EA) with or without tracheoesophageal fistula is the consequence of an abnormal septation of the foregut into esophagus and trachea. The incidence is estimated at about 1/3,500 live births. The molecular bases for isolated EA are currently unknown and probably multifactorial. Mandibulofacial dysostosis (MFD) is the consequence of an abnormal development of the first and second branchial arches. The core phenotype is represented by malar and mandibular hypoplasia and dysplastic ears. Conductive hearing loss, lower lid anomalies and/or cleft palate are frequent associated features. The molecular bases of syndromic EA or MFD are being identified at a rapid rate. In particular the *EFTUD2* gene (MIM603892) encoding a protein of the spliceosome complex has been found mutated in patients with MFD and microcephaly (MFDM, MIM610536). Until now, no syndrome featuring both MFD and EA has been clearly delineated although EA is reported in about 5% of patients diagnosed with oculoauriculovertebral spectrum (OAVS, MIM 164210). We report on 2 cases presenting MFD and EA and a de novo 17q21.31 deletion encompassing *EFTUD2*, *CCDC103*, *FAM187A* and *GFAP* not reported in the Database of Genomic Variants. Under a contiguous gene syndrome model, *GFAP* and *CCDC103* were considered unlikely to explain EA while FAM187A could not be excluded. Subsequently, we identified a de novo, heterozygous *EFTUD2* loss-of-function mutation in 4 unrelated patients with MFDM and EA. These data exclude a contiguous gene syndrome, broaden the spectrum of clinical features ascribed to *EFTUD2* haploinsufficiency, define a novel syndromic EA entity and emphasizes the necessity of mRNA maturation through the spliceosome complex not only for global growth but also within specific developmental fields.

3173W

Al-Awadi/Raas-Rothshild/Schinzel-Fuhrman spectrum phenotypes in patients with no mutations in WNT7A. *O. Caluseriu[1], E. Sherridan[2], K.M. Girisha[3], J. Parboosingh[1], A.M. Innes[1], F.P. Bernier[1].* 1) Department of Medical Genetics, Alberta Children's Hosp and University of Calgary, Calgary, AB, Canada; 2) Leeds Institute of Molecular Medicine, Leeds, UK; 3) Genetics clinic, Department of Pediatrics, Kasturba Medical College and Hospital, Manipal, India.

Al-Awadi/Raas-Rothshild/Schinzel Syndrome (AARRS;MIM 276820) and Fuhrman Syndrome (FS; MIM 228930) are autosomal recessive disorders characterized by severe symmetric limb anomalies . All three axes of limb growth are affected: proximodistal, anteroposterior and dorsoventral, and mutual interplay of signalling pathways govern this complex process including fibroblast growth factors (FGFs), wingless (Wnt), sonic hedgehog (Shh), and bone morphogenetic protein (BMP). Mutations in WNT7A have been previously associated with AARRS and FS (Woods et al, 2006). Despite intra-familial variability and a range of anomalies found in the initial affected individuals, genetic homogeneity for AARRS has been proposed (Garavelli et al, 2010). We present a cohort of 5 patients with limb anomalies in the AARRS-FS spectrum with no mutations in WNT7A gene. Three cases of FS-like have been identified. Cases 1 and 2, are sisters born to consanguineous parents of Libyan origin. Case 1 had a single bone forearm with 2-ray ectrodactyly on the right, and 5 rays on the left hand with dislocation of the radial head. Case 2 had bilateral congenital radial head dislocation, 5 rays on both hands and 4 rays on both feet. Both sisters shared generalized nail hypoplasia/aplasia with suggestion of an anterior-posterior gradient, and palmar/plantar creases on the dorsal aspect of both hands and feet. Lower limbs showed hypoplasia of distal femurs that were fused to tibias, and no patellae and fibulae. Spine and pelvis were unremarkable. Case 3 is a boy born to nonconsanguineous parents previously reported (Girisha et al, 2011). A further 2 cases have a proposed diagnosis of AARRS. We provide detailed clinical description and imaging of these patients and discuss their relationship to the AARRS-FS spectrum and implications regarding possible clinical and genetic heterogeneity for this group of conditions.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3174F

A Long polyphenilalanine repeats expansion in the RUNX2 gene in a patient with Cleidocranial dysplasia. *M. Michelson-Kerman[1,2], E. Leshinsky-Silver[1], D. Lev[1,2], A. Singer[2,3], C. Vinkler[1,2].* 1) Inst Med Genetics, Wolfson Medical Ctr, Holon, Israel; 2) Genetic Clinic, Maccabi Health Services, Israel; 3) Medical Genetic Unit, Barzilai Medical Center, Ashkelon. Israel.

Cleidocranial dysplasia (CCD) is a autosomal dominant disorder characterized by skeletal anomalies such as patent fontanels, late closure of cranial sutures with Wormian bones, late erupting secondary dentition, rudimentary clavicles, and short stature. The locus for this disease was mapped to chromosome 6p21. RUNX2 is a member of the runt family of transcription factors and its expression is restricted to developing osteoblasts and a subset of chondrocytes. Deletions, insertions, nonsense and splice-site mutations, as well as missense mutations mutations in the RUNX2 gene have been shown to cause CCD. CCD is one of nine developmental disorders associated with expansions of polyalanine (polyA) tracts. The gene harbors a polyalanine stretch of around 17 GAG repeats. We present a 33y old patient with a severe phenotype of CCD and an expansion of the polyalanine stretch in RUNX2 to 29 repeats. The proband is a 33y old man born to healthy unrelated parents. After birth he had persistently open skull sutures with delayed closure, and hypolastic clavicles. At the age of 10y he presented in an Orthopedic clinic with short stature and lordosis. He also suffered of persistent first dentition which required surgical procedures. Skull X ray revealed Wormian bones and chest X ray show shortening of the lateral parts of both clavicles . A clinical diagnosis was made of Cleidocranial dysplasia (CCD). Molecular genetic testing sequencing the RUNX2 gene with fragment length analysis of the poly(GAG) repeats showed that our patient has 29 (GAG) repeat in one allele and 11th (GAG) repats in the second allele. The encoded protein RUNX2 is a transcriptional factor involved in skeletal morphogenesis. RUNX2 is essential for the osteoblast differentiation and chondrocyte maturation during fetal development. It has been previously suggested that expansion of polyalanine repeats may inhibit protein function. However, the exact mechanism of cellular toxicity for polyA-expansions is not known. One previous report of an affected family has been described, with minor clinical findings of CCD segregating with a 27 repeats expansion of a polyalanine stretch. Our patient has an expansion of 29 GAG repeats of the polyalanine stretch in the RUNX2 gene and a severe phenotype of Cleinocranial dysplasia.

3175W

Novel Mutation in the ADAMTSL4 gene in a child with bilateral Ectopia Lentis and Aortic Root Dilatation: expanding the phenotype? *R.L. Sanchez[1], Z. Ammous[2], P. Jayakar[1,2].* 1) Miami Children's Hospital, Miami, FL; 2) University of Miami, Miami, FL.

Ectopia lentis is a highly heterogeneous entity presenting as subluxation of the lens due to disruption in the zonular fibers. It can be an isolated finding but it is frequently seen in association with disorders such as Marfan syndrome, Weill-Marchesani syndrome, or Homocysteinuria. Isolated autosomal dominant ectopia lentis is caused by mutations in the FBN1 gene while autosomal recessive cases are due to ADAMTSL4 mutations. ADAMTSL4, a member of the ADAMTS like gene family, codes for a disintegrin and metalloproteinase with thrombospondin motifs. Mutations in this gene are associated with several congenital ocular anomalies, usually unrelated to other systemic abnormalities suggestive of connective tissue disorders. However a mild evidence of connective tissue instability was previously proposed but no phenotype-genotype correlation has been evident. We report the case of a 10 year old female born to non-consanguineous parents who presented with bilateral lens dislocation, mild aortic root dilatation, scoliosis, and flat feet hinting towards a connective tissue disorder. FBN1, TGFBR1, and TGFBR2 full gene sequencing was negative as well as a basic work-up that included a chromosomal microarray analysis and metabolic tests. ADAMTSL4 mutation analysis revealed a compound heterozygous state for two discrete changes. The first is c.2021_2022delCT in exon 12 known as a disease causing mutation. The second is a c.2977C>T transition in exon 18 which has not been previously reported. ADAMTSL4 is expressed in many tissues but it has been presumed that its role is unique for the composition of zonular fibers of the lens due to absence of other systemic manifestations in previously reported cases. Little is known on the exact function of ADAMTSL4. Moreover, mutations in other members of the ADAMTSL superfamily give rise to other connective tissue disorders. Our patient not only had bilateral lens dislocation but also mild aortic root dilatation, flat feet and scoliosis suggestive of a systemic connective tissue involvement. In addition the mutations detected in our patient are rare and one of them is a novel mutation. Thus there may be an abnormal gene product that is affecting other systems in addition to the eye. Given the unknown consequences on posttranscriptional level we cannot predict the phenotype. What was previously known as ADAMTSL4-related isolated ectopia lentis could be rather a systemic connective tissue disorder yet to be named and defined.

3176F

Mutations in DMD gene identified during clinical evaluation of patients with autism and/or global developmental delay. *J. Moeschler[1, 2], S. Upton[1], J. Ozmore[1].* 1) Dept Pediatrics, Gen, Dartmouth-Hitchcock Med Ctr, Lebanon, NH; 2) Geisel School of Medicine at Dartmouth, Hanover, NH.

We describe a series of four patients referred for global developmental delay (GDD) or autism spectrum disorder (ASD) who were ultimately diagnosed with mutations in the DMD gene. In two cases, the DMD mutations were identified by chromosome microarray analysis (CMA). In one boy referred for GDD, DMD mutation was identified by Next Gen sequencing. In a third case, a boy with ASD, signs and symptoms for Duchenne muscular dystrophy were noted and supported by marked elevation of serum CK. Case 1. 2 year 3 month old boy referred for GDD and short stature. Developmental ages at 32 months ranged 10–15 months. No history of weakness. Diagnostic evaluation included normal microarray but on 92-gene panel for X-linked intellectual disability demonstrated a stop codon in DMD gene: c.9403C>T (pQ3135X). On this basis, serum CK was obtained with elevation of 18,780 U/L. This patient had no calf psuedohypertrophy, negative Gower sign nor signs of weakness. Case 2. 2 year 9 month old boy referred for ASD, GDD and hypospadias. Developmental delays first noted at 9 months of age. ASD diagnosed in second year of life. Walked late at 24 months. As part of the diagnostic evaluation, CMA demonstrated two pathogenic deletions: 1p36.21-1p36.13 of 1.42 Mb and containing 33 OMIM genes; and, Xp21.1 involving DMD gene. Sequencing of DMD gene demonstrated deletion of exons 49–51. Serum CK was mildly elevated at 351 U/L. Examination was negative. Case 3. 11-year healthy boy whose sister was evaluated for GDD at another center and noted on CMA to have Xp21.1 deletion involving the DMD locus. Asymptomatic mother also had this deletion. Her 11-year old son was evaluated after serum CK elevation of 1081 U/L. This boy had normal neuromuscular exam. Case 4. 4 year 6 month old boy with autism referred for diagnostic evaluation. Mother noted delays and differences in first year of life. The diagnosis of ASD was established in his 3rd year. There was a six-month history of weakness. When seen on the first visit, Gower sign and other signs of weakness were present. Serum CK was 11,456 U/L. Deletion of exon 46 of DMD gene noted on deletion/duplication analysis. These cases demonstrate that: mutations in DMD gene may participate in the pathogenesis of GDD and ASD; “Secondary findings” challenge us to make proper clinical interpretation of genotype-phenotype relationships; and, serum CK screening in all boys presenting with global developmental delays and/or autism be done.

3177W

Copy number variants in monozygotic twins with neurofibromatosis 1. *E. Schorry[1], E. Sites[2], D. Viskochil[3], D. Stevenson[3], N. Ullrich[4], T. Smolarek[1], L. Martin[1].* 1) Div Human Gen, Cincinnati Childrens Hosp, Cincinnati, OH; 2) Children's Memorial Hospital, Chicago, IL; 3) Univ. of Utah, Salt Lake City, UT; 4) Children's Hospital, Boston, MA.

The gene for neurofibromatosis type 1 (NF1) displays extremely variable expressivity. Causes for this variable expression, which occurs within members of the same family and even within monozygotic (MZ) twins, are not yet well understood. We have undertaken a study of copy number variants (CNV) in monozygotic twins with NF1 to investigate possible differences in CNVs underlying the variable expression of NF1 features within twin pairs. We obtained DNA from blood and/or saliva from 11 pairs of proven MZ twins with NF1, who have been well characterized phenotypically and genotyped for NF1. Twin pairs were concordant for many traits such as pectus deformity, learning disability, and overall numbers of café-au-lait spots, but were discordant for other features including optic nerve glioma, scoliosis, plexiform neurofibromas, and MPNST. We tested samples of twins and their parents on the Illlumina 610K SNP-based CGH microarray, which detects microdeletions or microduplications of >80 kb. CNVs were compared within twin pairs and their parents to investigate for de novo changes or association with specific complications. RESULTS: All twin pairs analyzed contained CNVs, with a mean of 8.6 conservative CNVs per twin pair. On average, 3.6 CNVs per twin contained genes. No differences in CNVs to date have been found between co-twins. Several CNVs were located close to genes of potential interest in the Ras pathway, including PTPN20A/B and MAP2K3, and can be investigated as possible modifiers of NF1. One twin pair with severe involvement had a large duplication in chromosome 17 in trans to NF1, located 4.1 Mb from the NF1 gene, which we postulate could make the second allele more susceptible to rearrangement and hence tumorigenesis. SUMMARY: Our study of MZ twins with NF1 has identified several genes and genomic regions which may be of interest for future investigation as modifiers of NF1 phenotype.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3178F

A Novel Synonymous Mutation Causing Aberrant Splicing on the SLC26A4 Gene in a Korean Patient with a Hearing loss. *Y. Kim[1], J. Kim[1], J.Y. Choi[2], K-A. lee[1].* 1) Department of Laboratory Medicine, Yonsei University College of Medicine, Seoul, South Korea; 2) Department of Otorhinolaryngology, Yonsei University College of Medicine, Seoul, South Korea.

Mutations in PDS (SLC26A4) cause both Pendred syndrome (PS) and DFNB4, two autosomal recessive disorders that share hearing loss as a common feature. PS/DFNB4 is a genetically homogeneous disorder caused by bi-allelic SLC26A4 mutations. Here, we report a novel mutation (c.1803G>A, p.Lys601Lys) causing an aberrant splicing in a Korean boy who is clinically considered as PS/DFNB4 with congenital hearing loss and dilated vestibular aqueduct (DVA) in temporal bone computed tomography. After extracting DNA from whole blood using standard procedures, 21exons and flanking introns of SLC26A4were PCR amplified. In sequence analysis, we detected a known splicing mutation (c.919-2A>G, heterozygote), but the detection of this single known mutant on SLC26A4 allele was not sufficient to explain PS/DFNB4 phenotype of our patient. Thus, we considered a novel variant (c.1803G>A, p.Lys601Lys) detected in our patient in exon 16 of the SLC26A4 to be a disease causing mutation in our case. According to in silico analysis using the Berkeley Drosophila Genome Project (BDGP)(http://www.fruitfly.org/), the novel variant (c.1803G>A) affected canonical splice donor nucleotide position. To define the effect at transcript level of this novel 1803G>A variant, we have performed exon trapping.

3179W

Overlapping anophthalmia syndromes: SOX2, the new kid on the block. *A. Schneider, T. Bardakjian.* Dept Gen, Albert Einstein Med Ctr, Philadelphia, PA.

Anophthalmia is a rare birth defect which is thought to be on the severe end of a continuum of ocular anomalies with microphthalmia and coloboma on the milder end (MAC). MAC can occur as an isolated finding or associated with other anomalies. Many syndromes are known to be associated with these ocular anomalies. Two well described syndromes where anophthalmia may be seen are Septo-Optic-Dysplasia (SOD) and Lenz microphthalmia syndrome. Initial descriptions of mutations in the SOX2 gene were noted to be associated with bilateral severe eye findings including anophthalmia in most cases. The phenotypes of these three syndromes overlap significantly and as more individuals have been tested it is clear that the ocular phenotype associated with SOX2 mutations is very variable. Molecular diagnostics for SOX2 is clinically available and mutations are identified in up to 20% of individuals with bilateral ocular involvement. Molecular testing for SOD and Lenz syndrome is not as successful. We propose that perhaps the diagnostic paradigm should switch to one which utilizes molecular testing for SOX2 immediately in cases clinically thought to be SOD or Lenz syndrome with significant eye findings. Using data from the Anophthalmia / Microphthalmia Clinical Registry and DNA screening project at Einstein Medical Center in Philadelphia, we will demonstrate a significant proportion of individuals who carried a diagnosis of Lenz syndrome or SOD who were identified as having mutations in, or deletions of the SOX2 gene.

3180F

Distinct clinical and neurological features in a Korean boy with Schinzel - Giedion syndrome caused by a de novo *SETBP1* mutation. *J. M. Ko, B. C. Lim, K. J. Kim, Y. S. Hwang, C. H. Chae.* Pediatrics, Seoul National University Children's Hospital, Seoul, South Korea.

Schinzel - Giedion syndrome (SGS) is a rare multiple congenital malformation syndrome defined by characteristic facial features, profound developmental delay, severe growth failure, and multiple congenital anomalies including skeletal, renal, genital and cardiac malformations. Most affected individuals die in early childhood mainly due to progressive neurodegeneration and respiratory failure. The causing gene of SGS, SETBP1, has been identified quite recently, and there has been no report of Korean SGS patients with molecular confirmation. In this study, we present a 6-month-old boy with SGS complicated by epilepsy and profound developmental delay. Typical facial features, multiple congenital anomalies, and associated neurological findings could lead to the clinical diagnosis of SGS. The SETBP1 gene analysis revealed a pathologic mutation, p.Gly870Ser, and our patient was molecularly confirmed as SGS first in Korea. Although this syndrome is extremely rare, it is important to consider SGS in the differential diagnosis of infantile-onset neurodegenerative conditions, especially in patients with intractable epilepsy.

3181W

Crisponi Syndrome in a Turkish newborn : a possible founder mutation in the CRLF1 gene? *V. Benoit[1], P. Hilbert[1], M. Deprez[2], A. Charon[3], I. Maystadt[1], S. Moortgat[1].* 1) Centre de Génétique Humaine, Institut de Pathologie et de Génétique, Charleroi, Belgium; 2) Département de Neuropédiatrie, Grand Hôpital de Charleroi, Charleroi, Belgium; 3) Département de Néonatologie, Grand Hôpital de Charleroi, Charleroi, Belgium.

Crisponi syndrome (CS) is a rare, severe, autosomal recessive disorder characterized by congenital contractures of the facial muscles in response to slight tactile stimuli or during crying. The phenotype is complicated by intermittent hyperthermia unrelated to infectious agents, and by respiratory and feeding difficulties. Febrile crises frequently lead to death within the first months of life. Mutations in the cytokine receptor-like factor 1 (CRLF1) gene have recently been described in association with CS. We report here a Turkish patient with clinical presentation and dysmorphic features consistent with CS. Sequencing analysis of the CRLF1 gene showed a non-sense homozygous mutation (c.829C>T, p.Arg277X). This mutation has already been reported in another patient from Eastern Turkey. Considering that our family comes from the central part of this country, we could hypothesize that c.829C>T is an ancient founder mutation. Interestingly, the patient previously described by Okur et al. presents with additional features of velopharyngeal insufficiency, incomplete cleft palate and thin corpus callosum, which were not present for our patient, highlighting the difficulty to establish a phenotype-genotype correlation.

3182F

Interference with IS-PCR assays for Inv22 Testing of Hemophilia A. *T. Lewis[1], G. Pont-Kingdon[1], Y. Louie[2], J. Swenson[2, 3].* 1) ARUP Institute for Clinical and Experimental Pathology, Salt Lake City, Utah; 2) ARUP Laboratories, Salt Lake City, Utah; 3) Department of Pathology, University of Utah Health Sciences Center, Salt Lake City, Utaha.

Introduction: Hemophilia A is a common bleeding disorder affecting 1/5000 male newborns caused by mutations in the F8 gene. In severe hemophilia A (<1% factor VIII), an F8 intron 22-1 inversion is the causative gene defect in approximately 48% of the cases. A 9.5 kb sequence (int22h-1) occurring in intron 22 of the F8 gene and again twice more at telomeric positions allow for intrachromosomal homologous recombination, disrupting the gene. The cascade strategy for individuals affected with severe hemophilia A, or a family history of such, is targeted inversion analysis, followed by full gene sequencing, and finally deletion/duplication analysis. For inversion analysis, many diagnostic laboratories utilize the inverse shifting-PCR (IS-PCR) based approach described by Rossetti et al. in 2005 or the more discriminatory IS-PCR assay detailed by Rossetti et al. in 2008. Description: An asymptomatic African-American female, with a family history of severe hemophilia A, was submitted for carrier testing. Utilizing the IS-PCR strategy, inversion testing revealed apparent homozygosity for the inversion 22 band. As this result conflicted with the clinical information, further investigation was undertaken. Additional testing indicated the presence of two X chromosomes, no observable deletion within the inversion primer regions, and no point mutations under the primers or restriction site. A database search for SNPs in the 21kb wild-type Bcl1 fragment was positive for 144 SNPs, 26 of which were common. One of the common SNPs, rs73563631, creates a novel BclI restriction site. Primers were designed to amplify the region containing this variant in our patient and the product sequenced; the results indicated that our patient was positive for the SNP variant. Implications: The creation of a novel BclI restriction site in intron 22 of F8 by the SNP rs73563631 results in two BclI fragments. Ligation of the fragments creates two products, neither of which contains both primer sites for the inverse shifting-PCR assays. The effect is such that the wild-type product will not be observed in an inversion carrier female or a non-inversion male. This was the case in an additional patient, an African-American male, who had no observable bands on the inversion 22 assay. In a non-inversion female, the wild-type band will be seen but will be mistakenly interpreted as two copies. The SNP rs73563631 has an overall population frequency of 1% but is observed in 4.5% of Africans.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3183W

Complex Ear Abnormalities, Choanal Atresia, Coloboma and Renal Hypoplasia in a Patient with Mutations in *CHD7* and *EYA1* and Microdeletion of 2q23.1. *R. Badilla-Porras[1], L. Dupuis[1], T. Stockley[2], D.J. Stavropoulos[2], R. Mendoza-Londono[1].* 1) Division of Clinical and Metabolic Genetics, Hospital for Sick Children and University of Toronto, Toronto, Canada; 2) Division of Molecular Genetics, Department of Paediatric Laboratory Medicine, The Hospital for Sick Children and University of Toronto, Toronto, Canada.

Abnormal ear morphology is a feature of many genetic syndromes. In combination with iris coloboma and renal hypoplasia, it raises concern for three common diagnoses: CHARGE syndrome, Branchio-Oto-Renal syndrome (BOR) and chromosomal aberrations. CHARGE syndrome is an autosomal dominant condition characterized by coloboma, heart defects, choanal atresia, retarded growth and development, genital hypoplasia and ear anomalies/deafness. *CHD7* is the only gene currently known to be associated with CHARGE syndrome. BOR syndrome is characterized by malformations of the outer, middle, and inner ear associated with conductive, sensorineural, or mixed hearing impairment. Patients have branchial fistulas and cysts, and renal malformations ranging from mild renal hypoplasia to bilateral renal agenesis. Molecular genetic testing of the *EYA1* gene detects mutations in 40%; of individuals with this clinical diagnosis. We present the case of a female that developed perinatal oligohydramnios. At birth she was found to have bilateral choanal atresia, cupped ears with middle and inner ear abnormalities (absent semicircular canals and hypoplastic oval window), bilateral chorioretinal colobomas, atrophic optic nerves, facial asymmetry, right preaxial polydactily and hypoplastic kidneys and external genitalia. Family history was positive for branchial cyst and duplication of renal collecting system in the patient's father. Sequencing of *CHD7* revealed a pathogenic *de novo* mutation in exon 11 of the *CHD7* gene c.2905_2906del (p.Arg969fs), confirming a diagnosis of CHARGE syndrome. Given the family history, we performed sequencing of *EYA1* in the father and identified a mutation in exon 15 c.1460C>T (p.Ser487Leu) confirming a diagnosis of BOR. This mutation was inherited by his daughter. Array comparative genomic hybridization revealed a paternally inherited 0.059 Mb microdeletion on chromosome 2q23.1 encompassing intronic sequences of the MBD5 gene. At this time the developmental implications of the microdeletion and the contribution to the overall phenotype in this patient is being investigated. To the best of our knowledge, our patient is the first example of mutations in *EYA1* and *CHD7*and a chromosomal microdeletion in the same individual. This case highlights the clinical variability and overlap in the phenotype of patients with CHARGE and BOR syndromes. Detailed analysis of the phenotype demonstrates the additive effect of these two mutations on normal development.

3184F

Severe CHST3 mutations in two Brazilian families with spondyloepiphyseal dysplasia with congenital joint dislocations. *W.A.R. Baratela[1], T.F. Almeida[1], G.L. Yamamoto[1], J.H. Marques[1], O. Letaif[2], A.C. Pereira[3], C.A. Kim[1], D.R. Bertola[1].* 1) Unidade de Genética, Instituto da Criança - Hospital das Clínicas - Universidade de Sao Paulo, Sao Paulo, SP, Brazil; 2) Orthopedics Department, Hospital das Clínicas, Universidade de Sao Paulo, Sao Paulo,SP, Brazil; 3) Laboratório de Genética e Cardiologia Molecular, InCor, Universidade de Sao Paulo, SP, Brazil.

Spondyloepiphyseal dysplasia with congenital joint dislocations (SED-CJD) has been previously named as recessive Larsen syndrome, humerospinal dysostosis, spondyloepiphyseal dysplasia - Omani type, and chondrodysplasia with multiple dislocations, and is characterized by dislocations of elbows, hips, and knees at birth, with typical lumbar vertebrae changes, normal carpal age, recessive inheritance, and absence of facial flattening. Carbohydrate sulfotransferase-3 deficiency caused by CHST3 mutations is responsible for SED-CJD. We describe 3 individuals from 2 distinct consanguineous families from Brazil, with SED-CJD, and positively tested for CHST3 mutations. Patient 1 is an 8y4mo boy, born from a first cousin couple, by cesarian section, 2kg of BW, 44cm of BL, with dislocated knees and clubfeet. At current age he is 108 cm tall (<5th%ile), camptodactyly, bilateral radioulnar subluxation, dextroscoliosis are seen, and his knees are still dislocated. Patient 2 is the youngest sister, with 4y4mo. Positive nuchal translucency along with oligohydramnios, were seen during prenatal ultrasound. She was born preterm, cesarian section, BW 1910g, BL 35cm. Bilateral hip and right knee dislocation, and right clubfoot. At her last follow up, she was 81cm tall (<5th%ile), OFC was 48cm, and she had phalangeal, elbows, and knee joints contractures. Patient 3 is a 3y11mo boy, born from a first cousin couple, at term, BW 2500g, BL unknown, and OFC 34cm. He had bilateral hip and knee dislocation, and left side clubfoot. At current physical exam he is 81cm (<5th%ile), 50cm OFC, brachycephaly, opalescent and brittle teeth, and small ears are seen. Skeletal surveys of the 3 patients were strikingly similar, with a mildly progressive epiphyseal involvement. There is an increase and posterior decrease, in the interpedicular distance and vertebral endplate irregularities. Molecular studies for the 3 patients were directed for CHST3 mutations analysis. The 2 siblings had a IVS2-1G>C and c.141G>C homozygous mutation, with R47S substitution and putative missplicing. Patient 3 showed a c.718A>T p. K240X, nonsense mutation at homozygosity. The two families present a typical skeletal involvement, with dislocations at different sites evident at birth and severe mutations, giving further support to the fact that profound effects on CHST3 protein are responsible for the full-blown phenotype.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3185W

Computer-aided facial recognition of individuals with FG (Opitz-Kaveggia) syndrome caused by p.Arg961Trp mutation in MED12. *L. Basel-Vanagaite[1,2,3], L. Karlinsky[2], L. Wolf[2,3], M. Shohat[1,2,3], C. Skinner[4], C. Rogers[4], R. Stevenson[4], C.M. Schwartz[4], J.M. Graham, Jr.[5].* 1) Schneider Children's Medical Center of Israel, Raphael Recanati Genetics Institute, Rabin Medical Ctr and and Felsenstein Medical Research Center, Petah Tikva, Israel; 2) FDNA Ltd., Herzlyia, Israel; 3) Tel Aviv University, Tel Aviv, Israel; 4) J.C. Self Research Institute, Greenwood Genetic Center, Greenwood, SC, USA; 5) Medical Genetics Institute at Cedars-Sinai Medical Center, Los Angeles, CA, USA.

FG (Opitz-Kaveggia) syndrome is an X-linked recessive syndrome characterized by relatively large head, frontal upsweep, hypertelorism, ptosis, small ears, broad and flat thumbs, imperforate anus, hypotonia, agenesis of the corpus callosum and moderate intellectual disability. The classical form of FG syndrome is caused by a recurrent p.Arg961Trp mutation in the MED12 gene. FG syndrome is misdiagnosed and recent literature confirmed that several patients with similar features had microdeletion/microduplication syndromes and syndromes caused by FLNA mutations. With the advent of novel automatic face analysis techniques, our ability to analyze facial morphology from photographs has improved significantly. In this study we examined whether a computer-based dysmorphological analysis can be used in order to discern between FG patients and non-FG patients that present superficially similar facial features. For this, we used a collection of 18 genetically verified FG cases and 18 non-FG cases that were clinically suspected to be FG but do not carry mutations in the MED12 gene. A modern face analysis system that was developed specifically for dysmorphological analysis was used. The system is fully automatic and starts by detecting the face in the image. Then, 130 facial fiducial points are localized and various measurements are taken. The final classification is based on these measurements as well as on a global “gestalt” detector that estimate the probability of the subject having FG based on the appearance of the entire facial image. A statistical technique called cross validation was used to estimate the recognition capability of the computer system. At each one of 20 rounds, the data was split randomly to training and testing data, each comprising of 50% of the samples. The system was trained to separate between the two classes using the train dataset, and evaluated on the test data. The statistical power of the test in correct recognition of FG patients with the p.Arg961Trp mutation was estimated using the conventional recognition metric of average area under the ROC curve and was 90%, which is considered very high. The gestalt of the face as captured by multiple local patterns of facial texture contributed significantly to the correct recognition of the individuals with FG syndrome. Overall, we have demonstrated that computer-based analysis can be successfully used in supporting experts for the correct recognition of patients with FG syndrome.

3186F

Hand abnormalities in Loeys-Dietz syndrome: Expanding the clinical spectrum. *B. Chung[1], A. Hinek[2], T. Bradley[3], L. Grossse-Wortmann[3], S. Blaser[4], D. Chitayat[5,6].* 1) Paediatrics & Adol Med, The University of Hong Kong, Hong Kong; 2) Division of Cardiovascular Research, The Hospital for Sick Children, Toronto; 3) Division of Cardiology, The Hospital for Sick Children, Toronto; 4) Division of Neuroradiology, The Hospital for Sick Children, Toronto; 5) Division of Clinical and Metabolic Genetics, The Hospital for Sick Children, Toronto; 6) The Prenatal Diagnosis and Medical Genetics Program, Mount Sinai Hospital, Toronto.

Loeys-Dietz syndrome is an autosomal dominant connective tissue disorder characterized by clinical manifestations including hypertelorism, bifid uvula and cleft palate, and arterial tortuosity. We report a patient with LDS with a mutation in the TGFBR2 gene who presented prenatally with clenched fist and clubbed feet and postnatally was found to have digital abnormalities including brachydactyly, camptodactyly, partial syndactyly and absent phalanges. The proband was the first baby born to a healthy non-consanguineous couple. The pregnancy was initially uneventful and fetal ultrasound at 21 wks showed bilateral club feet, small penis, mild right pelviectasis and a clenched right hand with abducted thumb. Amniocentesis showed a normal male karyotype and normal aCGH. Delivery was at 39 weeks gestation and uncomplicated. His birth weight was 4.490kg (>+2SD), the length was 50cm (mean) and his OFC was 39cm (> +2SD). He had ridges over multiple sutures, hypertelorism, retro-micrognathia, umbilical hernia, a right inguinal hernia and clenched hands. His left hand had a proximally inserted thumb and 3rd-5th finger camptodactyly. His right hand had a small, proximally inserted thumb with a hypoplastic distal phalanx. The 2nd- 4th fingers were bi-phalangeal with symphalangism and the 5th finger had a short mid-phalanx. There was skin syndactyly connecting the 2nd-3rd and 3rd-4th fingers. There were bilateral club feet, abnormal cervical vertebrae and mild thoracic scoliosis. Echocardiography showed a quadricuspid pulmonary valve, aortic root and main pulmonary artery dilatation and a large ductal aneurysm. Mutation analysis of the TGFBR2 gene showed a known pathogenic mutation c.1583G>A/p.Arg528His which confirmed the diagnosis. Subsequent imaging demonstrated generalized arterial tortuosity. Immunohisto-chemical staining of his skin fibroblasts showed intracellular accumulation of collagen type I in the presence of otherwise normal elastic fiber production, compatible with LDS associated with TGFBR2 mutations. The most common clinically important skeletal abnormalities in LDS include cervical vertebrae anomalies and instability, spinal cord abnormalities and foot deformities [Erkula et al., 2010]. Syndactyly and brachydactyly have not been reported in LDS and expand the spectrum of this condition. Recognition of all of the skeletal changes in LDS is important as they can be the earliest manifestation in this multi-system connective tissue disorder.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3187W

Gorlin Syndrome: Three unrelated female Mexican cases. *NO. Davalos[1,4], SA. Alonso Barragan[2,5], IM. Salazar-Davalos[3], MA. Aceves-Aceves[3], SA. Ramirez-Garcia[1], LR. Topete -Gonzalez[1,4], DA. Fierro-Rodriguez[1,4], DM. Prado[1,4], AR. Rincon-Sanchez[1], IP. Davalos[1,5], D. Garcia-Cruz[1].* 1) Institut of Human Genetics, Centro Universitario de Ciencias de la Salud,, Universidad Guadalajara, Guadalajara, Mexico,; 2) Doctorado en Genetica Humana, CUCS UdeG; 3) Facultad de Medicina, CUCS, U de G; 4) Genetica, Hospital Regional VGF, ISSSTE; 5) Genetica, Centro de Investigacion Biomedica, CIBO IMSS, Guadalajara Jalisco Mexico.

Introduction: Gorlin syndrome (GS) or Nevoid basal cell carcinoma syndrome (NBCCS), is charaterized by basal cell nevi, facial dysmorphology, odontogenic keratocysts and skeletal anomalies, is an autosomal dominant inheritance disease, with 97% penetrance and variable expressivity; about 40% of cases represents de novo mutations. The prevalence reported,1/57,000, with a sex ratio M:1:F:1. Report: We present three unrelated female cases, with other clinical findings not previously reported. Case I: 10 years-old, with family history of GS. At 11 months-old umbilical hernia is diagnosed. At 2 years-old, was diagnosed adduct foot and genu valgus. Oral findings shown dental malocclusion, within evidence of jaw cysts. Skin with nevi in face neck and multiple plantar pits detected. Dyslexia and intellectual disability presents. The CT scan revealed frontal-cortical atrophy and dilated lateral ventricles due to atrophy. Case II: 19 years-old. At 15 years-old detected nevi on scalp, removed and histolgically shown basal cell carcinoma. At 16 years-old detected maxilar and mandibular cysts. Radiologically shown calcification of the falx cerebri, broad ribs; calcification of vertebral bodies and rotation at T1-T3 and bifid spine. Case III: 40 years-old. At 10 years old maxilar and mandibular cysts diagnosed. Ophthalmologically presents right eye chalazion with sebaceous secretion. The proposita became pregnant three times, which ended in abortion, detected by US bicournate uterus. CT scan shown posterior fossa calcification, on straight sinus and cerebellar vermis that resemble a “Y” image; cortical atrophy detected and subcortical vascular lesions of the small vessels; mild hydrocephalus. Neuropsychological assessments reported, pause speech alterations. Discussion. We diagnosed three unrelated female cases with Gorlin Syndrome. Shown the clinical and radiological findings of the syndrome. In our cases we presented clinical and radiological findigs not described in Mexican cases, such as, cortical atrophy, presents in case II and III, multiple cortical calcifications. Neuropsychological abnormalities, dyslexia and neuropsychomotor deficit are an uncomun findings; as well as rotation of the lumbar vertebral bodies, that could be diagnosed in childhood. The bicornuate uterus could be the response of the repetitive abortions. This cases shown the wide spectrum of the GS, and suggests a detailed multidisciplinary study for search new findings in patients with GS.

3188F

Diamond-Blackfan Anemia: A Case Report and Seven-Year Follow Up. *M.G. González-Mercado[1,2], L. Bobadilla-Morales[1], N.O. Dávalos[1], D. García-Cruz[1], A. Corona-Rivera[1], C.E. Monterrubio-Ledezma[1], I.M. Salazar-Dávalos[3], A. González-Mercado[1,2], M.A. Aceves-Aceves[3], I.P. Dávalos[1,2].* 1) Doctorado en Genética Humana, Instituto de Genética Humana, CUCS, Universidad de Guadalajara; 2) División de Genética, Centro de Investigación Biomédica de Occidente, IMSS; 3) Facultad de Medicina, CUCS, Universidad de Guadalajara, Guadalajara, Jalisco, México.

Diamond-Blackfan Anemia (DBA) (OMIM #105650) is an autosomal dominant entity characterized by congenital erythroid aplasia and in 30–40% of patients have other congenital anomalies. Approximately 10–25% of DBA cases are familial. Genetic heterogeneity has been reported in DBA and it is related to *RPS19* gene located on 19q13.3. Blood test could show high adenosine deaminase activity (ADA). **Case report:** The proposita aged 7 years old, was born from non-consanguineous parents, mother apparently healthy, the father was referred as affected with anemia. At birth was detected with acrocyanosis and the echocardiogram reported left ventricular dilatation with mild dysfunction and myopericarditis, also she presented hepatomegaly and decompensated anemia, treated with blood transfusion in two occasions. At 3 months she had recurrence of severe anemia without a primary cause (Hb 3.4, Hto 9.9) treated with blood transfusion. A bone marrow analysis was reported with decrease of the dysplastic erythroid normoblasts with some asynchrony in maturation. White blood cells reported all stages of maturation of normal appearance and an appropriate platelet production. At 7 months, a haematological study reported no alteration of Hb or red cell membranes and the diagnosis of pure red cell aplasia was confirmed and was started treatment with prednisone. Physical examination at 4 years 11 months she presented weight of 12kg (<p3), height of 99cm (p10) and OFC of 46.5cm (<p3), generalized pallor, microcephaly, small forehead, sparse eyebrows, almond shaped eyes, large philtrum, thin lips, micrognathia, thenar hypoplasia and digitized thumbs. The psychomotor development was normal. At 7 years presented weight of 16kg (<p3), height of 110cm (p10) and OFC of 47.7cm (<p3). The karyotype and chromosomal instability study were normal. The ADA quantification was 17.8 U/L (0-15U/L). She has been treated with prednisone, folic acid and B complex until now. **Conclusions:** The clinical and laboratorial studies lead us to diagnose DBA in the present case. Differential diagnosis includes ribosomopathies, Shwachman-Diamond syndrome and Fanconi anemia. The treatment in DBA is based on corticosteroids, blood transfusion therapies and stem cell transplantation, although it is known that only bone marrow transplant is a cure for DBA. This patient has been treated with steroids and blood transfusions with a favorable outcome.

3189W

Familial Beare-Stevenson cutis gyrata syndrome in the absence of craniosynostosis. *E. Leon[1], J. Jung[2], M. Ririe[2], D. Stevenson[1].* 1) Genetics, University of Utah, Salt Lake City, UT; 2) Dermatology, University of Utah, Salt Lake City, UT.

Beare-Stevenson cutis gyrata syndrome consists of skin furrows of corrugated appearance, acanthosis nigricans, craniofacial anomalies (particularly craniosynostosis and ear anomalies), anogenital anomalies, skin tags, and prominent umbilical stump. Cutis gyrata variably affects the scalp, forehead, face, preauricular area, neck, trunk, hands, and feet. Craniosynostosis is present in most of the cases and only rare cases have been reported with normal skull development. This syndrome has been reported to be caused by mutations in the *FGFR2* gene. We report a 1-month-old baby girl (proposita) and her 29-year-old mother with variable degree of cutis gyrata and normal cranium. The proposita was born full term to a 29-year-old primagravida mother and a nonconsanguineous 31-year-old father. Prenatal history was remarkable for preeclampsia. Delivery was uncomplicated. Apgar scores were 7/8. Growth parameters: birth length (50th centile), weight (10th centile) and head circumference (75th centile). At her first evaluation a skin tag on the scalp was noticed along with skin furrows of corrugated appearance with vertical orientation on the forehead, and circular skin furrows on posterior aspect of her scalp. Her ears had multiple preauricular creases, large lobules, and prominent antitragus. A nevus flammeus on forehead and deep creases on palms and soles were also present. Her mother, who is cognitively normal, had the same skin finding on the forehead when newborn which are now attenuated as well as the preauricular creases. *FGFR2* gene testing is pending. The above family represents a likely autosomal dominant cutis gyrata condition. The skin tags and location and appearance of the cutis gyrata are strikingly similar to the cases of Beare-Stevenson syndrome reported in the literature. We propose that these two individuals are part of a broader phenotypic spectrum of this rare disorder. Few patients with Beare-Stevenson syndrome have been tested for *FGFR2* gene analysis and to our knowledge none of the patients with normal craniums had genetic testing. In addition, this family would be the first reported familial case of Beare-Stevenson syndrome. Documentation of a pathogenic mutation in *FGFR2* in our patients will provide evidence of phenotypic variability. However, at this time genetic heterogeneity cannot be excluded.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3190F

Report of a novel mutation in the SLC26A2 gene found in a Colombian adult patient with Diastrophic Dysplasia. *T. Pineda[1], A. Rossi[2], L. Bonafé[3], A. Superti-Furga[3], H. Velasco[4].* 1) Genetics Institute, National University of Colombia, MD; 2) Department of Biochemistry "Alessandro Castellani", University of Pavia, Pavia, Italy; 3) Centre Hospitalier Universitaire Vaudois, Lausanne, Switzerland; 4) Morphology Department, Genetics Institute, National University of Colombia, MD, MSc,E.

Diastrophic Dysplasia is an osteochondrodysplasia belonging to the group of dysplasias caused by mutations in the diastrophic dysplasia sulfate transporter. This syndrome is a micromelic dysplasia with multiple bone deformities of the hands, feet, knees and spine. In this paper we report the case of a Colombian adult patient with diastrophic dysplasia whose clinical diagnosis was confirmed at the molecular level. In this first report of diastrophic dysplasia in Colombia we found that the patient was compound heterozygote for the already reported Arg279Trp substitution and an unpublished mutation, a Ser157Thr substitution in the SLC26A2 gene. Bioinformatic analysis on the latter mutation suggested that it could correspond to a deleterious mutation because it is found in a highly conserved domain of the sulfate transporter.

3191W

Intragenic NSD1 duplication of exons 14–17 in a patient with Sotos syndrome. *S. Ramanathan, R.D. Clark.* Pediatric Genetics, Loma Linda University, San Bernardino, CA.

We describe the second report of Sotos syndrome caused by an intragenic duplication in the NSD1 gene. In Sotos syndrome, characteristic facial features, overgrowth, and developmental/cognitive disabilities are caused by haploinsufficiency of the NSD1 gene. Gene sequencing detects mutations in 83–90% of clinically diagnosed individuals. Microdeletions of chromosome 5q35, encompassing the NSD1 gene, are found in 10% of non-Japanese patients. Partial NSD1 gene deletions of one or more exons are reported in about 5% of patients. Intragenic duplications, however, are rare. Our patient, now a 2-year old Hispanic male, was the product of a pregnancy notable for a 6 lb weight gain. He was born at 40 4/7 weeks by vacuum-assisted vaginal delivery with a birth weight of 4.77 kg, birth length 61 cm and head circumference 38 cm, all of which were over the 98th percentile. He was in the NICU for 3 weeks for respiratory distress and tremors. He had a tall forehead with dolicocephaly, downslanting palpebral fissures with puffy periorbital tissue and hypertonia. Increased extra-axial space, seen on brain MRI, later resolved. An abdominal ultrasound was normal, with no organomegaly. Bilateral strabismus has resolved. The diagnosis of Sotos syndrome was suggested by his clinical features. Sequencing of the NSD1 gene was negative. Chromosome analysis and oligo-SNP microarray were normal. MLPA analysis of the NSD1 gene detected an intragenic duplication of exons 14–17. Parents declined testing, but have no suggestive features of Sotos syndrome. The patient continues to be over the 98th percentile on all growth parameters with global developmental delays. We found only one published report of Sotos syndrome in a patient with an intragenic duplication of exon 4 (Saugier-Veber P et al. 2007). There is also one report of a 4-year-old girl with clinical features of Sotos syndrome with a 264 kb duplication on chromosome 5q35, downstream of, but not including, the NSD1 gene (Kasnauskiene J et al., 2011). On the contrary, patients with microduplications that include the entire NSD1 gene, spanning 650 kb to 6.5Mb, have microcephaly and growth retardation (Zhang H et al., 2011). In our patient, a duplication of exons 14–17 produced the classic findings of Sotos syndrome. This case illustrates the importance of including analysis for NSD1 deletions or duplications when gene sequencing and microarray studies are negative.

3192F

Craniosynstosis as a bone anomaly of Kabuki syndrome. *A. Shimada[1], I. Tamada[2], H. Yoshihashi[3].* 1) Div of General Practice, Tokyo Metropolitan Children's Hos, Fuchu-shi, Japan; 2) Div of Plastic and Reconstructive Surgery, Tokyo Metropolitan Children's Hos, Fuchu-shi, Japan; 3) Div Med Gen, Tokyo Metropolitan Children's Hos, Fuchu-shi, Japan.

Introduction: Kabuki syndrome(KS) is acongenital syndrome characterized by distinctive facial features, skeletal anomalies, persisting of fetal fingertip pads, postnatal growth deficiency, and intellectual disability. MLL2 (12q12-q14) was identified as causative gene of KS in 2010. As noted above, skeletal anomalies are one of the characteristic features of KS. Although various skeletal anomalies, such as vertebral defects and rib anomalies, were widely recognized in KS, craniosynostosisis had not been highlighted as comorbid skeletal anomaly of KS. There had been five cases of KS with craniosynostosis reported so far, but none of these cases were genetically diagnosed. We present two cases of KS with craniosynostosisand compare with the clinical features of previously reported cases. Material and methods:Case1. 5mth-male. 37w0d, 2630g. He had distinctive face, clinodactyly of fifth digits, developmental delay, postnatal growth deficiency, ASD, and hearing difficulties. A mutation analysis of MLL2 revealed a mutation in exon 39(c.12368C>A,S 4 123X). His head X-ray had showed mild craniosynostosis requiring no surgical treatments. Case 2. 16y-female.38w0d, 2860g. She had typical facial features, craniosynostosis, scoliosis, persistence of fetal finger tips,intellectual disability, postnatal growth deficiency, ASD, and left hydronephrosis. She was followed as craniosynostosis syndrome with multiple anomalies until 16 year-old. A surgical intervention for craniosynostosis was made in other hospital in 4mth-old.Despite of surgical treatments, her head deformity remained. Her parents brought her to our plastic surgeon and the diagnosis of KS was firstly made when she was 16 year old. Surgical treatmentis planned in our institution for cosmetic reason.Discussion:Craniosynostosis has not been recognized as striking structural anomalies in KS but it could be relatively frequent comorbid structural anomaly than it is noted before. KS should be considered when evaluating children with intellectual disabilities and postnatal growth deficiency in addition to craniosynostosis. Craniosynostosis in KS varies in severity as we experienced. Long-term follow up will be recommend since a surgical intervention is necessary in severe case. Further accumulation of cases and genetic study may provide insight into the clinical characteristics of craniosynostosis in KS.

3193W

The variable spectrum of SMAD4 mutations. *J. van den Ende[1], N. Van der Aa[1], T. Boiy[2].* 1) Dept Med Gen, Univ Hospital, Antwerp, Edegem, Belgium; 2) Dept Pediatrics, Univ Hospital Antwerp, Edegem, Belgium.

The SMAD4 gene on chromosome 18q21.1, previously known as MADH4, encodes a protein involved in signal transduction of the TGF-beta superfamily and Bone Morphogenetic Proteins (BMP`s). SMAD 4 comprises 11 coding exons and pathologic allelic variants in the different domains can have various consequences. Several syndromes have been described as a result of mutations in SMAD4. The SMAD4 gene was first identified as a tumor suppressor gene involved in human pancreatic carcinoma (Hahn et al. 1996) and later as one of the causes of Juvenile Polyposis Syndrome (JPS) (Howe, 1998). This syndrome is characterized by predisposition to hamartomatous polyps in the gastrointestinal tract, sometimes with malignant transformation. It is caused by mutations in BMPR1A or in SMAD4. About 15–22% of patients with JPS due to SMAD4 mutations also show symptoms of Hereditary Hemorrhagic Telangiectasia (JPS/HHT). HHT is characterized by the presence of multiple arteriovenous malformations, in the skin (telangiectasia), nose (frequent nose bleedings), lungs, liver and brain. The majority of cases is caused by mutations in the ENG gene or in the ALK1 gene. 1–2% Of persons with clinically HHT show a SMAD4 mutation. Probably variable expressivity and/ or age related penetrance play a role and all patients with a SMAD4 mutation might be at risk for both conditions. The most recent syndrome attributed to SMAD4 mutations is the Myhre syndrome, a developmental disorder with short stature, facial dysmorphism, muscular hypertrophy, deafness and developmental delay. Patients with this syndrome show heterozygous mutations in the MH2 domain of the SMAD4 gene, necessary for SMAD oligomerizaton and TGFB/ BMP signal transduction. We present 3 patients with SMAD4 mutations, one with JPS/HHT and 2 patients with Myhre syndrome, and discuss the role of SMAD4 in the different disorders.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3194F

A new frontonasal dysplasia syndrome associated with SIX2 deletion.

S.L. Zimmerman, H.M. Saal, Z.M. Ahmed, R.B. Hufnagel. Cincinnati Children's Hospital Medical Center, Cincinnati, OH.

Craniosynostosis is premature closure of the cranial sutures. Several genes including FGFR1, FRFR2, FGFR3, TWIST and SIX3 have been implicated. However, mutations of these genes have only rarely been implicated in sagittal suture synostosis. We present a 22 month old female with sagittal synostosis and dysmorphic features including frontal bossing, hypertelorism, ptosis with epicanthus inversus, wide lateral palatine ridges, depressed nasal bridge and a broad nasal tip. Genetic analysis showed a normal female karyotype or 46,XX and normal molecular studies for the genes most commonly known to cause craniosynostosis. However, SNP microarray identified a 108.3 kb deletion on chromosome 2p21 that included the SIX2 gene. The SIX2 gene is a homeobox containing transcription factor shown to play an essential and evolutionary conserved role during early development in mice, Xenopus and Drosophila. Significantly, her mother had similar dysmorphic features including, ptosis, dental diastasis, broad nasal tip, frontal bossing and was identified with having the same SIX2 hemizygous deletion. Studies in Six2-/- mice document similar craniofacial features and in addition renal dysplasia However, Six2 heterozygote mice did not show the craniofacial or renal defects. Our data supports that loss of the SIX2 function affects craniofacial development and lead to a novel frontonasal dysplasia syndrome. The specific mechanism of how the SIX2 gene affects craniofacial development is yet to be determined and further studies will include looking at regulation of SIX2 and a neighboring SIX3 gene in this patient's derived cell lines.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3195W

Importance of Orofacial Features and The Role of Dentistry in The Early Diagnosis of Developmental Disorders. *Z. Öncel Torun[1], D. Torun[2], K. Karaer[2], RO. Rosti[2].* 1) Dept. of Restorative Dentistry and Endodontics, Gulhane Military Medical Academy, Ankara, Turkey; 2) Dept. of Medical Genetics, Gulhane Military Medical Faculty, Ankara, Turkey.

Dental development disorders are seen as quite often in patients with developmental disabilities. But in the field of dentistry, patients with prominent craniofacial features are occasionally defined. Aim of the present case series was to provide oral data from three unrelated patients affected with cleidocranial dysplasia, pycnodysostosis and hallermann-streiff syndromes. All patients underwent regular dental investigation in the first examination. Unusual facial appearance and dental abnormalities has given rise to thought developmental impairment. Consultation with specialist in medical genetics and further laboratory assessments allow to the right diagnosis and treatment planning of the patients. This study suggests that dentist's familiarity with dental features of genetic diseases may affect the process of diagnosis and treatment of patients. In this process, consultation with clinical geneticist allows a valuable estimate of prognosis and the anticipation of complications.

3196F

Renin receptor/ATP6AP2 gene mutation: An X-linked cause of mental retardation, postnatal microcephaly and intractable seizures. *D. Chitayat[1,2], K. Sirewanda[2], R. Mendoza[2], S. Blaser[3], J. Raiman[2], J. Jessen[1], E. Donner[4], C. Schwartz[5].* 1) Prenatal Diag & Med Gen, Mount Sinai Hosp, Toronto, ON, Canada; 2) Department of Pediatrics, Division of Clinical and Metabolic Genetics, The Hospital for Sick Children, Toronto, Ontario, Canada; 3) 1Department of Pediatrics, Division of Neuroradiology, The Hospital for Sick Children, Toronto, Ontario, Canada; 4) Department of Pediatrics, Division of Neurology, The Hospital for Sick Children, Toronto, Ontario, Canada; 5) Greenwood Genetic Center, 106 Gregor Mendel Circle, Greenwood, SC, USA.

The (pro)renin receptor [(P)RR] is abundantly expressed in heart, brain, placenta and eye. A mutation in the(P)RR gene, resulting in frame deletion of exon 4[4-(P)RR] was reported in association with X-linked mental retardation (XLMR) and epilepsy pointing to a novel role of (P)RR in brain development and cognitive function. We report the second case of XLMR and seizures associated with a de novo mutation in the (P)RR gene. The patient was the first child born to a 31-year-old healthy couple of Ashkenazi Jewish descent. The couple was healthy and non-consanguineous and their family history was non-contributory. The pregnancy was uncomplicated and delivery was at 39 weeks via emergency C-section. He was noted to have mild facial dysmorphism including a short forehead, bitemporal narrowing, flat and broad nasal bridge, bilateral epicanthic folds, almond shaped eyes, long eye lashes, long philtrum, microretrognathia and high arched palate. His ears were prominent and he had multiple flat hemangiomatous lesions on the face and body. At two weeks of age he was noted to have episodes of eye blinking, head turning and upper and lower limb twitching lasting 5–15 seconds that occurred up to 10 times a day. At one month of age his head circumference (OFC) was at the mean, his weight at the 50th-75th centile and his length at the 50th centile. He was noted to be alert and active and moved all four limbs symmetrically. His muscle bulk and tone were normal and his DTRs were +2. During the next 6 months he developed intractable seizures and at 6 months he was hypotonic with complete head lag, was not interactive and was unable to roll over or grab objects. He did not fixate or track and his OFC was at -1.5SD. He developed limbs' spasticity, his intractable seizures continued and his OFC centiles decreased being <-2SD at 3 years. His initial MRI showed diffuse supratentorial cerebral volume loss and a repeat showed delay in myelination with significant cerebral atrophy. DNA analysis for the XLMR panel showed a de novo variant in the ATP6AP2 gene [c.301-11_301-10delTT]. Analysis of membrane associated-receptor molecular forms showed the presence of full-length and truncated proteins in the patient. Our findings confirm the importance of the renin-angiotensin system and the renin receptor in brain development and cognitive function.

3197W

SCN1B sequence variations in Iranian Patients with Epilepsy, causality or susceptibility? *A. Ebrahimi[1,2], M. Moghaddasi[3], M. Houshmand[4], S. Zeinali[5], S.H. Tonekaboni[6], M.S. Fallah[5], M. Mamarabadi[3], Moghaddasi, M.* 1) Hematology Research Center, Shiraz University of Medical science, Shiraz, Fars, Iran; 2) Parseh medical Genetics Center, Tehran, Iran; 3) Rasool Akram Hospital, Tehran University of Medical Sciences, Tehran, Iran; 4) National Institute of Genetic Engineering and Biotechnology (NIGEB), Tehran - Karaj Highway, Tehran, Iran; 5) Kawsar Human Genetics Research Center, KAWSAR Genomics & Biotech Center, Tehran, Iran; 6) Shahid Beheshti University of Medical Sciences, Tehran, Iran.

Epilepsy is a common chronic neurological disorder that is characterized by recurrent unprovoked seizures. Molecular studies of candidate genes can help us to define a correct differential diagnosis. So we studied SCN1B gene in Iranian patients with Idiopathic Epilepsy includes Febrile Seizure, Generalized Epilepsy with Febrile Seizures (GEFS+) or Dravet syndrome, diagnosed clinically, to explain genotype-phenotype correlation. Materials and Method: We screened 34 selected epileptic unrelated Iranian probands for all coding regions of SCN1B by PCR amplification and direct Sequencing. All families and probands were previously screened for SCN1A and mtDNA mutations. Results: PCR amplification of whole coding regions and splicing sites of SCN1B followed by direct sequencing revealed two novel sequence variations in patients (p.R248S, p.L210P) which did not detected in the healthy normal family members. Conclusions: According to final results it seems that these two novels SCN1B variations are not causative mutations in epileptic patients' but they can act as genetic predisposition factors in epileptic phenotypes which introduce a susceptibility especially in response to antiepileptic drugs.

3198F

A novel KCNQ2 mutation in a mexican-american family with benign familial neonatal convulsions. *A. Sznewajs[1], E. Rider[2], E. Sherr[1, 2].* 1) Pediatrics, University of California, San Francisco, San Francisco, CA; 2) Child Neurology, University of California, San Francisco, San Francisco, CA.

Background: Benign familial neonatal convulsions (BFNC) is a rare autosomal dominant epileptic disorder that results from mutations in two voltage-gated potassium channels, KCNQ2 (20q13.3) and KCNQ3 (8q24). The majority (92%) of reported mutations are found in KCNQ2. The purpose of this study was to perform mutation screening of KCNQ2 and KCNQ3 in order to identify a disease-causing mutation in a three-generation Mexican-American family with BFNC. Methods: A BFNC family was identified at the University of California, San Francisco Children's Hospital after admission of the proband, a 3-day old female infant with a 1 day history of unprovoked, nonfebrile neonatal seizures. The family history was significant for a mother, two maternal uncles and a maternal grandmother with a history of BFNC. After informed consent, mutation screening of KCNQ2 and KCNQ3 was performed in affected family members and the proband's father in order to identify the disease-causing mutations. Results: A novel heterozygous 65 base pair deletion was found in the 3 prime region of exon 3 in KCNQ2, encompassing the splice site region of all 5 affected family members. Restriction enzyme analysis was performed using the unaffected family member as a control, thereby confirming the deletion. No other mutations were found in KCNQ2 or KCNQ3. Conclusions: Here we report a three-generation Mexican-American family with a large novel KCNQ2 deletion. These findings are consistent with prior studies which show that large deletions in KCNQ2 (49-479 kbp) comprise a substantial amount (16%) of KCNQ2 mutations. At 1 year of age, the proband's seizures have resolved on phenobarbital. No seizures outside the neonatal period have been reported in the other 4 affected family members. This KCNQ2 mutation has implications for diagnosis and prognosis of BFNC. Its presence suggests a benign course with good prognosis and its identification can spare patients and physicians the need for extensive investigations or therapy.

3199W

Coronal craniosynostosis and radial ray hypoplasia: a third report of Twist mutation in a 33 weeks fetus with diaphragmatic hernia. *P. Juliette[1], C. Collet[2], F. Arbez-Gindre[3], L. Van Maldergem[1].* 1) Centre de Génétique Humaine, Université de Franche Comté, Besançon, France; 2) Service de Biochimie et Biologie Moléculaire, CHU Lariboisière, Paris, France; 3) Service d'Anatomie Pathologie, Université de Franche-Comté, Besançon, France.

We describe a multiple malformation syndrome comprising coronal craniosynostosis, unilateral radial ray hypoplasia and diaphragmatic hernia in a 33w female fetus born to a 46 y-old male with an alleged personal and family history of Crouzon syndrome. By identifying an already described c.445C>T TWIST missense mutation, we were able to reassign the diagnosis of the family condition to Saethre-Chötzen syndrome. The present report illustrates clinical variability of a dominantly inherited TWIST mutation, adds diaphragmatic hernia in the spectrum of TWIST-related malformations and provides a third example of Baller-Gerold/Saethre-Chötzen overlapping phenotype.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

† Trainee Award Finalist

3200F

Identification of a novel gene causing fetal akinesia deformation sequence. *M.M. Weiss[1], G. Tan[1], M. Smit[1], J.I. de Vries[2], J.P. vd Voorn[3], I. Kluijt[4], E. Sistermans[1], H. Meijers-Heijboer[1,4], S. Groffen[1,5], Q. Waisfisz[1].* 1) Clinical Genetics, VU University Medical Center, Amsterdam, Netherlands; 2) Obstetrics and Gynaecology, Research Institute MOVE, VU University Medical Center, Amsterdam, Netherlands; 3) Pathology, VU University Medical Center, Amsterdam, Netherlands; 4) Clinical Genetics, Academic Medical Center, Amsterdam, Netherlands; 5) Functional Genomics, Center for Neurogenomics and Cognition Research, VU University, Amsterdam, the Netherlands.

Fetal akinesia deformation sequence (FADS) refers to a group of disorders characterized by congenital malformations related to decreased or absent fetal movements, such as multiple joint contractures and lung hypoplasia, which is often lethal. The etiology is heterogeneous, but underlying neuromuscular disorders represent an important cause of fetal akinesia. Embryonic skeletal muscle development involves a proper organization of acetylcholine receptors at neuromuscular junctions, a process that is triggered by agrin and involves several downstream genes, including DOK7 and RAPSN both of which are previously implicated in FADS. Here we describe a consanguineous family with two affected children with FADS. SNP array analyses of the affected children revealed overlapping regions of homozygosity. Sequencing analysis of candidate genes detected a homozygous missense variant in another known component of the DOK7 / RAPSN signaling pathway. Because the parents of the affected children originated from a genetic isolate, we suspected that the mutation could be a founder mutation. We tested four other families with an affected child, who originated from the same genetic isolate, and confirmed the homozygous mutation in each of the affected children. The five families had common ancestors traced back 6 to 7 generations. To our knowledge, this is the first report of a mutation in this gene causing a fetal akinesia lethal phenotype.

3201W

Identification of a new syndrome with severe type of cutaneous photosensitivity, mild mental retardation and short stature caused by KIAA1530 (UVSSA) gene. *R. Sharifi, A. Ahmadi, E. Ozkan, R. Maroofian.* Medical Genetics, St George's University of London, London, United Kingdom.

Here we are reporting for the first time a new syndrome characterized by severe type of cutaneous photosensitivity, mild mental retardation and short stature without developing skin carcinoma. Using linkage studies and exome sequencing of several extensive families with autosomal recessive inheritance, we revealed that mutations in KIAA1530 (UVSSA) gene causing this syndrome. The undergoing investigations are focusing on the possibility RNA profile in affected cases in comparison with carriers and non-affected individuals which may provide fundamental understanding of transcription-coupled nucleotide-excision repair (TC-NER) pathway and explain the different clinical features across these TC-NER-deficient disorders.

3202F

Infantile cerebral and cerebellar atrophy population screening using PCR-RFLP method. *V. Adir, E. Shahak, E. Golinker, N. Ekhilevitch, Z. U. Borochowitz.* Dept Molecular Genetics, Bnai Zion Medical Ctr, Haifa, Israel.

Introduction: Infantile Cerebral and Cerebellar Atrophy (ICCA) patients present microcephaly of postnatal onset, epilepsy, and psychomotor retardation. Head circumference percentiles decline with age, and brain MRI shows cereberal and cerebellar atrophy with severe myelination defect. Screening of patients with ICCA identified the same homozygous mutation L371P (c.1112T>C) in the MED17 gene, in all affected individuals who were of Caucasus Jewish origin, indicating a founder effect. The c.1112T>C mutation was found in the heterozygous state in 4 of 76 unaffected individuals in this population, Kaufmann et al. (2010). Research Goal: Developing simple method for population screening for the c.1112T>C mutation in the MED17 gene, and screening the Caucasus Jewish origin population to validate the high frequency carriers in this population. Method: We developed a PCR mediated site directed mutagenesis (PSDM) followed by restriction enzyme digestion test that would enable easy analysis of the c.1112T>C mutation in the MED17 gene. We designed PCR primers in which the 5'-to-3' (forward) primer bordered the c.1112T>C mutation site and included one nucleotide change so that in the normal allele an Xba1 restriction site will be created (TCTAGA). MED17 F CCGGAGGACCACCTTTATGTtC MED17 R GCTATGTGAATAATCAAGACTCAGCA. Using this primer set for amplification from genomic DNA gave a 157bp fragment and generated Xba1differential cleavage products of the mutant allele (157bp) versus the wild-type allele (137bp +20bp). Results: Our screening results indicated lightly higher carriers frequency than what was found by Kaufmann et al, we suggest that the reason for that is the small sample size in both our and Kaufmann's screening. If we combine the results from Kaufmann's and ours screening the carries frequency is 6.25% (1:16). The severity of the ICCA disease and the combined results (which definitively show that there is a high frequency of carriers among Caucasus Jews), resulted in the recommendation by the Medical Geneticists Association of Israel that this genetic test be included in the recommended screening genetic tests for Caucasus Jewish couples who plan pregnancy .

3203W

Genotype-phenotype correlation of enlarged vestibular aqueduct syndrome. *R. Birkenhager[1], S. Arndt[1], W. Maier[1], A. Aschendorff[1], E. Löhle[1,2], R. Laszig[1].* 1) Department of Otorhinolaryngology, Head and Neck Surgery, University Medical Center Freiburg, Killianstrasse 5, D-79106 Freiburg, Germany; 2) Department of Otorhinolaryngology, Head and Neck Surgery, University Medical Center Freiburg, Section Phoniatry and Pedaudiology.

The enlarged vestibular aqueduct syndrome (EVA) is one of the most common inner ear malformations, a distinct clinical entity characterized by progressive sensorineural hearing impairment. It is often associated with Pendred- Syndrome (PDS), an autosomal-recessive disorder characterized additionally with goiter. PDS is associated with temporal bone abnormalities ranging from isolated enlargement of the vestibular aqueduct (EVA) to Mondini dysplasia, a more complex malformation that also includes cochlear hypoplasia. Hearing loss is prelingual in the majority of the cases; only subset of patients has a progressive hearing loss later in life. EVA and PDS are caused by mutations in the SLC26A4 gene. The SLC26A4 gene is expressed in the non-sensory epithelia of the inner ear. Additionally the genes FOXI1 and KCNJ10 are described which may also be responsible for Pendred-Syndrome and inner ear malformations. The FOXI1 gene is involved in the transcriptional control of the SLC26A4 gene. The KCNJ10 gene encodes for a K+-channel located in the stria vascularis. In this study we analysed 152 patients with EVA and hearing loss to accomplish a phenotype genotype correlations for enlarged vestibular aqueduct related deafness. Individual exon and intron transitions of the SLC26A4, FOXI1 and KCNJ10 genes of patients were sequenced. Audiometric thresholds were performed, and radiologically, the vestibular aqueduct midpoint and opercular width were measured. In the analysed patients with EVA, a total of 32 different SCL26A4 mutations were detected, mutations could not be detected in 68 % of the cases. FOXI1 and KCNJ10 analyses reveal only known polymorphisms but no mutation. The present results obtained in patients with EVA indicate that this phenotype is likely to be complex. Moreover, analysis of FOXI1 and KCNJ10 genes in our cohort of enlarged vestibular aqueduct hearing loss suggests a minor implication of these two genes in the phenotype of non-syndromic EVA and PDS.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3204F

Molecular diagnosis of congenital muscular dystrophies with defective glycosylation of alpha dystroglycan using next generation sequencing technology. *J. CHAE, B. Lim, J. Ko, J. Choi, M. Wo, W. Park, B. Min.* Seoul National University Hosp, Seoul, South Korea.

Background Targeted resequencing using next generation sequencing technology is being rapidly applied to the molecular diagnosis of human genetic diseases. Muscular dystrophies as a group may be the appropriate candidate, because this group showed genotype-phenotype heterogeneity. As a proof of concept study, we selected the four congenital muscular dystrophies with defective glycosylation of alpha-dystroglcyan patients. **Methods** A custom solution-based target enrichment kit was designed to capture whole genomic regions of the 26 muscular dystrophies related genes including 6 genes implicated with alpha dystroglycanopathies. A multiplexing strategy, wherein four differently bar-coded samples were captured and sequenced together in a single lane of the Illumina Genome Analyser, was applied. **Results** Although the approximate 95% of both coding and non-coding regions were covered with more than 15 read depths, parts of the coding regions of FKRP and POMT2 were insufficiently covered. The homozygote and compound heterozygote POMGnT1 mutations were found in two patients. Three novel non-coding variants of FKTN were identified in one patient who had retrotransposon insertion mutation of FKTN in only one allele. Family study revealed two of the three non-coding variants segregated in the other allele with unknown mutation. The possible causes of sub-optimal coverage in a small portion of coding regions were estimated to be high GC ratio and insufficient capture probe design **Conclusion** The current targeted resequencing strategy showed the promising results for the extension of this method to other muscular dystrophies. Since sub-optimal coverage in a small subset of coding regions may affect the diagnostic sensitivity, complementary Sanger sequencing may be required.

3205W

Various clinical manifestations in 40 cases with type 1 neurofibromatosis. *CK. CHEON.* Department of Pediatrics, Pusan National University, Children's Hospital, South Korea.

Background: Neurofibromatosis type 1 (NF1) is one of the most common autosomal dominant disorder caused by mutations in the NF1 gene. The classic features of the clinical phenotype include the presence of café-au-lait spots, neurofibromas, axillary and inguinal freckling, Lisch-nodules and deformities of the skeletal system. However, it is known from the literature that the phenotypic variability can pose a huge diagnostic difficulty. Objective: We present a clinical and molecular study of Korean patients with NF1. Molecular genetic testing of the NF1 gene has been performed since 2009; during this period several unusual phenotypic variants were found. Methods: A total of 40 patients, 14 females and 26 males, were enrolled in this study at a single center. All the NF1 patients met the diagnostic criteria proposed by NIH Consensus Conference in 1987. Clinical manifestations, clinical courses and results of genetic study were reviewed. Results: Approximately 72% of all NF1 patients lack a family history of the disease. Subcutaneous and cutaneous neurofibroma was seen in 15 patients (37.5%). Lisch nodule was found in 7 patients (17.5%). Kyposcoliosis was detected in 12 patients (30%). Developmental delay or mental retardation was shown in 4 patients (10%). Brain involvement of neurofibromatosis was detected in 13 patients (32.5%). Among them, 5 patients (12.5%) showed harmatoma on cerebellum and basal ganglia. Two patients revealed hydronephrosis.. One patient showed malignant neural sheath tumor at the age of 41. One patient presented with detected severe hypertension at the age of 54. She was found to have elevated serum and urinary catecholamines levels. This patient underwent right adrenalectomy. Histology was diagnostic for pheochromocytoma. A 27-year-old adult having NF-1 was diagnosed with polycythemia vera. The mutational analysis of the NF1 gene revealed three novel mutations (Leu1944PhefsX6, Leu1773Pro,Cys2371X). Conclusions: Clinical manifestations of NF1 are variable, but “café au lait” spots and neurofibromas are the frequently symptoms. On some cases studied by us, the NF1 is associated with pheochromocytoma, polycythemia vera, kyphoscoliosis, and developmental delay or mental retardation. Even if there is currently no effective drug treatment for NF1, genetic counseling is very important and some of the symptoms of NF1 can be treated individually.

3206F

Mutation Analysis Results of 5α Reductase Type 2 Enzyme Deficiency Patients. *H. Onay[1], D. Goksen[2], A. Aykut[1], F. Hazan[3], S. Darcan[2], F. Ozkinay[1,2].* 1) Medical Genetics Dept, Ege University School of Medicine, Izmir, Türkiye; 2) Pediatrics Dept, Ege University School of Medicine, Izmir, Türkiye; 3) Dr. Behcet Uz Children's Hospital, Izmir, Türkiye.

The 5α reductase type 2 enzyme deficiency is an autosomal recessive disorders and is one of the major cause of 46,XY disorders of sex differentiation. This enzyme is encoded by the SRD5A2 gene on chromosome 2 and catalysis the conversion of testosterone (T) to dihydrotestosterone (DHT) and the gene is expressed mostly in external genital tissues and prostate. DHT is the essential androgen for normal development of external genitalia in the male fetus. Males affected with this deficiency usually present perineoscrotal hypospadias, microphallus and cryptorchid testes. In this study the mutation analyses results of the 151 patients with 5α reductase type 2 enzyme deficiency who referred to Medical Genetic Department of Ege University Hospital between 2008–2012 were evaluated. SRD5A2 gene mutation analysis was performed by sequencing of the coding exons of the gene. No mutation was found in 136 patients (90,1%). Ten of 15 patients (6,6%) carried homozygous mutations and remaining (3,3%) carried mutations in heterozygous state. The most common mutation (in 4 patients) detected in homozygous patients was L55Q as expected and second one was G196S. As a result mutation detection rate was very low in our patient group. Careful examination of the clinical signs and the laboratory findings before SRD5A2 gene mutation analysis would higher the detection rate.

3207W

Griscelli Syndrome with RAB 27A mutation and prenatal diagnosis. *I. Panigrahi, R. Suther, B. Behera, A. Rawat, R. Marwaha.* Pediatrics, PGIMER, Chandigarh, Chandigarh (UT), India.

One year old female, the first child of parents, presented with febrile episode 1 month earlier, followed by seizures and developmental regression for 15 days. Examination revealed dolichocephalic head with silvery white hair, round face, silvery eyelashes, brown-black irides and tented upper lip. She had hepatosplenomegaly, liver was 5cm below right costal margin and spleen was 6cm below left costal margin. CNS examination revealed normal tone with elicitable deep tendon reflexes. Since she had acute seizures s he was treated as a case of meningitis. MRI brain done revealed multiple extensive hyperintense lesions in periventricular/subcortical /cerebellar white matter along with hydrocephalus. Leptomeningeal enhancement was seen along cerebellar folia and convexities. Her WBC counts were low (lowest of 1800/mm3 with 57% polymorphs). She was put on phenytoin, midazolam, leviracetam for seizures and ultimately on thiopentone infusion at 4mg/kg/hour with increase upto 5mg/kg/hour. Ketamine infusion was also given. IV methylprednisolone was also given. RAB 27A gene mutation analysis revealed splice mutation 239+1 G-T in intron 3. Thus, the child was confirmed to have Griscelli syndrome type2 (GS2), in which neurological and immunological abnormalities co-exist, leading to early mortality. She developed intractable seizures during hospital stay and expired around D14 of hospitalization. Mother is pregnant and prenatal diagnosis is planned by amniocentesis and DNA analysis.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3208F

Molecular diagnosis of rare Mendelian diseases using whole exome sequencing. *B. Rodríguez-Santiago[1,2], S. Boronat[3,4], J. Argente[5,6,7], I. Valenzuela[4], M. del Campo[2,4,8], L.A. Perez-Jurado[2,8], L. Armengol[1,2].* 1) qGenomics, Barcelona, Spain; 2) Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Instituto de Salud Carlos III, U-735, Barcelona, Spain; 3) Servicio de Neuropediatria, Hospital Universitari Vall d'Hebron, Barcelona, Spain; 4) Programa de Medicina Molecular i Genètica, Hospital Universitari Vall d'Hebron, Barcelona, Spain; 5) Departamento de Pediatría, Universidad Autónoma de Madrid, Madrid, Spain; 6) Servicios de Pediatría y Endocrinología, Instituto de Investigación La Princesa, Hospital Infantil Universitario Niño Jesús, Madrid, Spain; 7) Centro de Investigación Biomédica en Red de la Fisiopatología de la Obesidad y Nutrición (CIBEROBN), Instituto de Salud Carlos III, Madrid, Spain; 8) Unitat de Genètica, DCEX, Universitat Pompeu Fabra, Barcelona, Spain.

The development in the recent years of the so-called next generation sequencing technologies based on massive parallel methods currently allows the production of millions of DNA sequences at an unprecedented speed with an increasingly reduced cost per nucleotide. These technologies are producing very significant scientific achievements, with the identification of new genes and the resolution of the genetic basis of Mendelian diseases at the forefront. The potential of this technology is being used to create new applications and biological tests that are soon going to revolutionise the pre- and postnatal diagnosis of genetic disorders. In order to assess the utility of this technology in molecular diagnostics we have applied whole-exome sequencing to two groups of patients with normal aCGH results. Group 1 included patients diagnosed with rare Mendelian disorders with genetic heterogeneity and well defined phenotype whereas group 2 was formed by patients with severe phenotypes but no specific diagnosis. Among group 1, patients with primary autosomal recessive microcephaly (7 causative genes known), Bardet-Biedl syndrome (15 genes), Joubert syndrome (10 genes), microcephalic dwarfism (4 genes) and Walker-Warburg syndrome (6 genes) were analysed. Potential causative biallelic mutations were identified and validated in parental samples in 4 out of 6 cases (66%) from group 1. Candidate mutations putatively related to the phenotype were also detected in all group 2 patients, but follow-up studies are required to define their pathogenic role. Our data indicates that exome sequencing in a single patient is highly effective (and likely cost-effective) as the initial diagnostics tool to identify causative pathogenic mutations in rare Mendelian disorders while in phenotypes without specific diagnosis the technology is a research tool that provides new candidate pathogenic variants.

3209W

Next-generation sequencing for low and high-bone density disorders. *G. Sule[1], P. Campeau[1], S. Nagamani[1], B. Dawson[1], M. Grover[1], C. Bacino[1], J. Lu[2,3], E. Lemire[4], R. Gibbs[2], D. Cohn[5], V. Zhang[1], L. Wong[1], B. Lee[1,6].* 1) Molecular and Human Genetics, Baylor College of Medicine, Houston, TX; 2) Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX, USA; 3) Department of Structural and Computational Biology & Molecular Biophysics, Baylor College of Medicine, Houston, TX, USA; 4) Department of Pediatrics, University of Saskatchewan, Saskatoon, Canada; 5) Department of Molecular, Cell, and Developmental Biology, University of California, Los Angeles, CA; 6) Howard Hughes Medical Institute, Houston, TX.

Osteogenesis imperfecta (OI), Ehlers-Danlos syndrome (EDS) and osteopetrosis (OPT) are collectively relatively common inherited skeletal diseases. Evaluation of subjects with the aforementioned conditions often includes molecular testing which has important counseling, therapeutic and even legal implications. Since multiple genes have been implicated in these conditions, Sanger sequencing of individual genes results in significant increases in cost and time to reach a molecular diagnosis. For example, a comprehensive evaluation for genetic causes for OI would entail sequencing of at least 9 genes totaling more than 150 exons. In order to circumvent these problems, we have designed a next-generation sequencing (NGS) platform that would allow sequencing of multiple genes implicated in OI, OPT, EDS, and other conditions leading to low or high bone mass on a single diagnostic platform. Our assay uses a liquid-phase library that captures 602 exons (~100 kb) of 33 genes. NGS of the captured exons by Illumina HiSeq2000 resulted in an average coverage over 900X. The platform was successfully validated by identifying mutations in 6 patients with known mutations. Moreover, in 4 patients with OI or OPT without a prior molecular diagnosis, the assay was able to detect the causative mutations. In conclusion, our NGS panel provides a fast and accurate method to arrive at a molecular diagnosis in patients with high or low bone mass diseases.

3210F

Molecular diagnosis of autosomal dominant polycystic kidney disease using massively parallel sequencing. *Y. Tan[1], A. Michaeel[1], G. liu[1], J. Blumenfield[2, 3], S. Donahue[3], T. Parker[3], D. levine[3], H. Rennert[1].* 1) Dept Pathology/Laboratory Med, Weill Cornell Medical College, New York, NY; 2) Dept of Medicine, Weill Cornell Medical College, New York, NY; 3) The Rogosin Institute, New York, NY.

Genetic testing of PKD1 and PKD2 is useful for diagnosis and prognosis of autosomal dominant polycystic kidney disease (ADPKD), particularly in those without a family history. However, PKD gene mutation analysis is complicated by the duplication of PKD1 exons 1–33 as six pseudogenes on chromosome 16, a high level of allelic heterogeneity (genetic variations), and the high cost of Sanger sequencing. We have developed and validated a clinical assay to analyze both PKD genes, using next-generation sequencing (NGS), by multiplexing barcoded libraries from long-range PCR (LR-PCR). We used this approach to characterize a cohort of patients with PKD gene mutations previously defined by Sanger sequencing. PKD1 and PKD2 genes were amplified in a total of ten LR-PCR reactions (1.4–10.9 kb), using locus-specific primers for PKD1 duplicated regions. LR-PCR amplicons from the same patient were combined and indexed with a unique DNA barcode. Indexed libraries from 25 patients were pooled together and analyzed using either a flow-cell of MiSeq or a single flow-cell lane of HiSeq system (Illumina Inc., San Diego, CA). Sequencing results were then sorted according to the barcodes with FASTX toolkit and aligned against the reference sequence with BWA program. Sequence variants were called and filtered with GATK software (The Genome Analysis Toolkit, The Broad Institute) following Best Practice Guidelines recommended by GATK. The raw variant calls were then annotated with ANNOVAR software. The data analysis pipelines were automated with Unix shell scripts. A total of 241 genetic variants were identified within exonic and adjacent intronic regions of PKD1 and PKD2 in 25 patients with the ADPKD phenotype. When compared with conventional Sanger sequencing, this NGS based approach achieved a sensitivity of 99.2%, and a specificity of 99.9%. Of these genetic variants, 16 variants were classified as definite pathogenic mutations, including two small deletions (a 4-bp and a 10-bp), and a 24-bp insertion. This NGS based testing detected all pathogenic mutations, achieving a sensitivity and specificity of 100%. By using the MiSeq platform, 25 patients were genotyped within one week for less than $100 per patient (including all reagents and sequencing costs). NGS-based ADPKD genetic analysis is a highly accurate and reliable approach for mutation analysis, achieving high sensitivity and improved intronic coverage with faster turnaround time and lower cost.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3211W

Genotype-phenotype correlation in Bardet-Biedl Syndrome. *E. Forsythe[1], K. Sparks[2], M.S.B. Huda[3], J. Hazlehurst[4], S. Mujahid[3], P. Carroll[3], B. McGowan[3], J.W. Tomlinson[4], S. Mohammed[5], P.L. Beales[1].* 1) Molecular Medicine Unit, UCL Institute of Child Health, London, United Kingdom; 2) Clinical Genetics, Great Ormond Street Hospital, London, United Kingdom; 3) Department of Endocrinology, Guy's and St Thomas' NHS foundation Trust, London, United Kingdom; 4) Centre for Endocrinology, Diabetes and Metabolism, Institute of Biomedical Research, University of Birmingham, Birmingham, United Kingdom; 5) Clinical Genetics, Guy's and St Thomas' NHS foundation trust, London, United Kingdom.

Background: Bardet-Biedl Syndrome (BBS) is a rare autosomal recessive ciliopathy. It is a clinically heterogeneous condition characterized by retinal dystrophy, obesity, post-axial polydactyly, renal dysfunction, learning difficulties and hypogonadism. The diagnosis can be confirmed in 80% of patients by sequencing the 16 known disease-causing genes. In northern European populations the majority of pathogenic mutations are found in *BBS1* and *BBS10* accounting for 23.2% and 20% respectively.

Purpose: To explore genotype-phenotype correlations in individuals with pathogenic mutations in *BBS1* and *BBS10* based on a questionnaire addressing key patient concerns about BBS and data from the national BBS clinic services.

Methods: 115 patients responded to the questionnaire. 35 patients had confirmed pathogenic mutations in *BBS1* and 15 in *BBS10*. The questionnaire was devised in collaboration with the UK patient support group and includes questions about visual deterioration, endocrine, renal and mental health problems. Fasted blood samples, height, weight and blood pressure were taken from 239 patients attending BBS clinics. Blood samples included full blood count, glucose, renal, liver, lipid and endocrine profiling. 96 patients seen in clinic had confirmed mutations in *BBS1* or *BBS10* (72 and 24 respectively). Results were assessed using standard statistical methods and the significance level was set at 0.05.

Summary of results: The age of onset of night blindness was significantly higher in patients with *BBS1* mutations (P=0.008, n=45). Individuals with *BBS1* mutations also reported being registered blind later (p=0.0008, n= 40). The *BBS10* cohort were more likely to report renal impairment (p= 0.042, n=40). Results from the BBS clinics revealed that patients with *BBS1* mutations have better albumin-creatinine ratios (p=0.038, n=45) and higher levels of HDL cholesterol (p=0.004, n=70) despite this group being significantly older (p=0.03, n=96). No other markers reached statistical significance even when corrected for age.

Conclusion: This study examines key issues prioritized by individuals affected by BBS and data from national BBS clinics. A milder ocular, renal and lipid phenotype is observed in patients with *BBS1* mutations. This should help guide genetic counseling and testing of patients with BBS.

3212F

Photoreceptor loss in LCR-deletion associated Blue Cone Monochromacy: implications for gene therapy. *R.B. Hufnagel[1,2], A. Dubra[3,4], R.A. Sisk[5], J.C. Gardner[6], S. Riazuddin[1,7], A.J. Hardcastle[6], A.T. Moore[6,8], J. Neitz[9], M. Michaelides[6,8], M. Neitz[9], J. Carroll[3,4,10], Z.M. Ahmed[1,2].* 1) Human Genetics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH; 2) Pediatric Ophthalmology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH; 3) Ophthalmology, Medical College of Wisconsin, Milwaukee, WI; 4) Biophysics, Medical College of Wisconsin, Milwaukee, WI; 5) Cincinnati Eye Institute, Cincinnati, OH; 6) Ophthalmology, University College London, London, United Kingdom; 7) Pediatric Otolaryngology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH; 8) Moorfields Eye Hospital, London, United Kingdom; 9) Ophthalmology, University of Washington, Seattle, WA; 10) Cell Biology, Neurobiology, & Anatomy, Medical College of Wisconsin, Milwaukee, WI.

Blue cone monochromacy (BCM) is an X-linked retinal disorder of cone photoreceptors caused by mutation that either prevents expression of both *OPN1LW* (encoding red opsin) and *OPN1MW* (encoding green opsin) or interferes with function of the encoded opsins. Patients with BCM have reduced visual acuity, markedly reduced color discrimination, photophobia, and nystagmus. While BCM does not usually demonstrate progressive visual deterioration, little is known about the effects of opsin loss on photoreceptor structure and function, both critical for predicting efficacy of gene replacement therapies. Here, we describe genetic, clinical, and in vivo cellular analyses of four families with BCM.

Clinical diagnosis was based upon clinical assessment, color vision testing, family history, and full-field and S-cone electroretinography (ERG). DNA was purified from blood samples obtained from affected individuals, parents, and unaffected siblings. Genomic segments harboring the *OPN1LW* and *OPN1MW* exons and promoter sequences were analyzed by PCR amplification followed by direct Sanger sequencing. Gene arrangement and copy number were evaluated by RT-PCR. Macular and photoreceptor structure was evaluated by Ocular Coherence Tomography (OCT) and Adaptive Optics (AO) imaging.

Probands ranged in age from 8 to 35 years old. All four families have mutations deleting the enhancer Locus Control Region (LCR), obviating expression of any of the *OPN1LW* or *OPN1MW* genes in the array. Affected individuals had moderate impairment in best corrected visual acuity, impaired cone ERGs, and reduced central macular thickness with changes in the retinal pigmented epithelium (RPE). AO imaging of foveal photoreceptors indicated diminished outer segment structure.

In conclusion, these families with LCR deletions had similar phenotypic manifestations, including foveal thinning with a granular RPE. Disease severity did not correlate with age or visual acuity. Macular photoreceptors exhibited diminished outer segment structure, critical for visual transduction. These previously uncharacterized architectural features of BCM suggest structural defects that may have implications for cellular viability and efficacy of future gene therapy approaches.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3213W

VCP-associated Inclusion Body Myopathy with Paget Disease of Bone and/or Frontotemporal Dementia: Expanding Natural History and Genotype-Phenotype characterization. *M. Khare[1], S.G. Mehta[2], R. Ramani[1], G.J. Watts[3], B. Martin[4], M. Simon[5], K.E. Osann[6], S. Donkervoort[1], E. Dec[1], A. Nalbandian[1], A. Wang[7], T. Mozaffar[7], C.D. Smith[4], V.E. Kimonis[1].* 1) Pediatrics, University of California, Irvine, Irvine, CA; 2) East Anglian Regional Genetics Service, Addenbrookes Hospital, Cambridge, UK; 3) Biomedical Research Center, University of East Anglia, Norfolk;; 4) Department of Neurology and Sanders-Brown Center on Aging, University of Kentucky, Lexington, KY; 5) Mitomed Laboratory, MAMMAG, University of California, Irvine, CA; 6) Division of Hematology/Oncology, University of California, Irvine, CA; 7) ALS and Neuromuscular Center, University of California, Irvine, CA.

VCP disease associated with Inclusion body myopathy, Paget disease of the bone and frontotemporal dementia is a progressive autosomal dominant disorder caused by mutations in the VCP gene, a member of the AAA-ATPase superfamily. To establish the natural history and genotype-phenotype correlations we analyzed the clinical and biochemical markers from a database of 190 members in 27 families harboring ten missense mutations. Individuals were grouped into three categories: symptomatic, presymptomatic carriers and non-carriers. The symptomatic families were further divided into ten groups based on their specific VCP mutations. We identified a higher prevalence of FTD in females with a ratio of 2:1. Survival analysis by Kaplan Meier survival method revealed an average life span after diagnosis of myopathy and Paget of 18 and 19 years respectively, and after dementia only 6 years. There were marked intra and interfamilial variations; and significant genotype phenotype correlations were only apparent between two of the most common mutation groups. The R155C mutation was more severe than R155H mutation with a significant earlier onset of myopathy and PDB along with a significant decrease in mean survival (p=0.03). We identified amyotrophic lateral sclerosis (ALS) in thirteen affected individuals in our study database (8.9%). Electromyography studies revealed neurogenic features such as fasciculations in 24% of individuals. Five individuals were also identified with Parkinson's disease (3%). We analyzed the tumor incidence because previous studies have implicated VCP's role in predicting the metastatic potential and, thus, the prognosis of patients with wide variety of tumors. We identified several patients with rare tumors including malignant thymoma, frontotemporal pleomorphic xanthoastrocytoma (a primitive neuroectodermal tumor), and peripheral nerve sheath spindle cell tumor of the thigh along with typical breast and colon cancers. This study represents the largest dataset of patients with VCP disease and has a potential to yield better understanding of the natural history, and possible genotype-phenotype correlations with atypical features such as ALS, Parkinson's disease, cancers and other features.

3214F

A novel short stature syndrome caused by defects in glycosaminoglycan synthesis due to altered xylosyltransferase 1 activity. *J. Schreml[1], B. Durmaz[2], O. Ço{{gab}}ulu[2], K. Keupp[1], F. Beleggia[1], E. Pohl[1], E. Milz[1], G. Nürnberg[3], P. Nürnberg[3], J. Kuhn[4], M. Coker[5], S. Kalkan Ucar[5], F. Özkinay[2], B. Wollnik[1].* 1) Institute of Human Genetics, University of Cologne, Cologne, Germany; 2) Department of Medical Genetics, Ege University, Bornova, Izmir, Turkey; 3) Cologne Center for Genomics, University of Cologne, Cologne, Germany; 4) Institute for Laboratory and Transfusion Medicine, Heart and Diabetes Center North Rhine-Westphalia, Ruhr University Bochum, Bad Oeynhausen, Germany; 5) Ege University, Department of Pediatric Metabolism, Bornova, Izmir, Turkey.

Glycosaminoglycans (GAGs) are essential components of various physiological processes that are indispensable for embryonic development. GAG synthesis is initiated by the sequential addition of a “linker chain“, comprising of four sugar residues, to a specific region of a core protein. During this complex process, xylosyltransferases 1 and 2 (XT1 and XT2), which are expressed in a tissue and time dependent manner and are mainly located in the cis-golgi network, initiate the important linker chain formation. Here we present a novel syndrome characterized by severe short stature, facial dysmorphism, alterations of fat distribution, and moderate intellectual disability. X-rays in affected individuals of a consanguineous Turkish family showed mild skeletal dysplasia in the form of metaphyseal widening, short femoral neck and thickened ribs. Using an innovative whole-exome sequencing strategy combined with exome variant analysis for homozygous stretches and independent homozygosity mapping, we identified the homozygous p.R481W missense mutation in *XYLT1* encoding xylosyltransferase 1 in both affected individuals of the family. The parents were heterozygous carriers and the healthy brother did not carry the mutation. *XYLT1* is located within the largest linked region on chromosome 16 and the missense mutation affects a highly conserved residue in close proximity to a functional domain of the enzyme. Therefore, the mutation is predicted to lead to a functional impairment of enzyme activity. Indeed, immunostaining of primary patient fibroblasts pointed towards a mislocalization of XT1 with a diffuse staining pattern compared to the distinct staining in the golgi region of wild-type control fibroblasts obtained from a healthy sibling. Moreover, Western blot analysis of the core protein decorin in cell culture supernatant showed evidence of glycosylation differences between patient and control cells. These data provide the first evidence that functional alterations of xylosyltransferase 1 cause a syndrome associated with short stature and intellectual disability in the investigated family.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3215W

Case studies for the clinical diagnosis of rare, congenital, pediatric disorders by whole genome sequencing. *S. Szelinger[1,2], V. Narayanan[2,3], H. Boman[4], A.F. Hahn[5], J.J. Corneveaux[1,2], A.L. Siniard[1,2], A.A. Kurdoglu[1,2], M.J. Huentelman[1,2], D.W. Craig[1,2].* 1) Neurogenomics Division, Translational Genomics Research Institute, Phoenix AZ., USA; 2) Childhood Rare Disorder Center, Translational Genomics Research Institute, Phoenix Arizona, USA; 3) Pediatric Neurogenetics Center, Barrow Neurological Institute, Phoenix Arizona, USA; 4) Center for Medical Genetics and Molecular Medicine Haukland University Hospital Bergen, Norway; 5) Department of Clinical Neurological Sciences London Health Sciences Centre University of Western Ontario London, Ontario, Canada.

Improved next-generation sequencing technology and reduced costs make whole genome analysis an attractive clinical diagnostic tool in hard-to-diagnose genetic disorders. We report two independent cases of neurogenetic disorder, where whole genome sequencing and candidate variant analysis was integrated with clinical information to identify one case as Crisponi Syndrome/CISS1 (Cold-Induced Sweating Syndrome1) and Mitochondrial encephalomyopathy (MtE) in the other case. The Crisponi Syndrome case was paired-end sequenced by Illumina Genome Sequencing Service on a Genome Analyzer IIX (GAIIX) to a mean genomic coverage of 39X. The MtE case was paired-end sequenced by Illumina's FastTrack Sequencing Service on a HiSeq2500 to a mean genomic coverage of 33X. Quality assessment, sequence alignment, short nucleotide polymorphism (SNP), insertion, deletion discovery were performed by Illumina. Variant annotation and candidate coding variant prioritization followed a tiered approach, integrating data from the 1000 Genomes Project, Ensembl, NCBI RefSeq, OMIM, Consensus CDS, ClinVar, and the Human Gene Mutation Database. Evaluation of candidate gene list in both cases resulted in the identification of novel heterozygous coding SNPs that were deleterious. However neither candidate variant lists contained a second hit in the selected candidate genes that would be expected under autosomal recessive inheritance. In the Crisponi case, the second hit, a frameshift deletion in exon1, was discovered after Sanger sequencing CRLF1 candidate gene, which is known to contain a hard-to-sequence first exon with high GC content. In the MtE case, the second hit, a frameshift deletion, in POLG gene was discovered after application of alternative bioinformatic approach. In both cases, availability of detailed clinical information for the interpretation and prioritization of genomic data was critical to discover the causative genes and make the diagnosis. However, the inability of the GAIIX to obtain reads in high GC content regions, and the false negative call of the POLG deletion highlight some of the challenges faced by next-generation sequencing in its routine application for clinical diagnostics. Although these problems are significant, the fact that quality and speed of next-generation sequencing is constantly improving with sequencing and analysis turnaround time of about two weeks, the main question is not if whole genome analysis will be routinely used in the clinic, but when.

3216F

Variable phenotypic spectrum in a cohort of patients with ARID1B mutations. *S. Boulanger, L. Mariage, D. Lederer, S. Moortgat, A. Destree, P. Hilbert, I. Maystadt.* Human Gen Ctr, Inst Pathology & Gen, Charleroi, Belgium.

In 1970, Coffin and Siris reported 3 girls with severe mental retardation, growth retardation, lax joints and absent or hypoplastic fifth distal phalanges and nails. Later, the Coffin-Siris syndrome (CSS) has been characterized by developmental delay, severe speech impairment, coarse facial features, hypertrichosis, hypoplastic or absent fifth fingernails or toenails and agenesis of the corpus callosum. Very recently, dominant mutations in ARID1B gene have been found by whole-exome sequencing in CSS patients (Santen et al, 2012). In that study, it was proposed that CSS could be consider in all individuals with intellectual disability, speech impairment and agenesis of the corpus callosum. In parallel, loss-of-function mutations in the ARID1B gene were found in a cohort of patients with non specific cognitive impairment, suggesting that haploinsufficiency of ARID1B is a common cause of intellectual disability (Hoyer et al, 2012). We screened a cohort of patients with developmental delay, moderate to severe intellectual disability, severe speech impairment, and some clinical features reported in Coffin-Siris syndrome (short stature, coarse face, flat nasal bridge, bulbous nasal tip, large mouth, sparse scalp hair contrasting with body hirsutism, nail dysplasia, hypotonia and joint laxity). The patients carrying an ARID1B mutation present with an atypical phenotype, not specific enough to allow for an unambiguous diagnosis of Coffin-Siris syndrome.

3217W

Whole Exome Sequencing in patients with Intellectual Disabilities. *I. Madrigal[1,2], U. Liljedahl[3], MI. Alvarez[1,2], O. Karlberg[3], L. Rodriguez-Revenga[1,2], A. Mur[4,5], AC. Syvänen[3], M. Mila[1,2].* 1) Biochemistry and Molecular Genetics, Hospital Clinic, Barcelona, Barcelona, Spain; 2) Centre for Biomedical Research on Rare Diseases (CIBERER), ISCIII, Barcelona, Spain; 3) Department of Medical Sciences, Molecular Medicine, Uppsala University, Uppsala, Sweden; 4) Servicio de Pediatría, Hospital del Mar, Barcelona, Spain; 5) Departamento de Pediatría y Obstectricia de la UAB, Barcelona, Spain.

Intellectual disabilities refers to a generalized disorder, characterized by substantial limitations in intellectual, functioning and adaptive behavior, diagnosed before 18 years of age and affecting about 1–3% of the general population. Genetically, the high heterogeneity and the unexpected great complexity of the genetic basis should be highlighted. From 50–60% of patients with intellectual disability remain undiagnosed. The use of Whole Exome Sequencing in families with some affected members increases the rate of diagnosis and facilitates the identification of new genes. The aims of this study were: to identify mutations in known intellectual disability genes, to identify new genes responsible for intellectual disability, to establish the phenotype-genotype correlation and to provide genetic counseling. We sequenced 32 individuals from eight families of self-reported European ancestry with unexplained moderate to severe intellectual disability using the Illumina HiSeq 2000 Sequencing System. We identified two new mutations in known genes responsible for intellectual disability: the c.2013 mutation delGfs*3 in the UBE3A gene and the c.1405C>T p. R469C mutation in the SMC1A gene; and one previously described mutation in the OCRL gene (c.1567 G>A p.D523N). In the remaining 5 families, several candidate mutations have been identified and more studies are being performed. In conclusion, this study confirms the high heterogeneity and difficulty in the clinical diagnosis of intellectual disability and demonstrates that Whole Exome Sequencing is a very efficient, reliable and cost-effective method which should be incorporated to routine diagnosis in the near future. Acknowledgements: FP7/2007–2011, grant agreement n°262055 (ESGI project).

3218F

Identification of the p.A140V mutation in MECP2 gene in a family with non specific X-linked mental retardation. *I. Maystadt[1,2], S. Boulanger[1], P. Vrielynck[3], S. Moortgat[1], A. Destree[1], D. Lederer[1].* 1) Ctr Human Genetics, Inst Pathology & Genetics, Gosselies, Belgium; 2) FUNDP, Namur, Belgium; 3) Service de Neurologie, Hôpital de Jolimont, Haine-Saint-Paul, Belgium.

Mutations in MECP2 gene, a gene located on the X chromosome and encoding a methyl CpG binding protein, have been identified in patients with Rett syndrome. This neurodevelopmental disorder, classically associated with normal early development followed by loss of acquired skills and deceleration of head growth, affects almost exclusively females because of the lethal effect of the mutations in hemizygous affected males. Here we report on the identification of the p.A140V mutation in the MECP2 gene in 4 males and 3 females of a family affected with X-linked mental retardation, and we provide a phenotypical description of the patients. Females present with mild cognitive impairment and speech difficulties. Males have severe intellectual disability, impaired language development, sociable affect, slowly progressive spastic paraparesis/pyramidal signs and dystonic movements of the hands. Two of them have microcephaly. None of them developed epilepsy. The p.A140V has been reported to preserve the methyl-CpG binding function of the MeCP2 protein while compromising its ability to bind to the mental retardation associated protein ATRX; and the p.A140V mouse model showed increased cell packing density and decreased neuronal dendritic arbors. This mutation was already described in 4 families with X-linked mental retardation and in two sporadic male patients with intellectual disability, indicating that MECP2 mutations are not necessarily lethal in males, and that this specific mutation seems to be recurrent and may be under-diagnosed in families with X-linked intellectual disability.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3219W

Whole exome sequencing identifies titin (*TTN*) mutations as a cause of centronuclear myopathy. *O. Ceyhan, P.B. Agrawal, K. Schmitz, E. DeChene, K. Markianos, A.H. Beggs.* Genetics, Children's Hospital, Harvard Medical School, Boston, MA.

Centronuclear myopathies (CNMs) are a group of rare and genetically heterogenous congenital myopathies, characterized by an abnormally high number of muscle fibres with internalized nuclei. To date, mutations in four genes involved in membrane trafficking and T-tubule assembly have been linked to CNMs; however, the genetic defects in 20% of the patients still remain unknown. We performed whole exome sequencing (WES) in a cohort of 28 CNM patients with no known genetic defects, and identified pathogenic mutations in the gene encoding the giant sarcomeric protein, titin, in 4 patients. All 4 patients are compound heterozygotes for unique frameshift, nonsense, or splicing mutations in the I- or A-band regions of the titin protein. Muscle biopsies show an increase in the number of cells with multiple internal nuclei, type I fiber hypotrophy/predominance, and disorganized sarcomeres. Clinical presentations range from birth-onset severe hypotonia with respiratory difficulties to milder childhood-onset muscle weakness, and do not involve cardiac abnormalities. Our results expand the wide phenotypic spectrum associated with *TTN* mutations, while suggesting novel mechanisms in the pathophysiology of CNMs. In addition, we demonstrate the value of WES for discovering mutations in genes that were previously insufficiently analyzed due to their large size.

3220F

Unexpected *EGFR3* variants confounding Achondroplasia mutation detection. *A. Millson[1], A.F. Rope[1,2], E. Lyon[1,3].* 1) ARUP Inst Clin & Exp Path, ARUP Labs, Salt Lake City, UT; 2) Department of Pediatrics, University of Utah, Salt Lake City, UT; 3) Department of Pathology, University of Utah, Salt Lake City, UT.

Achondroplasia is caused by a mutation in the fibroblast growth factor receptor 3 (*FGFR3*) gene located on chromosome 4p16.3 and is phenotypically expressed as short-limb dwarfism. Achondroplasia occurs in about 1/25,000 births worldwide. It is an autosomal dominant disorder with the majority of cases resulting from de novo sporadic mutations in the fibroblast gene. Homozygosity for the mutation results in a severe skeletal dysplasia that is lethal. The c.1138G>A mutation accounts for ~98% of sporadic cases with the c.1138G>C mutation causing ~1%. Both mutations result in G380R (Glycine to Arginine) amino acid substitution, giving rise to Achondroplasia. ARUP Laboratories offers testing for Achondroplasia utilizing PCR followed by Fluorescent Resonance Energy Transfer (FRET) probe analysis. Our protocol uses two reactions per sample, each reaction containing a FRET probe pair, with the reporter probe a perfect match to the wild type c.1138G or to the most common mutation c.1138A. Recently our clinical laboratory received a sample which after routine analysis was reported as wild type. The requesting clinician was suspect of the wild type result and sent the sample to another facility for sequencing, which found the less common mutation c.1138C. Subsequent sequencing at our facility confirmed the c.1138 mutation as well as identifying a variant under our clinical laboratory assay's forward primer, c.1097C>G, residing four bases from the 3' end. Presence of this variant inhibited amplification of the c.1138C mutation allele, therefore the 2 reside *in cis*. A search of various variant databases found no mention of c.1097C>G. This event led to an assay improvement incorporating a second set of unique primers to be used with the c.1138A matched FRET probe pair while keeping the original primer pair to be used with the c.1138G matched probe pair. During the assay improvement validation, which re-analyzed all prior samples sent in for testing (n=131), an aberrant melting pattern was seen with one heterozygous c.1138A sample. Sequencing of this sample found an intronic variant, c.1266+25C>T, located in the center of the newly designed reverse primer. A thorough search of variant databases failed to find mention of c.1266+25C>T. Here we describe two previously unreported *FGFR3* variants located near the c.1138G>A/C Achondroplasia causative mutation which have the potential to interfere with assay performance.

3221W

Analysis of C9orf72 repeat expansion in Japanese patients with ALS. *K. Ogaki[1], Y. Li[2], N. Atsuta[4], H. Tomiyama[1,3], M. Funayama[1,2], H. Watanabe[4], R. Nakamura[4], H. Yoshino[5], S. Yato[6], A. Tamura[7], Y. Naito[7,8], A. Taniguchi[7], K. Fujita[9], Y. Izumi[9], R. Kaji[9], N. Hattori[1,2,3], G. Sobue[4], the Japanese Consortium for Amyotrophic Lateral Sclerosis research (JaCALS).* 1) Dept of Neurology, Juntendo University School of Medicine, Tokyo, Japan; 2) Research Institute for Diseases of Old Age, Juntendo University School of Medicine, Tokyo, Japan; 3) Dept of Neuroscience for Neurodegenerative Disorders, Juntendo University School of Medicine, Tokyo, Japan; 4) Dept of Neurology, Nagoya University Graduate School of Medicine, Nagoya, Japan; 5) Setagaya Neurological Hospital, Tokyo, Japan; 6) Sayama Neurological Hospital, Sayama, Japan; 7) Dept of Neurology, Mie University Graduate School of Medicine, Tsu, Japan; 8) Dept of Neurology, Ise Red Cross Hospital, Ise, Japan; 9) Dept of Clinical Neuroscience, Institute of Health Biosciences, The University of Tokushima Graduate School, Tokushima, Japan.

Background and Purpose: Recently, a hexanucleotide repeat expansion in C9orf72 was identified as the most common cause of both sporadic and familial amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) in Western populations. The aim of this study was to investigate genetic and clinical characteristics of C9orf72 repeat expansion in a large cohort of Japanese patients with ALS. Methods: We analyzed 563 Japanese patients with ALS (552 sporadic and 11 familial) and 197 normal controls using fluorescent fragment-length analysis of C9orf72 and repeat-primed PCR analysis. Haplotype analysis was performed for 42 SNPs in patients with C9orf72 repeat expansion. Results: C9orf72 repeat expansion was found in two patients with sporadic ALS (2/552=0.4%) and no patients with familial ALS (0/11=0%). In the probands' families, one primary progressive aphasia (PPA) patient and one asymptomatic 76-year-old individual exhibited C9orf72 repeat expansion. The genotypes of all individuals with the C9orf72 mutation were detected for 24 or 25 consecutive SNPs in the 42 SNPs of original Finish risk haplotype. Conclusions: The frequency of the C9orf72 repeat expansion among Japanese patients is much lower than in Western populations. The existence of a 76-year-old asymptomatic carrier supported the notion of incomplete penetrance. The C9orf72 repeat expansion should be analyzed in sporadic ALS patients after determining their family histories not only of FTD but also of PPA. The haplotype bearing the C9orf72 repeat expansion was only shared in a narrow region between Western and Asian populations, suggesting that the C9orf72 repeat expansion may be an old mutation in human migration history from Western to East Asia.

3222F

Three pycnodysostosis cases with a novel mutation in cathepsin K gene. *T. Ozdemir[1], T. Atik[2], E. Karaca[1], H. Onay[1], F. Ozkinay[2], O. Cogulu[2].* 1) Ege University Faculty of Medicine, Department of Medical Genetics, Izmir, Turkey; 2) Ege University Faculty of Medicine, Department of Pediatrics, Division of Genetics, Izmir, Turkey.

Pycnodysostosis is a rare, autosomal recessive osteochondrodysplasia which is characterized mainly by short stature, abnormally high bone density, acroosteolysis, seperated cranial sutures with open fontanelles and bone fragility. It is caused by mutations in the cathepsin K gene (CTSK) which consists of 8 exons and results with impaired tissue specific expression of the lysosomal cysteine protease. Thirty-three mutations, associated with CTSK gene, have been reported in Human Gene Mutation Database. These mutations include particularly missense mutations, followed by frameshift, nonsense and splicing mutations. They mostly localize in exons 5 and 8, however only one case with frameshift mutation in exon 4 have been reported to date. Here we report 3 children (one female and 2 male) from 2 different families with c.354_355insT (p.V119Cfsx25) homozygous novel mutation in exon 4 of CTSK gene. The common clinical features of 3 cases were short stature, osteosclerosis, open fontanelles and sutures, beaked nose, hypoplasia of the jaws, obtuse mandibular angle, dental abnormalities, stubby hands and feet with acroosteolysis, dysplastic nails. Only one female case had a history of three fractures. In conclusion these cases make two contributions to the literature, firstly they are the first cases reported from Turkey in which a novel mutation in CTSK gene has been identified, secondly this novel mutation in three cases suggest a relatively mild clinical picture of pycnodysostosis.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3223W

Application of Next-Generation Sequencing for Mutation Detection in Autosomal Dominant Polycystic Kidney Disease. *H.C. Park[1,2], A. Kang[3], J.Y. Jang[3], Y. Hwang[2,4,5], H. Kim[1], M. Han[1], D.K. Kim[1], K. Oh[1], W. Park[2,6], H.I. Cheong[2,7], C. Ahn[1,2,3].* 1) Department of Internal Medicine, Seoul National University Hospital, Seoul, South Korea; 2) Research Center for Rare Diseases, Seoul National University Hospital, Seoul, South Korea; 3) Transplantation Center, Seoul National University Hospital, Seoul, South Korea; 4) Department of Internal Medicine, Eulji General Hospital, Seoul, South Korea; 5) Division of Genomic Medicine, Toronto General Hospital, Toronto, Canada; 6) Department of Biomedical Sciences, Seoul National University College of Medicine, Seoul, South Korea; 7) Department of Pediatrics, Seoul National University Children's Hospital, Seoul, South Korea.

Autosomal dominant polycystic kidney disease (ADPKD) is the most common hereditary renal disease characterized by multiple renal cysts and other extrarenal manifestations such as intracranial aneurysm and hepatic cysts. Although the genotype (PKD1 vs. PKD2) is the most important prognostic factor, it has not been so easy to detect mutation using conventional Sanger sequencing because of multiple duplicated regions (pseudogenes) along the chromosome 16 and a large mRNA transcript size of PKD1 with relatively high GC ratio. This study was performed to evaluate the efficacy of targeted exome sequencing in mutation detection of ADPKD compared to Sanger sequencing. We extracted DNA from whole blood of 24 ADPKD patients. ADPKD was typically diagnosed by family history and imaging such as ultrasonography or computed tomography with age-related criteria. We performed targeted exome sequencing using Agilent SureSelect in-solution enrichment system and verified the detected variants by Sanger sequencing methods. We designed the 120-mer RNA baits to cover the target exon three time (X3) and used hybridization target sequence capture strategy. The captured DNA was sequenced by ligation with the Illumina/GAIIx system. From the sequencing data, we aligned the sequences to the known reference sequence (GRCh37). In case of missense variants, we used PolyPhen-2, SIFT, and Mutation Taster to predict the functional effects. The process detected definitely and likely pathogenic variants in 21 (87.5%) out of 24 ADPKD patients. PKD1 mutations accounted for 81% (17/21) and PKD2 mutations for 19% (4/21). In addition, all of the mutations detected by novel next-generation sequencing method were confirmed by Sanger sequencing. Most of the PKD1 mutations were located at the exon 15 (12/25 mutations, 48.0%). Missense mutations accounted for 60% of PKD1 mutation; however, nonsense mutations and frameshift mutations were dominant in PKD2 mutations (80%). Conclusion: The targeted exome sequencing may be the promising tool to find out causative mutation in ADPKD in a high-throughput and an accurate way. We are currently working on the mutation data analysis from a larger ADPKD cohort and expect more comprehensive data to validate the efficacy of the targeted exome sequencing.

3224F

Whole exome sequencing identifies a missense mutation in *SEPT2* as a probable cause of a new autosomal dominant syndrome with distinctive face, ear anomalies, and learning disability. *A. Rump[1], K. Hackmann[1], A. Dahl[2], A. Fischer[3], M. Schweiger[3], M. Schilhabel[4], A. ElSharawy[4], A. Franke[4], E. Schrock[1], N. Di Donato[1].* 1) University of Technology Dresden, Institute of Clinical Genetics, Dresden, Germany; 2) University of Technology Dresden, Biotechnology Center, Dresden, Germany; 3) Max-Planck Institute for Molecular Genetics, Department of Vertebrate Genomics, Berlin, Germany; 4) Christian-Albrechts-University, Institute of Clinical Molecular Biology, Kiel, Germany.

Background: Recently, we have presented clinical data of three patients from two unrelated families with a unique combination of clinical features, including prominent eyes and bilateral ptosis. Other features include cleft palate, hearing loss, heart defects, and mild developmental delay. Since the condition is transmitted directly from mother to daughter (family 1) and observed in a male patient (family 2), autosomal dominant inheritance is suggested (Tyshchenko, N. et al. (2011) Am J Med Genet 155A (9):2060–2065). Method: Four members of family 1 (the affected mother, her affected daughter and her healthy parents) were analyzed by whole exome sequencing on a SOLiD4 system. Results: We identified a heterozygous missense mutation in the *SEPT2* gene. The mutation is present in the two affected individuals and absent in the healthy family members as well as in a large cohort of healthy, unrelated control individuals. According to both polyphen-2 and mutation taster, the mutation is damaging with a probability score of >0.999. The *SEPT2* gene encodes a filament-forming cytoskeletal GTPase. Unlike Ras-type GTP-binding proteins, septins form core oligomeric complexes and the human septin core structure is composed of a linear, nonpolar hexamer SEPT7-SEPT6-SEPT2-SEPT2-SEPT6-SEPT7. Sequencing of *SEPT2*, *SEPT6* and *SEPT7*, however, did not yield a plausible mutation in our second family. Conclusion: Septins have very recently been shown to regulate the collateral branching of axons (Hu et al., Curr Biol. 2012 May 16. [Epub ahead of print]). Since SEPT2 is also known to act as a scaffold of myosin II, the *SEPT2* gene is a plausible candidate for causing the clinical features of our patients upon mutation. Currently, we are investigating the effect of our *SEPT2* mutation on axon formation in retinoic acid induced SH-SY5Y cells. Since the affected boy of the second family did not show mutations in the septin genes mentioned above, family 2 is currently being sequenced on exome level.

3225W

Oculocutaneous Albinism (OCA1A/B) - founder mutations in the tyrosinase gene in Colombia. *O. Urtatiz, D. Sanabria, M.C. Lattig.* Universidad de los Andes, Bogota, Colombia.

Oculocutaneous albinism (OCA) is a group of inherited disorders of melanin biosynthesis characterized by a complete lack or generalized reduction in pigmentation of hair, skin and eyes. Mutations in the tyrosinase gene TYR on chromosome 11q14.3, are responsible of OCA1 (MIM203100) with OCA1A and OCA1B subtypes. The gene consists of 5 exons spanning about 65 kb of genomic DNA and encoding a protein of 529 amino acids. Mutations completely abolishing tyrosinase activity result in OCA1A, while mutations rendering some enzyme activity result in OCA1B allowing some accumulation of melanin pigment over time. In Colombia, the first study reported by Sanabria, et al. identified a missense G47D mutation in two unrelated individuals in heterozygous state. This mutation has already been reported in Puerto Rico and Sephardic/Moroccan Jews populations despite being two ethnically distinct populations. In this study we found the G47D mutation in 3 other unrelated Colombian individuals with OCA1A, and was not present in one individual with OCA1B. Two of the three OCA1A individuals were homozygous for the G47D mutation and the remaining individual was a compound heterozygote: G47D mutation and a novel nonsense mutation (S184X). This novel mutation is a c.551C-G nucleotide substitution in exon 1 resulting in an aminoacid change of serine to stop codon (S184X). The OCA1B individual also presented the novel mutation (S184X) in a heterozygous state, however the remaining mutation in the homologous allele was undetermined in TYR gene. According to this, we suggest that the G47D mutation in Colombian individuals is due to a founder mutation. This is reasonable if we take into account that haplotype analyses have shown that G47D mutation in individuals from Puerto Rico as well as Moroccan Jews share a common origin. Probably the G47D mutation found in Colombian individuals may have the same ancestral origin as the Sephardic Jews and Puerto Ricans, taking into consideration that Sephardic and Moroccan Jews had several waves of emigration arriving to America in the fifteenth century, fleeing persecution. The Canary Islands and Puerto Rico were part of the places they visited in their exploration trips thus probably brought the G47D mutation to the Caribbean and north of South America.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3226F

Ophthalmologic findings in Mexican patients with Myotonic Dystrophy Type I. *P. Barojas[1], JJ. Magaña[2], G. Ortega[1], N. Leyva[2], B. Cisneros[3], O. Hernández-Hernández[2], E. Barojas[1].* 1) Ophthalmology, National Rehabilitation Institute, Mexico City, Mexico City, México; 2) Department of Genetics, INR, Mexico City, México; 3) Genetics and Molecular Biology, CINVESTAV-IPN, Mexico City, México.

Myotonic dystrophy type 1 (DM1) is a dominantly neuromuscular inherited disorder caused by the expansion of an unstable CTG tract in the 3'UTR of the DMPK gene. Clinical manifestations of DM1 involve a great number of tissues including the ocular system. Several visual system alterations in DM1 have been described including cataracts, decreased intraocular pressure and ocular muscles alterations that result in ophthalmoplegia, bilateral motility disturbance, obicularis oculi, levator muscle weakness, and ptosis. To investigate visual system alterations in Mexican DM1 patients, an extensive ophthalmologic evaluation was performed. A sample of 44 individuals clinically diagnosed as DM1 were subjected to molecular test, applying a fluorescent PCR method in combination with capillary electrophoresis analysis. Patients were divided according with the CTG expansion on three groups: premutation, small mutation and greater than 100 CTG repeats. Opthalmologic applied tests included visual capacity, autokeratorefractometry, gaze positions and diplopia, examinations of lids, cornea, anterior chamber, iris and light response. Macula, retinal periphery, pachimetry, keratometries, lens density, endothelial cell count and macular optical coherence tomography (OCT) were also documented. Severe alterations were found in the group of higher CTG expansions: 50% of patients presented cortical punctate opacities in lens (previously undescribed), 30.7% showed impairment of epiretinal membrane (MER), meanwhile 11.5% showed neovascular membrane (MNV) damage. 80% of the patients presented cataract with a mean age of 35 years old, however we did not observe a christmas tree cataract in any subject. Interestingly, OCT examination showed hyperreflective dots on the inner layers of the retina in all of our patients including the premutation group. These results open a new avenue into the study of the cellular and molecular mechanisms to understand how hyperreflective dots in the retina are produced even in presence of small expansions. In addition, this sign on the retina could be used in the future as an early ophthalmologic marker of the disease.

3227W

Molecular diagnosis of Myotonic dystrophy type I by PCR Southern method. *GH. Kim[1], JJ. Lee[1], SH. Choi[1], JY. Lee[1], JM. Kim[1], YM. Kim[2], BH. Lee[1,2], HW. Yoo[1,2].* 1) Med Gen Clinic & Lab, Asan Med Ctr, Seoul, South Korea; 2) Asan Medical Center Children's Hospital, University of Ulsan College of Medicine, Seoul, Korea.

Myotonic dystrophy type 1 (DM1; MIM 176270) is a progressively degenerative skeletal muscular disease. Based on clinical severity, this condition is divided into three phenotypes, mild, classic, and congenital. It is inherited in an autosomal-dominant manner and caused by abnormal expansion of trinucleotide repeat, (CTG)n, in the 3'-UTR of Dystrophia myotonica-protein kinase (DMPK) gene located on 19p13.3. The aim of this study was to verify whether PCR Southern method is a useful molecular diagnostic tool for DM1. During the period of 2000 to 2011, 1,243 unrelated patients were requested for the genetic diagnosis of DM1. To assess the accuracy of our PCR Southern method, 54 DNA samples, obtained from Collage of American Pathology (CAP) molecular genetics survey program, were used as positive (18 samples) or negative (36 samples) controls. Genomic DNA was isolated from each patient's peripheral leukocytes. After PCR amplification, each product was separated by agarose gel electrophoresis, transferred to nylon membrane, and hybridized by biotin labeled (CTG)10 probe. Each band was exposed on X-ray film using streptoavidin conjugated alkarine phosphatase. The copy number of each band was measured according to DNA size marker. A total of 658 patients (52.9%) had abnormal (CTG)n in the DMPK. Twenty eight patients had congenital type (4.3%), 611 patients classic type (92.9%), and the remaining 15 patients had mild or premutation type (2.3%). The (CTG)n was inversely correlated with the request age of patients (R=-0.418, P<0.01). Maternally originated anticipation was observed in 15 families. The median fold change of (CTG)n in the offsprings to those in their mothers was 5.1. The sensitivity and specificity of PCR Southern method were all 100%, when this method was applied to positive and negative controls from CAP. The calculated copy numbers, however, were approximately 100 copies lower in our results than in CAP results, although these differences didn't influence the classifications of phenotypic subtypes according to our PCR Southern method. The limitation of PCR-based method is that large (CTG)n is difficult to detect by PCR method, even by triplet repeat primed PCR method, recently introduced to improve the detection efficiency of (CTG)n. In addition, the exact number of (CTG)n is difficult to measure due to cellular mosaicism. However, having very high sensitivity and specificity, PCR-Southern method is a proper molecular diagnostic tool for DM1.

3228F

Mutations in FOXP2 cause childhood apraxia of speech: report of a novel intragenic deletion in an Australian patient. *M.S. Hildebrand[1], S. Turner[2], J. Damiano[1], R.J.H. Smith[3], M. Bahlo[4], I.E. Scheffer[1,5], A.T. Morgan[2].* 1) Epilepsy Research Centre, Department of Medicine, University of Melbourne, Austin Health, Heidelberg, Victoria 3084, Australia; 2) Population Health, Genes and Environment Theme, Murdoch Childrens Research Institute & Department of Paediatrics, University of Melbourne, Parkville, Victoria 3052, Australia; 3) Department of Otolaryngology - Head and Neck Surgery, University of Iowa City, IA, 55242, USA; 4) Bioinformatics Division, The Walter and Eliza Hall Institute of Medical Research & Department of Mathematics and Statistics, The University of Melbourne, Parkville,Victoria 3052, Australia; 5) Florey Neuroscience Institutes and Department of Paediatrics, University of Melbourne, Royal Children's Hospital, Parkville, Victoria 3052, Australia.

Developmental speech and language problems are a heterogeneous group of disorders that can prove difficult to manage clinically. In most cases the etiology is unknown, although in a minority genetic variants have been implicated or shown to confer susceptibility. The first specific subtype to be directly linked to a molecular genetic defect was childhood apraxia of speech (CAS) due to mutation in the FOXP2 gene at the SPCH1 locus. We completed recruitment of eight Australian families where the proband had a primary diagnosis of motor speech disorder. Affected family members were phenotyped using a comprehensive assessment battery that comprised measures of speech, oromotor function, language, literacy skills and cognition. We screened the coding regions of the FOXP2 gene in all eight probands and identified novel variants in two. We then determined the inheritance pattern and segregation of the variant in the probands' families. Here we describe these variants including the first reported small, intragenic deletion of the FOXP2 gene. Our results support the current evidence that FOXP2 mutation is a well-defined cause of CAS, and extend the genotype-phenotype spectrum of FOXP2-related speech and language disorders.

3229W

Exome sequencing identifies a novel *MRE11* mutation in a patient with generalized myoclonic tremor. *R. Miyamoto[1,2], H. Morino[1], H. Maruyama[1], Y. Izumi[2], R. Kaji[2], H. Kawakami[1].* 1) Department of Epidemiology, Research Institute for Radiation Biology and Medicine, Hiroshima University, Hiroshima, Japan; 2) Department of Clinical Neuroscience, Institute of Health Biosciences, Graduate School of Medicine, University of Tokushima, Tokushima, Japan.

[Background] Ataxia-telangiectasia-like disorder (ATLD) is a rare autosomal recessive disorder caused by mutations in the *MRE11* gene. ATLD is characterized by slowly progressive ataxia, oculomotor apraxia, and cerebellar atrophy. To date, little is known about other clinical features associated with *MRE11* mutations. [Patient] The patient is a 52-year-old female from consanguineous parents. She developed a jerky neck movement at age 9. Three years later, she showed a hand tremor, which gradually expanded to the entire body. On physical examination at age 47, she presented short stature and bilateral ptosis with a masked face. She showed generalized myoclonic tremor that was exacerbated by emotional stress. Brain MRI demonstrated no signs of cerebellar atrophy. EEG was within normal limits. SEP and jerk-locked back averaging failed to show the evidence of cortical myoclonus. Surface electromyogram demonstrated 4Hz unsynchronized bursts in her limbs. [Methods] The genetic study was approved by the ethics committee, and the patient provided written informed consent. Genomic DNA libraries were prepared using the SeqCap EZ Human Exome Library v2.0. Sequencing was performed with 100bp paired-end reads on a HiSeq2000. The reads were aligned to the human reference genome with BWA. Variants were called with GATK, and were annotated with Annovar using dbSNP build 135. Pathogenicity predictions for variants were obtained from the algorithms, including SIFT and Polyphen-2. Sequencing depth was calculated using GATK. All reported genomic locations are from GRCh37/hg19. Validation of the sequence variant was assessed with Sanger sequencing. [Results] Exome sequencing produced 129 million sequencing reads for the sample, comprising 12 billion bases. Approximately 88% of the sequencing reads were aligned to hg19. The average depth was 116. We identified 80,852 variants located in exonic or splicing regions. After filtering these variants in a reasonable manner, a homozygous missense sequence variant in the *MRE11* gene (NM_005590:c.140C>T:p.A47V) was left as a candidate. This amino acid change was predicted to be damaging by all the prediction algorithms. [Conclusions] This case, with the novel *MRE11* mutation, predominantly showed myoclonic tremor. Clinical symptoms associated with *MRE11* mutations could be more numerous than previously thought. Exome sequencing could be a powerful method of diagnosing very rare diseases from a single sample.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3230F

Pathophysiological features of dermatan 4-O-sulfotransferase 1 (D4ST1)-deficient Ehlers-Danlos syndrome (DD-EDS). *T. Kosho[1], S. Mizumoto[2], M. Kobayashi[3], Y. Fujita[4], J. Nakayama[5], N. Miyake[6], Y. Nomura[7], A. Hatamochi[8], Y. Fukushima[1], K. Sugahara[2], N. Matsumoto[6].* 1) Medical Genetics, Shinshu University School of Medicine, Matsumoto, Japan; 2) Laboratory of Proteoglycan Signaling and Therapeutics, Hokkaido University Graduate School of Life Science, Sapporo, Japan; 3) Department of Food and Nutritional Environment, College of Human Life and Environment, Kinjo Gakuin University, Nagoya, Japan; 4) Division of Medical Research Engineering, Nagoya University Graduate School of Medicine, Nagoya, Japan; 5) Department of Molecular Pathology, Shinshu University Graduate School of Medicine, Matsumoto, Japan; 6) Department of Human Genetics, Yokohama City University Graduate School of Medicine, Yokohama, Japan; 7) Scleroprotein and Leather Research Institute, Tokyo University of Agriculture and Technology, Fuchu, Japan; 8) Department of Dermatology, Dokkyo Medical University, School of Medicine, Mibu, Japan.

Dermatan 4-O-sulfotransferase 1 (D4ST1)-deficient Ehlers-Danlos syndrome (DD-EDS), caused by recessive loss-of-function mutations in CHST14, is a recently delineated form of EDS [Kosho et al., 2011], independently identified as “adducted thumb-clubfoot syndrome” [Dündar et al., 2009], “EDS, Kosho Type” [Kosho et al., 2010; Miyake et al., 2010], and “Musculocontractural EDS” [Malfait et al., 2010]. DD-EDS is characterized by a unique set of clinical features consisting of progressive multisystem fragility-related manifestations (skin hyperextensibilty/fragility, progressive spinal/foot deformities, large subcutaneous hematoma) and various malformations (facial features, congenital eye/heart/gastrointestinal defects, congenital multiple contractures). The cause of multisystem connective tissue fragility is postulated to be impaired assembly of collagen fibrils based on glycobiological analysis on patients' skin fibroblasts showing loss of dermatan sulfate replaced by chondroitin sulfate in the decorin glycosaminoglycan (GAG) side chain. However, most of the pathophysiology is still uncovered including skin pathology related to decorin and glycobiological effects on cells except skin fibroblasts. In this study, we report novel pathological and glycobiological findings. Light microscopy on three patients' skin specimens by Azan staining showed waving of the epidermis and fineness of collagen bundles in the dermis, compared with a normal control. Electron microscopy on a patient's skin specimen using cupromeronic blue, directly staining GAG side chains of proteoglycan, demonstrated shorter GAG side chains than those in a patient with another type of EDS. Decorin staining on four patients' skin specimens using human decorin antibody, specifically binding to decorin core protein, demonstrated filamentous staining along collagen bundles, markedly different from a normal control showing uniformly stained collagen bundles. Array-based expression study demonstrated increased or decreased expression of several genes, though the changes were not significant. Glycobiological study on five patients' urine specimens, using anion-exchange HPLC after digestion with chondroitinases ABC/AC/B, demonstrated complete loss of dermatan sulfate. These data present direct evidence on conformational abnormalities between decorin core protein and GAG side chains on collagen fibrils and demonstrate that generalized complete loss of dermatan sulfate is the essence of this disorder.

3231W

Rupture of chordae tendineae as an initial presentation of an haploinsufficiency mutation in COL3A1. *Z. Xu[1], B.F. Griswold[1], L.J. Sloper[1], A.S. Shah[2], N.B. McDonnell[1].* 1) National Institute on Aging, Baltimore, MD; 2) Johns Hopkins University, Baltimore, MD.

COL3A1 haploinsufficiency mutations usually result in milder phenotypes of Vascular form of Ehlers Danlos syndrome as compared to missense or splicing mutations. We describe the case of a 31 year old male who developed acute mitral insufficiency due to rupture of chordae tendineae during jogging. During the repair the mitral valve, tissue fragility was noted by the surgeon and the patient was referred for testing for pathogenic mutations in COL3A1. Sanger sequencing was performed for the COL3A1 gene. A novel nonsense mutation c.2569C>T, p.Q857X was identified in one allele of COL3A1 gene, leading to haploinsufficiency of procollagen III. Subsequent work showed that this mutation was inherited from the patient's father who had a history of aortic aneurysm. The mutation was also present in the patient's two young daughters and his brother. This case illustrates an unusual initial presentation for the Vascular form of Ehlers Danlos syndrome. The observation of tissue fragility by the cardiothoracic surgeon was the key to the correct diagnosis and genetic counseling of this family. Rupture of chordae tendineae in an otherwise healthy and young individual may be the first key to the identification of a hereditary connective tissue disorder in a family.

3232F

Prenatal diagnosis and identification of heterozygous frameshift mutation in *PRRX1* in an infant with agnathia-otocephaly. *M. Donnelly[1], E. Todd[1], M. Wheeler[1], V. D. Winn[1], D. Kamnasaran[2,3].* 1) Department of Obstetrics and Gynecology, University of Colorado, Aurora, CO; 2) Pediatrics Research Unit, Centre de recherché du CHUL, Quebec, Canada; 3) Department of Pediatrics, Laval University Quebec, Canada.

Recently, several infants diagnosed with agnathia-otocephaly, a severe disorder of first pharyngeal arch development, have been discovered to have loss of function mutations in the Prrx1 gene, which encodes a protein in the paired homeobox family important for facial development. We report the 3-D ultrasound prenatal diagnosis of an infant with agnathia-otocephaly, as well as a novel heterozygous frameshift loss-of-function mutation of PRRX1 in this infant. With sequencing of genomic DNA of this infant, a heterozygous deletion of an adenosine (c. 267 del A) was discovered in a poly A tract in the PRRX1 gene. This deletion is absent in both parents, and in normal individuals (n=100 tested). Subsequent functional assays with the wildtype PRRX1 protein (245 amino acid isoform) demonstrated a significant repression of the human tenascin-c (TNC) promoter fragment (-516 to +195) in COS7 cells. Such repression is also seen in COS7 cells transfected with the pcDNA3.1 vector only, since basal levels of wildtype PRRX1 are expressed endogenously. However, the PRRX1 protein with the p.Lys90Argfs*131 frameshift mutation demonstrated a significant inability to repress the human tenascin-c promoter fragment, and is therefore a loss-of-function mutation. This case provides additional evidence that the Prrx1 gene is important in the development of agnathia-otocephaly, and further work may lead to development of testing for use in prenatal diagnosis beyond ultrasound imaging.

3233W

Identification of the disease causing gene in a familial autosomal recessive form of congenital nystagmus by exome sequencing. *B. Isidor, J. Albuisson, S. Bezieau.* Service de génétique médicale, CHU Nantes, Nantes, France.

Congenital or infantile nystagmus (CN) is a rare disabling genetic condition, presenting as conjugate, horizontal oscillations of the eyes. It may be associated with torticollis, decreased visual acuity, strabismus, astigmatism, and occasionally head nodding. This heterogenous disorder can be transmitted in an X-Linked fashion, as a consequence of the FRMD7 gene mutations, leading to a pure, neurologic CN. To date, several loci have also been identified in autosomal dominant forms of CN, but no other causative gene could be identified. Recessive CN was exceptionally described and has never been investigated by genetic analyses. We present here our strategy to identifiy the gene responsible for autosomal recessive CN in a French family. The CN kindred is represented by three girls from healthy, apparently unrelated parents. Two of the girls had CN. No mutation could be identify by Direct sequencing of the FRMD7 gene in one of the affected girl. We also believe that a new gene could be responsible fro this very rare form of CN. Exome sequencing in 2 affected patients of this family and one healthy sister will probably lead to the identification the disease causing gene.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3234F

Identification of X-Linked RPGR ORF15 Mutations in Females Diagnosed with Autosomal Dominant Retinitis Pigmentosa. *J.D. Churchill[1], S.J. Bowne[1], L.S. Sullivan[1], R.A. Lewis[2], D.K. Wheaton[3], D.G. Birch[3], K.E. Branham[4], J.R. Heckenlively[4], S.P. Daiger[1].* 1) Human Genetics Center, UTHSC Houston, Houston, TX; 2) Department of Ophthalmology, Baylor College of Medicine, Houston, TX; 3) The Retina Foundation of the Southwest, Dallas, TX; 4) Kellogg Eye Center, University of Michigan, Ann Arbor, MI.

The purpose of this study is to identify the fraction of families with a provisional diagnosis of autosomal dominant retinitis pigmentosa that have disease-causing mutations in the X-linked retinitis pigmentosa GTPase regulator (*RPGR*) gene. Retinitis pigmentosa (RP) is an inherited retinal degeneration that affects approximately 1.5 million people world-wide. RP can be inherited in an autosomal dominant, autosomal recessive, or X-linked manner. However, X-linked RP (XlRP) may be mistaken for autosomal dominant RP (adRP) since some X-linked mutations cause disease in female carriers. Mutations in the *RPGR* gene are the most common causes of XlRP, accounting for over 70% of XlRP, and approximately 60% of disease-causing mutations in *RPGR* are found in ORF15. However, screening ORF15 of *RPGR* in females is technically difficult due to the presence of polymorphic insertions and deletions that often make sequence alignment impossible. To facilitate testing of *RPGR* in “adRP” families, the repetitive and purine-rich ORF15 of *RPGR* was subcloned and sequenced in females from 16 unrelated families that had a diagnosis of adRP, no male-to-male transmission, and no males available for testing. Analysis identified a disease-causing mutation (a single-nucleotide duplication) in one of the families examined in this study. In previous studies, we selected families with a provisional diagnosis of adRP and no male-to-male transmission from a cohort of 237 families and screened these families for mutations in *RPGR*. Eighteen of these “adRP” families were found to have *RPGR* mutations. Combining these data with data from our recent study shows that approximately 8% (19 in 237) of families thought to have adRP truly have XlRP. These results have significant implications for calculation of recurrence risk, counseling, and treatment and illustrate the importance in screening families with a provisional diagnosis of autosomal inheritance and no male-to-male transmission for mutations in X-linked genes. Additionally, these results show that for families in which males are not available for testing, subcloning and sequencing of ORF15 is a reasonable option for molecular diagnosis.

3235W

A recurrent 143 kb duplication in TSPAN7 as a possible cause of intellectual disability. *J. Hoyer, M. Krumbiegel, C. Kraus, A. Reis.* Institute of Human Genetics, University Erlangen-Nuremberg, Erlangen, Germany.

We identified with molecular karyotyping on a high density microarray a microduplication of 143 kb on chromosome Xp11.4 comprising only Tetraspanin 7 (TSPAN7) in a seven 7 years old boy of German origin with mild intellectual disability (ID) and body measurements within the normal range. We detected the same duplication in another male with body measurements below the 3rd centile, mild ID, particularly difficulties in expressive and perceptive speech, strabismus and suspected Attention Deficit Hyperactivity Disorder (ADHD). Both boys had normal male karyotypes and Fragile-X syndrome was excluded. The duplications were shown to be inherited from normal mothers coming from Germany and Kyrgyzstan respectively. X inactivation studies indicated a random pattern in both women. Abidi and colleagues reported in 2002 the implication of TSPAN7 in nonsyndromic X-linked MR (MRX58). The affected males had mild to moderate mental retardation and no consistent dysmorphic features, while obligatory carrier females were normal and showed a random pattern of X inactivation. TSPAN7 is highly expressed in the central nervous system and encodes a member of the transmembrane 4 family of proteins, which are known to contribute in molecular complexes that include beta-1 integrins. Zemni and colleagues speculated that through this interaction, TSPAN7 may play a role in the control of neurite outgrowth. The duplication identified in our two patients was not found among a cohort of 820 controls and is not listed in the database of genomic variants suggesting the copy number variation to be a recurrent alteration playing a pathogenic role in male ID. To support this hypothesis systematic analyses of members of family 1 are ongoing.

3236F

Alpha thalassemia /mental retardation X linked (ATRX) - an uncommon presentation. *U.H. Kotecha[1], R. Puri[1], C. Badens[2], I.C. Verma[1].* 1) Center of Medical Genetics, Sir Gangaram Hospital, New Delhi, India; 2) Laboratoire de Génétique Moléculaire, Département de Génétique Médical, Hôpital d'Enfants de la Timone, Marseille, France.

X-linked alpha thalassemia /mental retardation syndrome (ATRX) occurs due to mutations in the ATRX gene. Cardinal features are mental retardation, characteristic facies, hematological and genital abnormalities. Sporadic cases without the above are difficult to diagnose, especially in infancy as the facial features are still evolving. We describe a child with an uncommon phenotypic presentation of ATRX. A one year old boy was brought for evaluation of global developmental delay. Developmental milestones were limited to social smile and some vocalization. He was born at term to a non-consanguineous Asian Indian couple and had no history of birth asphyxia. There were no seizures. The family history was non contributory. Examination revealed microcephaly, and subtle dysmorphism (telecanthus, bilateral epicanthal folds and depressed nasal bridge). Contractures of distal interphalangeal joints of the third and fourth fingers of the right hand and all the fingers of the left as well as bilateral cryptorchidism were present. An MRI of the brain and ophthalmological examination showed cortical atrophy and myopia respectively. Karyotype was normal. A possibility of distal arthrogryposes atypical facies syndrome was considered. Since no gene for the same is known, patient was advised supportive management and annual follow up. He came for review after three years. On examination a frontal upsweep, small forehead, low set ears, hypertelorism, epicanthal folds, tented upper lip with thick lower lip were present. There was no progression of joint involvement. ATRX syndrome was suspected clinically. Haematological investigations revealed hypochromic microcytic anemia but HbH inclusions were absent. Molecular analysis of ATRX gene confirmed a novel missense mutation in exon9 c.A224E. This mutation is located in the zinc finger domain and is expected to affect splicing. The zinc finger region is located in the ADD domain which is a highly conserved segment and represents a major mutation site. Mutations in this region are known to be associated with severe mental retardation. Though microcephaly and some coarsening of facies was noted at the first visit, distal joint contractures were a predominant feature and hence delayed the diagnosis of ATRX. Though fixed flexion deformities of distal joints have been noted in ATRX, the above case highlights the difficulty in diagnosis in infants with this presenting feature in sporadic cases.

3237W

Proteolipid protein 1 (PLP1) gene mutation in 44 Chinese patients with Pelizaeus-Merzbacher disease and prenatal diagnosis of five fetuses in three Chinese families with PMD probands. *J. Wang[1], D. Li[1], Y. Wu[1], H. Zhao[1,2], J. Shang[1,2], J. Qin[1], F. Fang[3], X. Wu[1], Y. Jiang[1].* 1) Dept Pediatrics, Peking Uni First Hosp, Beijing, China; 2) Department of Neurology, First Hospital Affiliated to Shanxi Medical University, Taiyuan; 3) Beijing Childrenès Hospital, Beijing, China.

Purpose The object of this study was to identify Proteolipid protein 1(PLP1) mutations in 44 unrelated Chinese patients (P1-44) with Pelizaeus-Merzbacher disease (PMD) and prenatal diagnosis of five fetuses in three Chinese families with PMD probands. Methods Genomic DNA was extracted from peripheral bloods samples. At 19 or 20 weeks gestation, amniotic fluid sampling (AFS) was performed. Gene dosage was determined by Multiplex Ligation-dependent Probe Amplification (MLPA). All 7 exons and exon-intron boundaries of PLP1 gene were amplified and analyzed by direct DNA sequencing. Results Of these 44 patients, there were 24 classical, 15 transitional, and 5 congenital PMD according to the clinical and radiological presentation. PLP1 duplications were identified in patients 1–30 with PMD, account for 75% (30/44). 8 hemizygous missense mutations including c.517C to T (p. P173S), c.709T to G(p.F237V), c.623G to T(p.G208V), c.467C to T(p.T156I), c.353C to T(p.T117R) and c.646C to T(p.P214S), c.96C to G(p.F32L), IVS5-1G to A were found in Patients31–38 with PMD(18.2%). 6 (13.6) were not found any changes with PLP1. For the results of prenatal diagnosis, the three fetuses in family 1 showed normal PLP1 dosage and chromosome analysis demonstrated a 46,XX and two 46,XY chromosome pattern, respectively. The fetus in family 2 revealed three cope of PLP1 which suggested a PLP1 duplication carrier after chromosome pattern was a 46, XX. In family 3, duplication of PLP1 was detected in the fetus imply as a PMD patient with 46, XY chromosome pattern. Conclusions We identified 30 genomic duplications and six missense mutation of PLP1 gene in 38 Chinese patients with PMD, account for 86.4% (38/44). Prenatal diagnoses for five fetuses in three PMD families were performed, which is useful and helpful for those families. This is the first report about PLP1 mutations and prenatal diagnosis in PMD patients in mainland of China.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3238F

Four Years of Data from the California Cystic Fibrosis Newborn Screening Program and Experiences of the Sutter CF Center. *M. Tsang[1], B. Chipps, MD, FAAP, FAAAI, FCCP[2], M. Kharrazi, PhD[3], K. Pearson, NPN, MS[1], S. O'Bra, BSHS[1].* 1) Sutter Medical Center 1625 Stockton Blvd. #112 Sacramento, CA 95816; 2) Capital Allergy and Respiratory Disease Center 5609 J. St. Suite C Sacramento, CA 95819; 3) Genetic Diseases Screening Program California Department of Public Health 850 Marina Bay Parkway Rm F175 Richmond, CA 94804.

The purpose of population-wide newborn screening for CF is to improve health and survival through early preventative care and therapeutic interventions. Every state in the US now screens for CF but the screening protocol differs for each state and as a result, the detection rates, false positives, and false negatives vary widely. California started screening newborns for cystic fibrosis in July 2007. The program is unique and includes comprehensive genotyping of the cystic fibrosis transmembrane conductance regulator gene (CFTR) and long term follow up of hypertrypsinogenemic infants with two or more CFTR mutations regardless of sweat chloride results. When indicated, parent testing is provided to determine if the mutations are in cis or trans phase. Sutter Medical Center is an approved specialty care center for CF located in Sacramento, CA. Staff have seen 56 newborns with a positive screen to date. Of these, 24 have a diagnosis of CF (including 2 Hispanic patients with novel mutations), 17 have CFTR-Related Metabolic Syndrome (CRMS), 13 are carriers, 2 are pending and 1 has transferred to another center. A designation of CRMS is made based on the patient's mild clinical manifestations, results from sweat chloride, fecal elastase and other tests, and may take several months to a year to determine. Every new patient at the Center receives a phone call by the genetic counselor (GC) to review results within days of a positive screen and to discuss preparation for the sweat chloride test. At the first visit, the family is seen by the GC, nurse practitioner, and dietician. At subsequent visits, the MD, social worker and respiratory therapist are also involved with the family. The novel mutations/variants have provided the most challenges for us in that there is very little information regarding symptoms, organ involvement, and degree of severity. Most of these patients have CRMS and are doing well. We have learned that keeping the families informed and being optimistic while stressing the importance of consistent follow-up with our CF Center and pulmonologist, so early treatment can be implemented, is the best approach. In conclusion, the CA CF Screening Program has a screened more than 2 million babies in 4 years with a 93% detection rate and 22 missed cases. The ratio of screen positives to cases detected is 2.9 to 1. The Sutter CF Center is a family oriented clinic dedicated to providing comprehensive care and treatment to our patients with CF and CRMS.

3239W

Polymicrogyria awareness is a critical first step in facilitating early diagnosis and intervention for those affected by this rare congenital brain anomaly. *C. Byrge[1], A. Perszyk[2], PMG Awareness Organization, Inc.* 1) PMG Awareness Organization, Inc., Jacksonville, FL; 2) Univ. of Florida - Jacksonville, Medical Genetics, Jacksonville, FL, USA.

The lack of polymicrogyria awareness is a major factor in the late diagnosis of those affected. Both patient and doctor education are needed to improve medical care and patient outcomes. Polymicrogyria is an under recognized medical condition that can be present as a congenital, developmental brain disorder. Both environmental and genetic etiologies have been described in the literature. Many forms have been identified as genetic and therefore have an increased risk of recurrence. Recently, we surveyed patients or families of those diagnosed with polymicrogyria. Ages at the time of diagnosis ranged from pre-term infants to adults. We report here the data on over 100 individuals. The majority of patients had more than a 6 month delay in diagnosis of their brain malformation. Feeding and aspiration risks, gastroesophageal reflux and the use of tube feedings were common. Hypotonia and non-specific developmental delays were seen in half. Seizures, often refractory, also present in approximately 35 percent; of those affected. Seizures that resulted in a referral to a pediatric neurologist typically expedited diagnosis. Family support and education is important. Among those diagnosed, few families recalled any specific education about the diagnosis, genetic testing, or seizure precautions. Of critical note, the survey identified that 85 percent of their pediatricians were not familiar with polymicrogyria; therefore, over 44 percent of families stated that their pediatricians were not able to assist them in managing the care of their children. The majority of them also report that there is no single physician who manages the overall care of their children. Polymicrogyria is associated with perinatally acquired cytomegalovirus infections as one of the known environmental causes. This information was lacking in 90 percent of those surveyed. This situation suggests a need for more education regarding CMV modes of transmission and prevention on early prenatal visits. Based on this survey earlier diagnosis could be achieved and this can help to prevent some of the potential effects of polymicrogyria. Of special concern are the long term sequelae of uncontrolled seizure activity and of aspiration-related events. Much more research needs to be done regarding the unique treatment difficulties that are present for those affected by polymicrogyria. Further investigations are needed to explore survival data and long term outcomes.

3240F

Novel KIF7 mutations extend the phenotypic spectrum of acrocallosal syndrome. *A. PUTOUX[1,2], S. NAMPOOTHIRI[3], V. CORMIER-DAIRE[1,2,4], N. LAURENT[5], P. BEALES[6], A. SCHINZEL[7], D. BARTHOLDI[7], C. ALBY[2,8], S. THOMAS[1,2], N. ELKHARTOUFI[4], J. LITZLER[4], F. ENCHA-RAZAVI[1,2,4], R. KANNAN[9], A. MUNNICH[1,2,4], L. FAIVRE[10,11], N. BODDAERT[12], A. RAUCH[7], M. VEKEMANS[1,2,4], T. ATTIE-BITACH[1,2,4].* 1) INSERM U781 et fondation IMAGINE, Hopital Necker Enfants Malades, Paris, France; 2) Université Paris Descartes, Paris Sorbonne Cité, Paris, France; 3) Department of Pediatric Genetics, Amrita Institute of Medical Sciences, Kerala, India; 4) Département de Génétique, Hôpital Necker-Enfants Malades, Assistance Publique Hôpitaux de Paris (AP-HP), Paris, France; 5) Service d'Anatomie et de Cytologie Pathologique, Hôpital du Bocage, Dijon, France; 6) Molecular Medicine Unit, University College London (UCL) Institute of Child Health, London, UK; 7) Institute of Medical Genetics, University of Zurich, Schwerzenbach, Switzerland; 8) Service de Gynécologie - Obstétrique, Hôpital Necker Enfants Malades, AP-HP, Paris, France; 9) Department of Radiology, Amrita institute of Medical Sciences, Kerala, India; 10) Service de Génétique Médicale, Hôpital du Bocage, Dijon, France; 11) EA GAD, IFR Santé - STIC, Université de Bourgogne, France; 12) Service d'Imagerie Médicale, Hôpital Necker-Enfants Malades, AP-HP, Paris, France.

Acrocallosal syndrome (ACLS) is a rare recessive disorder characterised by corpus callosum agenesis or hypoplasia, cranio-facial dysmorphism, duplication of the hallux, postaxial polydactyly and severe mental retardation. Recently, we identified mutations in KIF7, a key component of Sonic hedgehog pathway, as responsible for this syndrome. We sequenced KIF7 in five suspected ACLS cases, one fetus and four patients, based on facial dysmorphism and brain anomalies. Seven mutations were identified at the KIF7 locus in these five cases, six of which are novel. We describe the first four compound heterozygous cases. In all patients, the diagnosis was suspected based on the craniofacial features, despite the absence of a corpus callosum anomaly in one and of polydactyly in another. Hallux duplication was absent in 4/5 cases. These results show that ACLS has a variable expressivity and can be diagnosed even in the absence of the two major features namely polydactyly or agenesis or hypoplasia of the corpus callosum. Facial dysmorphism with hypertelorism, prominent forehead in all, as well as vermis agenesis with brainstem anomalies (molar tooth sign) strongly indicated the diagnosis. KIF7 should be tested in less typical cases in whom craniofacial features are suggestive of ACLS.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3241W

Goldberg-Shprintzen Syndrome: KBP, the *KIAA1279* encoding protein, is located to cytoskeleton, but not to mitochondria. *L. Drévillon[1,2], A. Megarbane[3], B. Demeer[4], C. Matar[1,6], P. Benit[5], A. Briand-Suleau[1,2], J. Ghoumid[1,2], M. Nasser[1,2], V. Bodereau[1], M. Doco-Fenzy[7], P. Rustin[5], D. Gaillard[7], M. Goossens[1,2], I. Giurgea[1,2].* 1) AP-HP, Groupe hospitalier Henri-Mondor, Service de Biochimie et Génétique, Créteil, France; 2) INSERM U955, Equipe 11, and Université Paris-Est Créteil, Créteil, France; 3) Laboratory of Molecular Biology and Cytogenetics, Faculty of Medicine, University Saint-Joseph, Lebanon; 4) Pôle pédiatrie, Hôpital Nord, CHU Amiens, Amiens, France; 5) INSERM U676, Hôpital Robert Debré, Paris, France; 6) INSERM U955, Equipe 07, and Université Paris-Est, Créteil, France; 7) Service de Génétique, CHU, UFR de médecine, IFR53, Reims, France.

Goldberg-Shprintzen syndrome (GOSHS, MIM #609460) is a rare an autosomal recessive disorder associating severe intellectual disability, facial gestalt particularities, and Hirschsprung disease. In 2005, *KIAA1279* was identified as the disease causing gene with homozygous nonsense mutations in two consanguineous families. The encoded protein, named KBP (KIF-binding protein) due to its related interaction with kinesin-like protein, was initially claimed to be located at the mitochondria, a subcellular localization not confirmed by a recent study. KBP was implicated in axonal structure, outgrowth and maintenance in zebrafish. Since 2005, no supplementary patients with GOSHS have been reported, and data concerning the underling mechanism are poor. In this report, in parallel with the detailed phenotype of five novel patients from three different families, we studied, at the mRNA and at the protein level, the functional consequences of the *KIAA1279* mutations (c.303C>T p.Arg90X, c.599C>A p.Ser200X, and c.604–605delAG p.Arg202IlefsX2) carried by these patients. We proved that a NMD (nonsense mediated decay) mechanism occurred for all three truncating mutations. Indeed, in fibroblasts of patients with homozygous mutations, *KIAA1279* mRNA was barely detectable, and in those with heterozygous mutations, it was two-fold decreased. The addition of a NMD inhibitor in cell culture medium restored partially the mRNA level in fibroblasts with homozygous mutations. We showed that the activities of respiratory chain complexes and the oxydative stress, measured by spectrophotometry, were normal in fibroblasts from all the patients. We studied the subcellular localization of KBP by confocal microscopy, and revealed KBP co-localization with microtubules and filamentary-actin, but not with mitochondria, in human fibroblasts. Microtubule - KBP and filamentary-actin - KBP interactions were confirmed by biochemical binding assays. Finally, by extinction (si-RNA) or over-expression of KBP in a neuroblastoma cells line, we confirmed KBP implication in neurite outgrowth. In conclusion, this study permitted advances in the understanding of GOSHS molecular mechanism: i) by demonstrating NMD mechanism in *KIAA1279* truncating mutations, ii) by showing that KBP absence did not influence mitochondrial respiratory chain complex activity in human fibroblasts, iii) by proving KBP interactions with cytoskeleton, and iv) by confirming *in vitro* the implication of KBP in axonal outgrowth.

3242F

Craniofacial and Dental Development in Cardio-facio-cutaneous Syndrome: The Importance of Ras Signaling Homeostasis. *A. Goodwin[1], S. Oberoi[1], M. Landan[2], C. Charles[2], J. Groth[1], A. Martinez[3], C. Fairley[2], W.E. Tidyman[1], L.A. Weiss[4], O.D. Klein[1,2,3], K.A. Rauen[3].* 1) Center for Craniofacial Anomalies, Department of Orofacial Sciences, UCSF, San Francisco, CA; 2) Program in Craniofacial and Mesenchymal Biology, UCSF, San Francisco, CA; 3) Department of Pediatrics, UCSF, San Francisco, CA; 4) Department of Psychiatry and Institute for Human Genetics, UCSF, San Francisco, CA.

Cardio-facio-cutaneous syndrome (CFC) is a RASopathy that is characterized by craniofacial, dermatologic, gastrointestinal, ocular, cardiac and neurologic anomalies. CFC is caused by activating mutations in BRAF, MEK1, MEK2 or KRAS, all of which are components of the Ras/mitogen-activated protein kinase (MAPK) signaling pathway, which, in turn, is part of receptor tyrosine kinase (RTK) signaling. RTK signaling is known to play a central role in craniofacial and dental development. However, to date, no studies have systematically examined a large cohort of CFC individuals to define key craniofacial features, and lacking are reports of the dental phenotype in CFC. In this study, we examined a large cohort (n=32) of individuals with a clinical and molecular diagnosis of CFC who attended the 2009 and 2011 CFC International Family Conference. The goal was to systematically evaluate the craniofacial and dental phenotype of CFC. We determined that there was a unique phenotype of CFC including malocclusion with open bite, posterior crossbite, teeth crowding and a high-arched palate. Individuals with CFC also presented with oral habits including tongue thrusting and open mouth posture. For the vast majority of CFC individuals, tooth number, size and morphology were normal. This thorough evaluation of the craniofacial and dental phenotype in CFC individuals provides a beginning to understand the role of RTK/MAPK signaling in human craniofacial and dental development and will aid clinicians who care for patients with CFC.

3243W

Sotos Syndrome Diagnosis Confirmed by aCGH. *F.I. Sahin[1], Y.K. Terzi[1], M. Derbent[2].* 1) Department of Medical Genetics, Baskent University Faculty of Medicine, Ankara, Turkey; 2) Department of Pediatrics, Baskent University Faculty of Medicine, Ankara, Turkey.

Sotos syndrome is an overgrowth syndrome caused by mutations or deletions of the NSD1 gene located at chromosome region 5q35.2-q35.3. The condition is characterized by a wide spectrum of clinical abnormalities. Overgrowth is a finding from the prenatal development through childhood as well as long, thin face, broad forehead, sparse frontotemporal hair, downslanting palpebral fissures and malar flushing. Cardiac and renal anomalies are associated features of the syndrome as well as hypotonia, scoliosis, seizures and advanced bone age. Here we report a 2 year 2 months-old girl with mental and motor retardation. The patient was the second child of healthy non-consanguineous parents. At the time of presentation, she was 8 months old and had mildly prominent and high forehead, depressed and broad nasal bridge, anteverted nares, low-set ears, micrognathia, and high-arched palate. Neurological findings were normal except for developmental delay and pronounced hypotonia. Chromosome analysis performed on peripheral blood sample resulted in a normal karyotype (46,XX). In order to explain dysmorphic features, array comperative genomic hybridization (aCGH) method was performed by Roche NimbleGen Human CGH 3×720K Whole-Genome Tiling v3.0 Array. For the analysis of raw data, web based Genoglyphix programme (Signature Genomics) was used. As a result of aCGH analysis, 4 chromosomal regions were marked as important regions. Three of them were copy losses and one was copy gain. However, only one region was found to be clinically significant. One hundred and twelve kb long region was located at 5q35.2-q35.3, and 11 probes marked this region as a copy loss. This region comprises NSD1 gene which has been reported to be related to Sotos syndrome. Molecular genetics methods are useful as additional tools to conventional cytogenetics in cases similar to our patient. Although the clinical findings were insufficient for diagnosis of Sotos syndrome, aCGH results supported and confirmed the diagnosis.

3244F

Expanding (or narrowing) the phenotype of Beckwith-Wiedemann syndrome. *E. Todd, G. Bellus.* Clinical Genetics and Metabolism, Children's Hospital Colorado, Aurora, CO.

This is a case report of a child with a molecular diagnosis of Beckwith-Wiedemann syndrome and an atypical phenotype. Perinatal course: This male patient was appropriate for gestational age, born at term. Upon prenatal ultrasound at 20 weeks gestation, the fetus was found to have a cystic retroperitoneal mass. No other anatomic abnormalities were identified. There were no complications in the perinatal period. Pertinent negatives include no hypoglycemia and no problems with feeding due to macroglossia. Studies: Abdominal computed tomography after birth revealed a low density inhomogeneous mass in the left upper abdomen measuring 4.2 × 4.1 × 4.5 cm with mass effect upon neighboring organs. Pathology following left adrenalectomy demonstrated a hemorrhagic macrocyst and adrenal cytomegaly, concerning for a known association with Beckwith-Wiedemann syndrome (BWS). LIT1/H19 molecular studies identified abnormal allele-specific hypermethylation of H19 in the 11p15.5 chromosomal region, consistent with a diagnosis of BWS. The methylation index detected was above a threshold for abnormal methylation of > 0.65 (2SD greater than the normal mean). No loss of normal imprinting-specific methylation for maternal or paternal alleles of the LIT1 gene was identified. Genetics encounters: At 3 months of age, the patient presented to Genetics clinic. Developmentally he was on track. The patient continued to have normal growth parameters. On examination, there were no physical features consistent with BWS. There was no evidence of neoplasia. Serial screening for tumors was recommended. At 15 months of age, the patient continues to have normal growth parameters and a normal physical examination. Developmentally, he has some mild delays. He is continuing to follow the standard tumor surveillance protocol for children with BWS. Discussion: BWS is a disorder of growth characterized by macrosomia, macroglossia, hemihyperplasia, embryonal tumors, anterior abdominal wall defects, neonatal hypoglycemia, and adrenocortical cytomegaly. It is interesting that while the adrenal cyst which our patient presented with is considered almost diagnostic for BWS, he does not demonstrate any of the other phenotypic manifestations of this disorder. This is, to our knowledge, the first report of a child with isolated adrenal cytomegaly with molecular BWS. His findings probably indicate that the phenotypic spectrum of abnormal imprinting of H19 at 11p15.5 is broader than initially believed.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3245W

Genetic Studies in VACTERL Association. *J. Winberg[1], P. Gustavsson[1], E. Sahlin[1,2], G. Annerén[3], E. Iwarsson[1], P-J. Svensson[4], F. Bradley[1], E. Nordenskjöld[1], N. Papadogiannakis[5], A. Nordgren[1], A. Nordenskjöld[4, 6].* 1) Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden; 2) Department of Clinical Genetics, Karolinska University Hospital, Stockholm, Sweden; 3) Department of Immunology, Genetics and Pathology, Uppsala University, Uppsala, Sweden; 4) Department of Pediatric Surgery, Karolinska University Hospital, Stockholm, Sweden; 5) Department of Laboratory Medicine, Division of Pathology, Karolinska Institutet, Stockholm, Sweden; 6) Department of Woman and Child Health and Center for Molecular Medicine, Karolinska Institutet, Stockholm, Sweden.

In order to identify genetic causes of VACTERL association (V vertebral defects, A anorectal malformations, C cardiac defects, T tracheoesofageal fistula E esophageal atresia, R renal anomalies and L limb deformities), we have collected DNA samples from 20 patients diagnosed with VACTERL or a VACTERL-like phenotype. In addition, we have collected samples from 20 aborted fetuses with VACTERL. To screen for gene dose alterations, array-CGH (comparative genomic hybridization) was performed on genomic DNA from all patients and fetal cases using a 180K oligonucleotide array with whole genome coverage. Patient and fetal DNA samples were also screened for mutations in three candidate genes using DNA sequencing. We present clinical findings in both patients and fetal cases. Results from array-CGH revealed two clinically significant gene dose alterations indicating that pathogenic copy number variants (CNVs) are not a frequent cause of VACTERL. DNA sequencing of candidate genes showed that three fetal cases and one patient carried mutations possibly associated with VACTERL.

3246F

A homozygous AHI1 mutation in a Moroccan consanguineous Joubert Syndrome family. *L. Baala[1, 2], S. Chafai Elalaoui[3], M. Chalon[1], N. Elkhartoufi[4], M. Mansouri[3], Y. Kriouele[5], O. Perche[1, 2], S. Briault[1, 2], T. Attie[4], A. Sefiani[3, 6].* 1) INEM- UMR7355 CNRS-Université Orleans, France; 2) Pole Biopathologie, Centre Hospitalier Régional d'Orléans (CHRO), Orléans, France; 3) Département de génétique médicale, Institut National d'Hygiène, Rabat, Morocco; 4) Département de Génétique et INSERM U781, Hôpital Necker-Enfants Malades, Paris France; 5) Service de pédiatrie IIA, Hopital d'Enfants, Rabat Morocco; 6) Centre de Génomique Humaine, Faculté de Médecine et de Pharmacie, Université Mohamed V Rabat, Morocco.

Joubert syndrome (JBTS, MIM 213300) is a highly genetically heterogeneous disorder, characterized by a developmental delay with hypotonia, ocular-motor apraxia, and breathing abnormalities. The diagnosis is based on brain imaging feature showing “molar tooth sign” (MTS) with vermis agenesis and brainstem accompanying components. Fifteen genes have been identified thus far, 14 of which are autosomal recessive and one X linked-recessive. All encode proteins involved in primary ciliary function including JBTS in the group of so-called ciliopathies. We described a consanguineous Moroccan family (parents are german cousins) with 3 affected sibs (2 boys of 18 and 8 years and one girl of 13 years) showing typical signs of JBTS: hypotonia, developmental delay, mental retardation, bilateral nystagmus, ataxia, retinitis pigmentosa and a cerebelar vermis hypoplosia and molar tooth sign on brain MRI. To identify the underlying genetic defect in this family, we performed a genome-wide homozygosity mapping using the high resolution Affymetrix SNP 6.0 array containing 906,600 polymorphic SNPs and more that 946,000 unique probes for CNV detection. We found that only three homozygous regions were shared between affected siblings, with a size of 1.7-Mb (at 5q33.3), 30.1 Mb (at 14q22-24) and 28.9 Mb (at 6q22.1-24) respectively. Overall, the three intervals contained 100 to 300 genes. We compared the detected homozygous regions with known JS loci and found that the homozygous region on chromosome 6q contains AHI1 considered as the gene at JBST3 locus. Direct sequencing of the 26 coding exons revealed a novel homozygous mutation (c.910dup, p.Thr304AsnfsX6) located in exon 6 (NM_017651.4) in all patients. This mutation was heterozygous in their parents and not found in controls in the same ethnic background. Otherwise, no pathological CNV (deletion or duplication) was detected. Interestingly, the c.910dup mutation was previously reported in a JS family originating from Spain suggesting a possible founder effect. Our study highlights that homozygosity mapping is a powerful method to rapidly detect the disease causing gene especially in genetically heterogeneous disorders such as JBTS. This result allows an accurate genetic counseling and molecular prenatal diagnosis in this family.

3247W

Twin Discordance and Congenital Hydrocephalus. *C. Lee[1], J.T.C. Shieh[1,2].* 1) Division of Medical Genetics, Department of Pediatrics, University of California San Francisco, San Francisco, CA; 2) Institute for Human Genetics, University of California San Francisco, San Francisco, CA.

Abnormalities that arise during human development such as congenital hydrocephalus may have an unclear etiology, however twin studies may help identify factors that lead to disease. Hydrocephalus is defined by dilated cerebral ventricles, and these may arise due to abnormal cerebral spinal fluid flow, obstruction or impaired fluid balance. The association of congenital hydrocephalus with VACTERL, genomic microarray abnormalities, and single gene alterations suggest that hydrocephalus may arise in the context of genetic alterations. Whether genetic factors underlie apparently isolated hydrocephalus is unclear. Here we present data on identical twin phenotypic discordance with hydrocephalus. We present clinical course, MRI findings and candidate gene sequencing as these may help us understand factors that contribute to disease. We examined a case of hydrocephalus in one twin of a discordant twin pair who was found to have ventriculomegaly with dilatation of bilateral lateral ventricles and third ventricle by fetal MRI at 23 weeks gestation. The co-twin was normal. The male twins were born prematurely at 33 weeks. Head size was large in the affected twin, and neurological examination at birth was unremarkable but mild tone abnormalities were noted in subsequent exams. Postnatal MRI revealed non-communicating hydrocephalus and aqueduct of Sylvius obstruction in the affected twin, and a shunt procedure was performed. The twins appeared non-syndromic, and although a family history of hydrocephalus was absent, L1CAM testing was discussed and test results are pending given its potential impact on prognosis. We also compare this case to other twin studies of hydrocephalus. We propose that further twin studies may help us understand heritability and the factors that lead to this disease.

3248F

Hallermann-Streiff Syndrome: Case report from Mongolia. *P. Erkhembulgan[1], M. Purevdorj[2], T. Altansukh[2], I. Purevdorj[2].* 1) National Center for Maternal and Child Health, Ulaanbaatar, Mongolia; 2) Health Sciences University, Ulaanbaatar, Mongolia.

Hallermann-Streiff Syndrome is a rare genetic disorder that is characterized primarily by head and face abnormalities. Hallermann-Streiff Syndrome was first described completely in 1948 by Hallermann, in 1950 by Streiff, and in 1958 by Francois. We report a Hallermann-Streiff Syndrome case from Mongolia. Proband B, 17-years-old female was admitted to the Center of Genetic counseling, Health Sciences University of Mongolia in November 2011. Physical examination revealed alopecia, beak-shaped nose, bilateral microphthalmia, hypognatia, brachycephaly, bird-like face, lack of hair in frontal and parietal parts of the head, congenital cataract, microstomia, high arched palate, abnormal dentition and natal teeth, severe caries, absence of teeth in lower jaw. Despite retarded growth, the mental development of the patient was normal. Examination of eyes revealed that the child had impaired vision (VOD=0.04, VOS=0.03), sparse eyebrows and eyelashes. The patient underwent lensectomy at 8 and 14 months. Seven essential signs (bird like face, abnormal dentition, hypotrichosis, atrophy of skin especially on the nose, congenital cataract, microphtalimia, proportional dwarfism) were described by Francois as diagnostic criteria for Hallermann-Streiff Syndrome. Our patient had six of these seven signs. Based on our findings the patient was diagnosed to have Hallermann-Streiff Syndrome. Sleep apnea, one of the life threatening complications among patient with Hallermann-Streiff syndrome was detected in our patient. Recently, she underwent nasal cavity surgery due to respiratory embarrassment. An interdisciplinary approach including early preventive-care programs, detailed oral hygiene instructions, dietary recommendations, counseling of the parents and regular dental visits are the essential procedures for the Hallermann-Streiff Syndrome.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3249W

Global Gene Profiling Identifies A Novel Cytoskeleton Pathway Involved in VCP-Associated Myopathy. *A. Nalbandian[1], S. Ghimbovschi[2], S. Radom-Aizik[1], E. Hoffman[2], V. Kimonis[1].* 1) University of California, 2501 Hewitt Hall, Irvine, CA; 2) Children's National Medical Center, 111 Michigan Avenue, NW, Washington, DC.

Inclusion body myopathy associated with Paget's disease of bone and frontotemporal dementia (IBMPFD) is an increasingly recognized disorder caused by Valosin containing protein (VCP) mutations and is now identified in 2% of familial amyotrophic lateral sclerosis (fALS). VCP disease is characterized clinically by adult onset of progressive proximal muscle weakness, Paget's disease of bone, frontotemporal dementia, and clinical evidence of motor neuron disease in approximately 10% of individuals. Patients may express only one or two phenotypic components and need not demonstrate all three phenotypic features. Pathology includes ubiquitin and TDP-43-positive inclusions also seen in ALS and other proteinopathies. VCP is a key regulatory protein and is involved in a plethora of cellular functions and signaling transduction pathways. Current theories suggest VCP disease mutations may cause disturbances in common signaling pathways, including autophagy, lysosomal, and ubiquitin-proteasomal pathways. We have conducted a microarray profiling study examining the muscles from patients with VCP mutations versus their first degree unaffected and non-mutation carrying relatives to identify the key disrupted pathways in VCP disease. With this global analysis, we have elucidated several dysregulated transduction cascades. Interestingly, the microarray results showed most significant changes in the regulation of actin cytoskeleton pathway including genes critically involved in cell communication, multiple structural and dynamic processes including maintenance of cell shape, cell movement, cell signaling, cell differentiation, and apoptosis. Downregulated genes in this pathway included alpha actinin 1, actin related protein 2/3 complex, myosin light chain 10, and fibroblast growth factor receptor 2, directly linked to cytoskeletal integrity. Upregulated genes in the actin cytoskeleton included protein tyrosine kinase 2, integin beta 1, protein-activated kinase 1, and protein phosphatase 1. These proteins are expressed in muscle and may be involved in attaching muscle tissue to the extracellular matrices. Thus, a dysregulation of these proteins may lead to aberrations in the cytoskeleton, thereby leading to myopathy. Dissecting the signaling intermediates of the actin cytoskeleton cascade will help in understanding the pathogenesis of VCP-associated disease and for the development of potential therapeutic targets in related neurodegenerative diseases.

3250F

External validation of the Bartholdi clinical scoring system for Silver-Russell syndrome: a report of a Brazilian cohort. *G.L. Yamamoto[1], A. Bonaldi[2], A.M. Vianna-Morgante[2], I. Gomy[1], C.A. Kim[1], D.R. Bertola[1].* 1) Genetica Medica ICR-FMUSP, Universidade de Sao Paulo, Sao Paulo, Sao Paulo, Brazil; 2) Instituto de Biologia, Universidade de Sao Paulo, Sao Paulo, Sao Paulo, Brazil.

Silver-Russell syndrome (SRS) is a genetic heterogeneous disorder with a variable phenotype characterized by pre and post-natal growth deficiency, typical facial features, body asymmetry and other malformations. SRS is caused by imprinting disorders in at least two different regions: chromosome 7 and the 11p15.5 locus. The clinical criteria for diagnosis and recommendations for molecular testing have been widely discussed in the literature without reaching a consensus. In this study, we report on 27 suspected SRS patients using molecular testing, and 18 of these patients (67%) were positive, with 14 patients (78%) possessing hypomethylation of Imprinting Center Region 1 (ICR1), one patient (6%) with a duplication of ICR2 and 3 patients (17%) with maternal uniparental disomy of chromosome 7. The clinical features and genotype-phenotype correlations were in accordance with the literature for the asymmetry and clinodactyly of fifth finger associated with ICR1 hypomethylation. Four different clinical criteria were applied, specifically, Bartholdi, Netchine, Price and Saal, to compare their efficacy in this population. The Bartholdi criteria performed better than the other tests, with 100% specificity and sensibility being observed for all positive and negative molecular SRS patients.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3251W

Copy number differences of putative regulatory elements shape primate expression profiles. *R. Iskow[1,2,3], O. Gokcumen[1,2], A. Abyzov[4,5], J. Malukiewicz[6], Q. Zhu[1,2], A.T. Sukumar[7], A.A. Pai[8], R.E. Mills[9], L. Habegger[4,5], D.A. Cusanovich[8], M.A. Rubel[10], G.H. Perry[11], M. Gerstein[4,5,12,13], A.C. Stone[14], Y. Gilad[8], C. Lee[1,2].* 1) Department of Pathology, Brigham & Women's Hospital, 221 Longwood Avenue Boston, MA 02115, USA; 2) Harvard Medical School, 20 Shattuck Street, Boston, MA 02115 USA; 3) MMSc in Biomedical Informatics, Harvard Medical School. 10 Shattuck Street, Boston, MA 02115 USA; 4) Program in Computational Biology and Bioinformatics, Yale University, P.O. Box 208009, New Haven, Connecticut 06520, USA; 5) Department of Molecular Biophysics and Biochemistry, Yale University, 260 Whitney Avenue P.O. Box 208114, New Haven, Connecticut 06520, USA; 6) School of Life Sciences, Arizona State University, 427 East Tyler Mall, Tempe, AZ 85287, USA; 7) Currently at Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, Texas 77030, USA; 8) Department of Human Genetics, The University of Chicago, 920 E. 58th Street, CLSC 5th floor, Chicago, IL 60637, USA; 9) Department of Computational Medicine and Bioinformatics, Department of Human Genetics, University of Michigan, 2017 Palmer Commons Bldg. 100 Washtenaw Avenue, Ann Arbor, MI 48109, USA; 10) Department of Anthropology, University of Pennsylvania, 3260 South Street Philadelphia, Rm. 325, PA 19104-6398 USA; 11) Department of Anthropology, Penn State, 409 Carpenter Building University Park, PA 16802 USA; 12) Department of Chemistry, Yale University, 225 Prospect Street PO Box 208107, New Haven, Connecticut 06520, USA; 13) Department of Computer Science, Yale University, P.O. Box 208285, New Haven, Connecticut 06520, USA; 14) School of Human Evolution and Social Change, Arizona State University, P.O. Box 872402, Tempe, AZ 85287, USA.

Gene expression differences are shaped by selective pressures and contribute to phenotypic differences between species. We identified 964 copy number differences (CNDs) of DNA segments that are highly conserved at the sequence level across human, chimpanzee, and rhesus macaque and examined their potential effects on gene expression profiles. Samples with copy number different genes had significantly different expression levels than samples with neutral copy number for those genes. In particular, genes encoding regulatory molecules, such as transcription factors, were enriched both for overlap with copy number differences and for significant expression level differences. We also identified 127 CNDs which were actually processed pseudogenes and some of which were found to be expressed. Furthermore, copy number different noncoding regulatory regions, such as ultraconserved elements and long intergenic noncoding RNAs with the potential to affect expression, were also identified. We postulate that CNDs of these conserved sequences fine-tune developmental pathways by altering the levels of RNA. Many of these conserved, copy number different, and differentially expressed regions are implicated in human disease.

3252T

Linkage disequilibrium, population structuring and genome-wide association studies reveal metabolism and craniofacial dysmorphology mutations in the domestic cat. *L.A. Lyons, H. Alhaddad, B. Gandolfi, R.A. Grahn.* Population Hlth Reprod/Vet Med, Univ California, Davis, Davis, CA.

Domestic cats are increasingly attracting the scientific community as a model organism for human hereditary and infectious diseases as well as attracting the general public as a legitimate household pet. There are more cats in the USA than people and approximately 30% of households own a cat. Our research efforts have focused on the genetic structuring of cat populations, indicating that SNPs and STRs can define eight worldwide “races” of cats and that the most common 29 breeds form only ~ 21 cat breed populations. The feline scientific community used this population data and additional information to support the sequencing efforts for the appropriate selection of cats for efficient SNP discovery and the development of an Illumina Infinium 63K Feline iSelect DNA array. The DNA array also contains SNPs that define phenotypic and known disease traits in the cat. Linkage disequilibrium (LD) studies of 1536 SNPs across ten 1 Mb regions have shown that some breeds such as Burmese have fairly extended LD, comparable to horses and some dog breeds, but the LD of random bred cats is more comparable to human populations. The initial testing of the DNA array considered 288 cats, including 10–20 cats per breed population, five trios and two random bred populations. The cats were selected to test the array's efficiency by remapping several known traits - with and without the causative mutations present. Several additional case-control projects, investigating hypokalemia in Burmese and coat traits in other breeds have been been successful, regionally localizing the traits with subsequent causative mutation identification with as few as 11 cases. Presented will be the highlights of the cat population and LD studies, initial studies of the array development, showing the power of trait mapping, and the subsequent successful case-control studies for feline hypokalemia and a craniofacial defect that also causes brachycephaly.

3253F

Genome-wide scan for identifying selection footprints in immune related pathways. *J. Molineros, S. Nath.* OMRF, 825 NE 13th, Oklahoma City, OK. 73104.

Immune-related genes have been previously identified in scans for positive selection. Genome-wide scans of selection are important to autoimmunity because they might be able to identify loci carrying deleterious mutations linked to positively selected loci. The effects of selection would be expected to be detected on the pathways as well. Our objective was to identify autoimmune genes that are under natural selection and their associated pathways. We used 1000 genomes project genotype data for the Illumina Omni2.5 SNP platform for European (CEPH, Finnish, British, Tuscan, Iberian) and African (Yoruba, Luhya and Massai) populations to estimate the newly reported spatial statistic (SPA), as well as Fst and iHS. Loci with evidence of selection were analyzed using gene set enrichment analysis to determine overrepresented pathways as well as overrepresented protein domains through their gene ontology. Adjusted p-values were obtained using the false discovery rate statistic. We identified selection signals on 2888 loci. Only 38 KEGG categories pathways passed the multiple testing correction threshold value FDR (P<0.05). Whereas 535 categories of gene ontology were significantly overrepresented, among those, metal-ion binding domains and extracellular structure organization domains were particularly enriched (P = 0.04, for both), as well as Zinc ion binding domains (P = 0.02). Of particular interest were loci *KCNH7*, *XKR6*, *CANCB4*, *STAT1* and *STAT4* that have been implicated in autoimmunity. Our results suggest that single genes rather than immune pathways were under selection.

3254W

Signatures of selection surrounding large insertions and deletions in coding regions identified between modern and Denisovan humans, as well as four other species of primates genome-wide. *W. Guiblet[1], K. Zhao[3], D. Ferrer-Torres[1], C.T. Ruiz-Rodriguez[1], A. Roca[3], S. Massey[2], J.C. Martinez-Cruzado[1], T. Oleksyk[1].* 1) Department of Biology, University of Puerto Rico at Mayaguez, Mayaguez, PR. 00680; 2) Department of Biology, University of Puerto Rico at Rio Piedras, San Juan, PR 00931; 3) Department of Animal Sciences, University of Illinois at Urbana-Champaign, Urbana, IL 61801.

Genes have highly conserved sequences and usually show very few differences between closely related species, such as humans and nonhuman primates. In this study, we focused on >10 bp insertions and deletions (indels) found when comparing modern humans, Denisovans, chimpanzee, gorilla, orangutan, and rhesus macaque reference genome sequences, with the purpose of testing indel flanking regions for the signatures of selection. From 36,422 indels identified by comparing reference genomes pairwise, we chose 186 indels within coding regions because of the potentially high impact on protein sequence. These fragments had been earlier validated in the laboratory by PCR and electrophoresis to distinguish real features from the computational artifacts, and further tested for the presence of polymorphisms in multiple human populations. We used five different algorithms (PAML package) to produce over 7,000 pairwise Ka/Ks ratios from 98 filtered genes. Among the 98 genes containing indels, ten showed significant Ka/Ks ratios between human and at least one primate sequence. We also search for and identify indels within candidate chromosomal regions for positive selection displaying unusually low multilocus heterozygosity and high divergence (FST) in pairwise comparisons between populations or continental groups from the the Human Genome Data Panel (HGDP). The comparisons include all the possible pairwise combination for the continental groups as well as populations within groups for all chromosomes. Our findings are evaluated against the random expectation by a resampling method, where exactly the same procedures and tests are performed with a dataset of randomly positioned indels matched by size distributed across the human reference genome. Those indels still polymorphic in humans for the alleles discovered in cross-species comparisons (ex: USP7), are further interrogated for recent selection signatures by evaluating extended haplotype homozygosity around ancestral and derived alleles. The genes examined in our study may have been shaped by selection in the human or other primate lineages, thus adding to the understanding of our recent evolution. Some of these may harbour adaptation to disease, and benefit discoveries in the future biomedical studies.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3255T

An optimized panel for Ancestry Informative Markers derived from the Mexican Genome Diversity Projec. *JC. Fernandez-Lopez[1], A.V. Contreras[1], MJ. Gomez-Vazquez[2], C. Rangel-Escareño[1], I. Silva-Zolleżzi[1,3].* 1) Computational Genomics Department, Instituto Nacional de Medicina Genómica, Ciudad de Mexico, Distrito Federal, México; 2) Universidad Autónoma de Nuevo León; Monterrey N.L. México; 3) BioAnalytical Science Department Nestle Ltd, Nestlé Research Center Lausanne, Switzerland.

Genome wide association studies (GWAS) were not common in Mexican and Latino populations until recently. These studies have contributed to the search for genetic variants underlying the inheritance and development of complex traits and diseases within Mexican mestizos. The process of identifying functional variants can be long and difficult, one of these difficulties is the huge number of false positive results observed in these studies due to population stratification between cases and controls groups. In recent years, a variety of methods that aim to minimize the effect of population stratification in Mexican and Latinos have been developed. One of these tools is the development of Ancestry Informative Marker (AIM) panels. Nevertheless, it remains extremely important to generate new additional panels to address the admixed populations using different feature-discriminators and approaches. With this work we decided to release the top 260 AIMs panel based on a list of 1,814 unpublished AIMs selection used in Silva-Zolezzi et al. (2009). That list of AIMs was created using unlinked SNPs based on allele frequency differences with (δ) $\geq$ 0.4. This AIMs selection captures East Asian structure in addition to European, African and Native-American structure not included in another Latino panels. The top 260 AIMs, was selected by taking all pairwise comparisons to discriminate between populations. Caucasian vs East-Asian (40), Caucasian vs African (40), East-Asian vs African (40), African vs Native-American (40), Caucasian vs Native-American (50) and East Asian vs Native-American (50). To validate this top 260 AIMs panel we genotyped 400 Mexican-mestizos from Age-related macular degeneration GWAS. The performance of this subset of markers captured the ancestry proportions as accurate as that using all genomewide data. Next, we tested using candidate gene approach 200 Chronic Obstructive Pulmonary Disease study in Mexican mestizos, the preliminary results were consistent with the ancestry expected in this selection group.

3256F

A fish-specific transposable element shapes the repertoire of p53 target genes in zebrafish. *M.N. Loviglio[1,2], L. Micale[2], M. Manzoni[3], C. Fusco[2], B. Augello[2], E. Migliavacca[1], G. Cotugno[2], E. Monti[3], G. Borsani[3], A. Reymond[1], G. Merla[2].* 1) Center for Integrative Genomics (CIG), University of Lausanne, Lausanne, Canton Vaud, Switzerland; 2) Medical Genetics Unit, IRCCS “Casa Sollievo della Sofferenza”, San Giovanni Rotondo, Foggia, Italy; 3) Department of Biomedical Science and Biotechnology, University of Brescia, Brescia, Italy.

Transposable elements (TEs) represent the largest genomic component of most eukaryotic organisms. Often erroneously referred as “junk DNA”, they actually play a central role in the structural organization and plasticity of genomes and are known to mediate the establishment of new cellular functions during evolution. Recent discoveries support the hypothesis that gene regulation can be influenced by TEs, which can directly mediate novelties at the network level by contributing to transcriptional and post-transcriptional modulation due to their proximity to nearby genes. For instance, TEs partecipate to the origin of new functional elements, as promoter sequences, transcription factors binding sites and enhancer, silencer and insulator elements. For example, binding sites of the pleiotropic master transcription factor p53 reside in LINE1, Alu and LTR repeats in the human genome. Similarly, here we demonstrate that regulatory information controlling expression of a repertoire of p53 target genes in zebrafish (Danio rerio) is coming from a mobile element, the non-autonomous DNA transposon EnSpm-N6_DR. Through their embedded p53 responsive elements, the multiple copies of the EnSpmN6_DR transposon drive in several instances p53-dependent transcriptional modulation of the adjacent genes, whose human orthologs were often previously annotated as p53 targets. The domestication of EnSpm-N6_DR repeats and their embedded p53REs is likely associated to the involvement of the targeted genes in shared pathways; in fact, these transposons define a set of target genes that contribute to axonogenesis, synaptic transmission and the regulation of programmed cell death. Consistent with these biological functions, the expression data of EnSpmN6_DR-colonized loci notably show an enrichment for genes expressed in both adult and fetal brain, as well as specific brain structures, such as amygdala, hippocampus, and brain cortex. Accumulating evidences demonstrate that p53, besides its well studied features as master regulator in cell-cycle and apoptosis, plays a leading role in the regulation of neural stem cell proliferation and differentiation, as well as axon outgrowth and regeneration. In this scenario, our data corroborate these recent findings and pinpoint a remarkable example of convergent evolution: the exaptation of lineage-specific transposons to establish networks of p53-regulated genes crucially involved in neuronal morphogenesis in both a hominid and a teleost fish.

3257W

Meiotic recombination hotspots in humans - dynamics and controlling factors. *L. Odenthal-Hesse, A.J. Jeffreys.* Genetics, University of Leicester, Leicester, Leicestershire, United Kingdom.

Meiotic gene conversion has a major impact on genome diversity. Both crossovers and non-exchange conversions cluster into distinct recombination-active regions that we call hotspots. Hotspot analysis in humans has focused on the description of crossover profiles and only few hotspots had been tested for crossover and non-exchange gene conversion. Therefore, very little was known about the frequency and distribution of non-exchange conversions and how well they correlate with crossing over. Six extremely active recombination hotspots were analysed by using small pool PCR approaches on sperm DNA to detect both types of recombinant molecules. Non-exchange conversions were detectable and arose at high frequencies (0.01–0.47%) per sperm, in addition to crossovers (0.2–0.70%). Conversion tracts were short and their distribution defined a steep conversion gradient, centred at the peak of crossover activity and probably marking the zone of recombination initiation. It was also observed that non-exchange gene conversion and crossover frequencies were positively correlated, not just between men at the same hotspot but equally when compared across several hotspots. On average, non-exchange gene conversions spanning a marker close to the centre of the hotspot occurred at 50% of the crossover frequency. Hotspot regulation in cis had been described previously to results in different initiation efficiencies between interacting haplotypes. Preferential initiation on a more active haplotype, in turn leads to the overtransmission of alleles from the less active haplotype. Additional hotspots that showed a phenomenon of biased gene conversion were described in this study, with crossovers and non-exchange gene conversions influenced to the same degree. More unusually, biased gene conversion specifically affecting non-exchange events was also observed at two hotspots. Here single SNP heterozygosities appear responsible for triggering the bias in cis. Crossovers were not affected, which may provide evidence for distinct crossover and non-crossover pathways operating at human hotspots.

3258T

The MHC linkage haplotype in the Finnish Population. *A. Wennerström, ML. Lokki, MHC Disease Research Group.* HLA Laboratory, Transplantation Laboratory, Haartman Institute, Helsinki, University of Helsinki, Finland.

Previously, the Major Histocompatibility Complex (MHC) has been associated with a variety of infections (e.g. HSV, Pogosta, otitis) and immunologically mediated diseases (e.g. sarcoidosis). The genes in the MHC region (6p21) encoding HLA molecules fall into two major classes, Class I (HLA-A, -B, and -C) and Class II (HLA-DP, -DQ, and -DR). The MHC Class III genes containing, e.g., complement C4, TNF-α, LTA and BTNL2 reside between the HLA Class I and Class II genes. Our aim was to explore the MHC linkage haplotypes in the Finnish population and utilize the information in disease association studies. The HLA class I (HLA-A and -B) and Class II (HLA-DRB1, -DQB1 and -DPB1) allele and haplotype frequencies of a Finnish population (n=150) were calculated. In addition, C4 gene copy numbers (CNV) and 89 MHC Class III tag-SNPs were genotyped. Linkage disequilibrium (LD) and Hardy-Weinberg distributions were analyzed and haplotypes were constructed with ARLEQUIN 3.11, GenePop 4.0.10 and PHASE v.2.1.1, respectively. The most common Finnish A-B-DR -haplotype (7 %) was not-so-usual-Caucasoid A*03-B*35-DRB1*01:01, which possibly reflects the isolated nature of the Finnish population. The haplotype has been previously suggested to be protective for diabetes but predispose to coronary artery disease. The most frequent Caucasoid haplotype AH8.1 including A*01-B*08-DRB1*03:01, was present at frequency 4%. As expected, AH8.1 haplotype had strong pairwise LD (D'> 0.8) covering Class I, Class II and Class III genes. However, when studying other haplotypes including Class III markers and HLA-DPB1, we noticed that the pairwise LD between alleles and loci varied suggesting conserved regions and potential recombination hotspots. Further, this data clearly shows the importance to study several MHC markers covering a long chromosomal region to distinguish different predisposing alleles and haplotypes. This data provides useful information for disease association studies and here, preliminary results from our Sarcoidosis project will be presented.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3259F

Fine-scale mapping of meiotic recombination in Mongolians and Koreans. *T. Bleazard[1,2], Y.S. Ju[1], J. Sung[1,3], J.S. Seo[1,4].* 1) Genomic Medicine Institute (GMI), Medical Research Center, Seoul National University, Seoul, Korea; 2) College of Natural Sciences, Seoul National University, Seoul, Korea; 3) Department of Epidemiology, School of Public Health and Institute of Environment and Health, Seoul National University, Seoul, Korea; 4) Department of Biochemistry and Molecular Biology, Seoul National University College of Medicine, Seoul, Korea.

Meiotic recombination causes a shuffling of homologous chromosomes as they are passed from parents to children. Finding the genomic locations where these crossovers occur is important for genetic association studies, understanding population genetic variation, and predicting disease-causing structural rearrangements. There have been several reports that recombination hotspot usage differs between human populations. But while fine-scale genetic maps exist for European and African populations, none have been constructed for Asians. We present the first Asian genetic map with resolution high enough to reveal hotspot usage. We constructed this map by applying a hidden Markov model to genotype data for over 500,000 single nucleotide polymorphism markers from Korean and Mongolian pedigrees which include 980 meioses. For direct comparison of genetic maps between ethnic groups, we also constructed a map for CEPH families using identical methods. We used these data to investigate hotspot usage and the hypothesized contribution of recombination problems to age-related aneuploidy. We found high levels of concordance with known hotspots, with approximately 72% of recombination occurring in these regions. Importantly, this shows that the proportion of human recombination which occurs in hotspots is considerably greater than previously though. Our large sample size allowed us to detect a weak but significant negative effect of maternal age on recombination rate. Our map fills an important gap in the understanding of recombination pattern variation and will be a valuable resource for future research in population genetics. This will improve the accuracy of linkage studies and inform the design of genome-wide association studies in the Asian population.

3260W

Heterogeneity in recombination among African populations. *M. Capredon[1], J. Hussin[1, 2], J. Quinlan[1, 3], Y. Idaghdour[1], L. Barreiro[1], T. de Malliard[1], JC. Grenier[1], E. Gbeha[1], P. Awadalla[1].* 1) Pediatrics, CHU Sainte Justine, Faculty of medicine, University of Montreal, Montreal, Quebec, Canada; 2) Biochemistry, Faculty of medicine, University of Montreal, Montreal, Quebec, Canada; 3) School of public health, Faculty of medicine, University of Montreal, Montreal, Quebec, Canada.

Heterogeneity in recombination can have substantial impacts on the shape of genetic variation within and among populations and the signature for selection in human genomes. For genome-wide association approaches to be effective and replicated among different ethnic groups or populations, heterogeneity in recombination rates needs to be factored. Here we present both a method and results for recombination rate variation across the genome for different African populations. Analyses were performed on genome-wide data for three populations from Benin (n=300), two cohorts from Central Africa (n=80, including a Pygmies sub-cohort) and also African Hapmap phase 3 population data and data from the HGDP. An analysis of the SNPs generated a set of common SNPs that were analyzed across populations. Recombination rates were inferred using both phased and unphased data. An IBD index using Plink was computed in each population in order to exclude most of the related individuals. Finally, recombination events were calculated with Recmin, and rates were inferred for data with a subset of common positions in all populations. Recombination rates and hotspots were inferred using the interval option of LDhat 2.1 and LDhot softwares. A CEU population from the Hapmap project was also used as control outgroup for our comparison of hotspots. Population structure analyses allowed us to make comparisons for recombination rates among populations that are more distantly related as well as among very closely related populations, both geographically and genetically. For example, with PCA we visualized a western Africa cluster. The first axis separated central and southern African populations and the second axis the eastern African populations and highlight a Bantu sub-cluster. In order to compare results, the recombination rates were normalized and hotspots predicted with LDhot were then compared between all the populations by their presence or absence. Naturally, we observed that mean map lengths were generally consistent across all populations. However, we were able to capture rapid evolution in hotspots even among populations that are geographically close. Although most events are shared, subtle changes can be detected, even after correcting for differences in sample size.

3261T

Linkage disequilibrium patterns in a Brazilian population: a comparison of X chromosome STR markers with autosomal and Y chromosome STR markers. *S. Oliveira[1], C. Mendes-Júnior[2], A. Trindade-Filho[3].* 1) Genetica e Morfologia, Universidade de Brasilia, Brasilia, Distrito Federal, Brazil; 2) Departamento de Química, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, São Paulo, Brazil; 3) Laboratório de Antropologia Forense, Instituto de Medicina Legal, Polícia Civil do Distrito Federal, Brasília, Distrito Federal, Brazil.

The X chromosome has a peculiar mode of inheritance: it is transmitted to males exclusively from their mothers, whereas it is transmitted to females by both parents. Such eccentricity makes the genetic markers located in this chromosome particularly useful in population genetics studies, especially to access the extension and distribution of Linkage Disequilibrium (LD). LD may be reduced in comparison to the estimates from the Y chromosome as X chromosomes recombine in females, but higher in comparison to autosomal chromosome. LD arises from genetic linkage, population stratification and natural selection. Here, we compare LD patterns estimated using autosomal, X-linked and Y-linked STR data in a Brazilian population sample. Ten STRs located on the human X chromosome (DXS8378, DXS9898, DXS7133, GATA31E08, GATA172D05, DXS7423, DXS6809, DXS7132, DXS9902, and DXS6789); 10 autosomal STRs (CSF1PO, D3S1358, D7S820, D8S1179, D13S317, D16S539, D21S11, TH01, TPOX, and vWA); and 10 Y-linked STRs (DYS456, DYS390, DYS389II, DYS458, DYS19, DYS439, DYS635, DYS392, DYS438, and DYS448) were analyzed in a sample of unrelated people from Brasilia (Brazilian Federal District), comprising 204 chromosomes (204 men for X and Y chromosomes; and 102 men for autosomal chromosomes). For each of these three aforementioned sample sets, there are 45 possible pairs of loci for which the LD Likelihood Ratio Test was calculated using the Arlequin 3.11 software. LD analysis revealed three pairs of autosomal loci, five pairs of X-chromosome loci and 35 pairs of Y chromosome loci with significant LD ($p \leq 0.05$). The 45 LD probability values revealed mean p-values of 0.5504, 0.4614 and 0.0420 for autosomal, X-linked and Y-linked STR pairs, respectively. The observed LD pattern for the autosomal loci reflects both random segregation and small population stratification, while the p-values among the Y-chromosome loci reflect physical attachment and lack of recombination among the loci. On the other hand, the X-chromosome STRs behave in an almost independent manner (such as the autosomal STRs), with slightly smaller LD probabilities. In conclusion, the X-chromosome STRs show an LD pattern that resembles the behavior of autosomal STRs and completely differs from that observed for Y chromosomes, whereby STRs are inherited as a cluster. Our results could assist current debates regarding the estimation of X-chromosome STR profile frequency in forensic casework. Acknowledgments: CAPES, CNPq, FAPDF.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3262F

Association of (ccttt)n polymorphism in the nos2a gene with tuberculosis in two population groups of india. *M. Jena[1,2], P. Das[2], R. Bamezai[1].* 1) National center applied human genetics,school of life sciences,JNU,New Delhi,; 2) Anthropology, Utkal University, Bhubaneswar,Odisha,India, India,.

NOS2 is a molecule that plays a key role in the immunological control of a broad spectrum of infectious agents.Investigation is hampered by difficulty in estimating in vivo production of nitric oxide(NO),but genetic studies provide a potential means of examining the relation between NO production and disease outcome.In this study we assess the influence of CCTTT pentanucleotide microsatellite repeat polymorphism in the NOS2A gene on the risk of developing pulmonary TB in two population(Odisha in eastern india and Chhattisgarh tribe in central india) groups of India .This study included a total of 658 individuals(186 TB patients and 254 controls from odisha and 78 TB patients and 140 controls from Chhattisgarh) .The polymorphism was analyzed by PCR-Electrophoretic method and by direct sequencing. Statistical analysis was done by using Graph pad instat v3.00 for windows (Graphpad software,San Diego CA,USA). Analysis revealed significant difference between patients and controls with NOS2A CCTTT polymorphism in both groups. The overall CCTTTn distribution showed statistically significant differences between TB patients and healthy controls in odisha(p=0.006) and chhatisgarh(p=0.007).We analyzed the possible contribution of the 10-repeat allele (186 base pairs) to TB predisposition in both the population .Decreased allele frequency of (CCTTT)10 in cases as compared to controls resulted in a significant association indicating protection against tuberculosis in odisha (p=0.0002, OR=0.47, 95%CI=0.31-0.70) and chhatisgarh(p= 0.0005, OR=0.37, 95%CI=0.20–0.66). This was confirmed when NOS2A alleles were stratified into short (CCTTT)7–11 and long (CCTTT)12–17 repeats, and a significant difference observed again with short repeats between patients and controls of odisha (p=0.04, OR=0.7, 95%CI=0.5–0.9) and chhatisgarh(p=0.01,OR=0.59, 95%CI=0.39–0.88). Individuals carrying long repeats (CCTTT)12–16 provided a very low risk towards TB in odisha population (p=0.04, OR=1.33,95%CI=1.01–1.74),but shows a significant difference in chhatisgarh population(p=0.01,OR=1.68,95%CI=1.13–2.50).Again combined analysis of short and long repeats in both the groups shows long repeat carriers provided risk(p=0.002, OR=1.43, 95%CI=1.14–1.79)towards TB. In conclusion, further replication studies in different populations are necessary to fully establish the role of this iNOS CCTTT'npolymorphism in TB and its relationship with other genes implicated in tuberculosis susceptibility.

3263W

Purifying selection in mitochondrial protein-coding genes is highly effective in mammals and congruent with evolution of nuclear genes. *S.I. Nikolaev[1], K. Popadin[1], T. Junier[1], M. Baranova[2], S.E. Antonarakis[1].* 1) Department of Genetic Medicine and Development, University of Geneva Medical School; 1 Rue Michel Servet, 1211 Geneva, Switzerland; 2) Faculty of Bioengineering and Bioinformatics, M.V. Lomonosov Moscow State University, Leninskiye Gory, GSP-2, building 73, Moscow, 119992, Russia.

The mammalian mitochondrial genomes differ from the nuclear genomes by maternal inheritance, absence of recombination and higher mutation rate. All of these differences decrease the effective population size of mitochondrial genome and make it more susceptible to accumulation of slightly-deleterious mutations. It was hypothesized that mitochondrial genes, especially in species with low effective population size, irreversibly degrade leading to decrease of organismal fitness and even to extinction of species through the mutational meltdown. To interrogate this hypothesis we compared the purifying selections acting on the representative set of mitochondrial (potentially degrading) and nuclear (potentially not degrading) protein-coding genes in species with high and low effective population size. For 21 mammalian species we calculated the ratios of accumulation of slightly-deleterious mutations approximated by Kn/Ks separately for mitochondrial and nuclear genomes. 75% of variation in Kn/Ks is explained by two independent variables: type of a genome (mitochondrial or nuclear) and effective population size of species approximated by generation time. Firstly, we observed that purifying selection is more effective in mitochondria than in the nucleus that implies strong evolutionary constraints of mitochondrial genome. Mitochondrial de novo nonsynonymous mutations have at least 5 fold more harmful effect as compared to nuclear. Secondly, Kn/Ks of mitochondrial and nuclear genomes are positively correlated with generation time of species, indicating relaxation of purifying selection with decrease of species-specific effective population size. Most importantly the linear regression lines of mitochondrial and nuclear Kn/Ks's from generation times of species are parallel, indicating congruent relaxation of purifying selection in both - mitochondrial and nuclear genomes. Thus our results provide evidence against any mitochondria-specific detrimental genetic processes in the mammalian species.

3264T

Molecular phylogeny of an autosomal region under natural selection. *V.A. Canfield[1], A. Berg[1], S. Peckins[1], S. Oppenheimer[2], K.C. Cheng[1].* 1) Penn State College of Medicine, Hershey, PA; 2) Oxford University, Oxford, UK.

The derived (A111T) variant of SLC24A5 is associated with lighter skin pigmentation compared to the ancestral allele. A111T is fixed or nearly fixed in most European, North African and Middle Eastern populations, extending east to Pakistan. In Europeans, a large genomic region of diminished variation on chromosome 15, nearly 150 kb in extent, includes SLC24A5. We analyzed the haplotypes in this region using existing genomic data. Eleven haplotypes, defined on the basis of 16 SNPs that span a 76 kb genomic region in which recombination was rare, account for 95% of the total. A single haplotype (here called C11) carries A111T, suggesting that its origin did not long predate the onset of selection. Haplotype C11 was the product of recombination between haplotypes C3 and C10, followed by the A111T mutation. C3 and C10 are both present in East Asia and the New World but virtually absent in Africa, suggesting that C11 originated outside of Africa, most likely in the Middle East. The current distribution of A111T is consistent with the view that it originated after the divergence between populations that settled Europe and those that settled East Asia.

3265F

MtDNA analysis of global populations supports that major population expansions began before Neolithic time. *H. Zheng[1], S. Yan[1,2], Z. Qin[1], L. Jin[1,2].* 1) School of Life Sciences, Fudan University, Shanghai, China; 2) Partner Institute for Computational Biology, Chinese Academy of Sciences.

Agriculture, as one of the most important innovations in human history, resulted in extensive population growths and human activities. However, whether major human expansions started after Neolithic Time still remained controversial. With the benefit of 1000 Genome Project, we were able to analyze a total of 1,277 samples without ascertainment bias from 15 populations in East Asia, Africa, Europe and Americas. From complete mtDNA of these random samples, we identified the expansion lineages and reconstructed the historical demographical variations in the four continents. In all the continents analyzed above, we found that most lineages underwent expansions coalesced before the first appearance of agriculture, but after the Last Glacial Maximum. Consistent to this observation, further analysis showed that the time of major population expansions was concordant to that of major lineage expansions, also between the Last Glacial Maximum and the Neolithic Time. Considering results in current and previous study, global mtDNA evidence showed that rising temperature after Last Glacial Maximum offered amiable environments and might be the most important factor for prehistorical human expansions. The initial population growth constituted a need for the introduction of agriculture, and might be one of the driving forces that led to the further development of agriculture.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3266W

Different DNA methylation of FOXP2 target genes in adult cortices of humans and chimpanzees. *U. Zechner[1], D. Seifert[1], E. Schneider[2], N. El Hajj[2], B. Navarro[3], I. Kondova[4], R.E. Bontrop[4], O. Bartsch[1], T. Haaf[2].* 1) Institute of Human Genetics, Mainz University Medical Center, Mainz, Germany; 2) Institute of Human Genetics, University of Würzburg, Germany; 3) Institute of Legal Medicine, University Medical Center of the Johannes Gutenberg University, Mainz, Germany; 4) Department of Comparative Genetics and Refinement, Biomedical Primate Research Center, Rijswijk, The Netherlands.

The molecular mechanisms governing language acquisition in the human brain are largely unknown. Up to now, mutations in the gene encoding the transcription factor FOXP2 (forkhead box P2) are the only reported monogenic cause of language and speech disorders in humans (Lai et al., 2001). A recent study analyzing differential gene expression in human neuronal cells overexpressing human or chimpanzee FOXP2 as well as in human and chimpanzee brain provided evidence for human-specific transcriptional regulation of FOXP2 target genes to be implicated in the development and evolution of language circuitry in humans (Konopka et al., 2009). We set out to investigate if the observed differential expression of FOXP2 target genes may be due to differential methylation of their promoter regions in human and chimpanzee brains. Overall, we quantified the methylation levels of 15 differentially expressed target genes in adult frontal cortices of 13 humans and 6 chimpanzees using bisulphite pyrosequencing. Two genes, NPTX2 and GJA12 (GJC2), displayed mean methylation levels that were significantly different in human and chimpanzee samples ($p < 0.05$). For both genes, the mean methylation levels were higher in the human (7.6% for NPTX2 and 37.3% for GJA12) compared to the chimpanzee (4.7% for NPTX2 and 27.5% for GJA12) brains. For the promoter region of the GJA12 gene whose mutations are associated with delayed speech development in Pelizaeus-Merzbacher-like disease (PMLD), we confirmed these results using a more in-depth methylation analysis of multiple CpG sites by classical bisulphite sequencing. The observed higher methylation levels in human brains correlate very well with the data of Konopka et al. (2009) that showed a downregulation of GJA12 expression in both neuronal cells overexpressing human FOXP2 and human brain samples. Thus, our findings suggest that differential methylation of regulatory regions is involved in regulating differential expression of at least a subset of FOXP2 target genes in human and chimpanzee brains. In addition, our study provides additional support for the use of comparative methylation analyses as an important methodology to detect evolutionary changes in gene regulation that have operated in human-specific brain development and function. Lai et al., 2001, Nature 413: 519–523 Konopka et al., 2009, Nature 462:213–217.

3267T

Skin Color Variation in the Orang Asli Tribes of Peninsular Malaysia. *K.C. Ang[1], M.S. Ngu[2], K.P. Reid[1], M.S. Teh[2], A.S. Zamzuraida[2], D.X.R. Koh[2], A. Berg[1], S. Oppenheimer[3], S. Hood[2], M.M. Clyde[2], B.M. Md-Zain[2], V.A. Canfield[1], K.C. Cheng[1].* 1) Penn State College of Medicine, 500 University Drive, Hershey, PA. 17033; 2) Universiti Kebangsaan Malaysia, Bangi. 43600. Selangor. Malaysia; 3) Inst. of Cognitive and Evolutionary Anthropology, University of Oxford, Oxford, United Kingdom.

Pigmentation is a readily scorable and quantitative human phenotype, making it an excellent model for studying multifactorial traits and diseases. Convergent human evolution from the ancestral state, darker skin, towards lighter skin colors involved divergent genetic mechanisms in people of European vs. East Asian ancestry. It is striking that the European mechanisms result in a 10–20-fold increase in skin cancer susceptibility while the East Asian mechanisms do not. Towards the mapping of genes that contribute to East Asian pigmentation there is need for one or more populations that are admixed for ancestral and East Asian ancestry, but with minimal European contribution. This requirement is fulfilled by the Senoi, one of three indigenous tribes of Peninsular Malaysia collectively known as the Orang Asli. The Senoi are thought to be an admixture of the Negrito, an ancestral dark-skinned population representing the second of three Orang Asli tribes, and regional Mongoloid populations of Indo-China such as the Proto-Malay, the third Orang Asli tribe. We have calculated skin reflectance-based melanin indices in 492 Orang Asli, which ranged from 28 (lightest) to 75 (darkest); both extremes were represented in the Senoi. Population averages were 56 for Negrito, 42 for Proto-Malay, and 46 for Senoi. The derived allele frequencies for SLC24A5 and SLC45A2 in the Senoi were 0.04 and 0.02, respectively, consistent with greater South Asian than European admixture. Females and individuals with the A111T mutation had significantly lighter skin ($p=0.001$ and 0.0039, respectively). Individuals with these derived alleles were found across the spectrum of skin color, indicating an overriding effect of strong skin lightening alleles of East Asian origin. These results suggest that the Senoi are suitable for mapping East Asian skin color genes. We will also report the results of ongoing skin color analysis of the Iban tribe from West Malaysia who are subjectively lighter than the Proto-Malay.

3268F

Exome and RNA Sequencing of French-Canadians reveals an excess of rare variants that are enriched at functionally important sites. *A. Hodgkinson, F. Casals, Y. Idaghdour, J. Hussin, V. Bruat, T. de Malliard, J-C. Grenier, J-P. Goulet, E. Gbeha, E. Hip-Ki, S. Girard, J-F. Spinella, V. Saillour, D. Sinnett, G. Rouleau, P. Awadalla.* Sainte Justine Research Centre, University of Montreal, Montreal, QC, Canada.

The French-Canadian population of Quebec are descendants of around 8,500 French settlers who colonised the province around 400 years ago. Since this migration event, the French-Canadian population has undergone massive population growth and has become genetically isolated from France with limited exchange with other non-French communities in the same geographical area. Subsequent colonisation events into new territories have also added to the rich demographic history of Quebec and it has been hypothesised that the high incidence of several genetic diseases are a consequence of regional founder effects. Here we analyse over 100 whole exome and over 1000 RNA sequencing datasets from French-Canadians and attempt to describe how this complex demographic scenario has shaped genetic variation in the region. Interestingly, we observe an excess of rare variants in the French Canadian population when compared to European populations and subsequent analysis points to an enrichment of alleles at functionally important sites in French-Canadians. Furthermore, we find intriguing differences in fundamental population genetic parameters between French Canadians and Europeans, highlighting potential variation in fitness and selection across the two groups.

3269W

Evidence for selection at the BDNF Val66Met polymorphism in 1000 Genomes Project populations. *C.V. Van Hout, A.G. Clark.* Dept of Molecular Biology and Genetics, Cornell University, Ithaca, NY.

The Val66Met SNP (rs6265) in the Brain Derived Neurotrophic Factor (BDNF) gene is of particular interest due to its pleiotropic effects in humans and model organisms, encompassing molecular, cellular, organ, and behavioral aspects (1). Previous studies of populations, including the Human Genome Diversity Panel, have reported that 1. population differentiation at this polymorphism is in the top 1% genome-wide, 2. extended linkage disequilibrium (LD) at this locus is consistent with recent positive selection of the Val allele, and 3. the Met allele likely arose after early humans migrated out of Africa (2). We evaluated the evidence for selection at the Val66Met locus in publicly available 1000 Genomes Project resources which, in comparison to previous studies, offers a larger number of individuals in subpopulations, and whole genome sequence, thus, increased resolution for signatures of selection. To elucidate gene flow at this locus, we applied the pairwise sequentially Markovian coalescent model (PSMC) (3) to genome sequence flanking the Val66Met variant, and estimated mean TMRCA for each subpopulation conditioning on the presence of the Met allele. Consistent with previous studies, population differentiation is high at the Val66Met SNP among continental groups as measured by Fst. Additionally, SNP specific and haplotype homozygosity based tests of selection (4) are consistent with recent positive selection of the Met allele. In contrast to previous reports, we did not observe evidence for positive selection favoring haplotypes harboring the ancestral Val allele. Furthermore, the haplotype carrying the Met allele is present (~2% MAF) in 1000 Genomes Project participants of native African ancestry, i.e., YRI and LWK, which is consistent with the arrival of the Met allele prior to migration out of Africa (or, with very low probability, back migration). We also find that the TMRCA for the Met haplotypes is far younger than expected given its frequency, a finding that is also consistent with positive selection for this haplotype. Understanding the evolutionary history of the Val66Met variant and its haplotype background in worldwide populations, is critical for the interpretation of the many studies of pleiotropic effects of this variant. 1. Bath and Lee, Cognitive, Affective, & Behavioral Neuroscience, 2006 2. Petryshen, et. al., Molecular Psychiatry, 2010 3. Li and Durbin, Nature, 2011 4. Voight, et. al., PLOS Biology, 2006.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3270T

Mobile element evolution of the *Callithrix jacchus* (common marmoset). *M.K. Konkel[1], J.A. Walker[1], B. Ullmer[2], R. Hubley[3], A.F.A. Smit[3], M.A. Batzer[1] for the Marmoset Genome Sequencing and Analysis Consortium.* 1) Department of Biological Sciences, Louisiana State University, Baton Rouge, LA; 2) Department of Computer Science, Center for Computation and Technology (CCT), Louisiana State University, Baton Rouge, LA; 3) Computational Biology, Institute for Systems Biology, Seattle, WA.

The common marmoset (*Callithrix jacchus*) is the first New World monkey (NWM) species for which a draft genome sequence is available. The purpose of this study was to investigate mobile element propagation and evolution in the NWM lineage through the lens of the common marmoset. Moreover, our aim was to examine NWM phylogeny and the population genetic composition of common marmosets raised in captivity in one of three U.S. Primate Centers (New England, Southwest, Wisconsin) using selected retrotransposon markers. Our analysis of the *C. jacchus* draft genome assembly [CalJac 3.2] revealed that the common marmoset has an overall repeat composition similar to other sequenced and analyzed primate genomes. For example, in the lineage leading to the common marmoset, two primary non-LTR (long terminal repeat) retrotransposon families, long interspersed elements (LINE1, L1) and short interspersed elements (SINEs) contribute the bulk of lineage-specific insertions. However, these retrotransposon families evolved uniquely in the lineage leading to *C. jacchus*. The NWM-specific L1 lineage intersects with the human-derived L1 tree between L1PA6f and L1PA7, and evolved in a mostly linear manner. In contrast, the reconstruction of the *Alu* subfamily evolution shows a `star-like' pattern with several subfamilies being active in parallel. The youngest *Alu* subfamilies are derivatives of *Alu*Ta15, a previously identified NWM-specific subfamily. Moreover, about half of the NWM-specific *Alu* elements appear to be derived from *Alu*Ta15 or its derivatives, indicating that a small burst of *Alu* retrotransposition coincided with the rise of *Alu*Ta15 and its derived subfamilies. Intriguingly, *Alu*Ta15 resulted from a gene conversion event of two *Alu* subfamilies, indicating that the new *Alu* subfamily may have initially escaped repression. In recent history, *Alu* activity appears to be slower than in humans and rhesus macaques. Our population genetic analysis with polymorphic *Alu* markers using the Structure software indicates the presence of population structure with at least two different clusters, indicating diversity among common marmosets. In addition, we investigated the evolution of *Alu* subfamilies with respect to the radiation of NWMs through the lens of the common marmoset. This study provides insights into the evolution of mobile elements in the lineage leading to the common marmoset and represents the first comprehensive analysis of the mobilome of an NWM species.

3271F

Association of GSTT1,MI & M3 gene polymorphism with age and quantity of smoke in Northern Indian COPD patients. *R.K Shukla[1], S. Kant[1], S. Bhattacharya[2], B. Mittal[3].* 1) Dept of Pulmonary Medicine, CSMMEDICAL UNIVERSITY, (ERSTWHILE KGMC), LUCKNOW, INDIA, India; 2) Dept of Physiology, CSMMEDICAL UNIVERSITY, (ERSTWHILE KGMC), LUCKNOW, INDIA, India; 3) Dept of Medical Genetics, Sanjay Gandhi Post Gratute Insitute of Medical Sciences, Lucknow, India.

Background: There is increasing evidence for a close relationship between aging and chronic inflammatory diseases. COPD is a chronic inflammatory disease of the lungs, which progresses very slowly and the majority of patients are therefore elderly. Smoking is the major risk factor in development of COPD. In the present study, we assess whether age and quantity of smoke affect in metabolism of phase 2 enzyme gene polymorphisms (GSTT1, GSTM1 & GSTM3) in northern Indian Population. Material & Methods: In this case Control study, we have enrolled 412 study subjects (208 Controls and 204 COPD patients). All the subjects were enrolled after the spirometry evaluation. Smoking quantity were taken in predefine questionnaire and genotyping were done by PCR-RFLP method. Results: The frequency of homozygous GSTM1 null was found to be significantly higher in COPD patients as compared with healthy controls (OR=2.08; 95%;CI=1.40–3.09; P=0.0001), but there were no significant differences in the distribution of genotypes of polymorphisms of GSTT1 and GSTM3 between COPD patients and healthy controls. In subgroup analysis in age we found GSTM1 Null/Present polymorphism significantly associated with the age group (50–60 years) for COPD development (43.5%/21.2%, p=0.015).In case of Smoking, quantity of smoke was not associated with the COPD patients. Conclusion: we conclude that GSTM1 null polymorphism was associated with COPD patients, and we also observe that GSTM1 null polymorphism was associated between age group 50–60 years COPD patients.

3272W

Polymorphic miRNA genomic sequences embedded in CNV sites: a prospective screening in Brazilian admixed population sample. *T.C.L. Lins[1], D.E. Jimenez[2], A.P.M. Barbosa[2], R.W. Pereira[1,2].* 1) Universidade de Brasília, Brasília, DF, Brazil; 2) Universidade Católica de Brasília, Brasília, DF, Brazil.

Copy number variation (CNV) has a major impact on gene expression in the human genome. Furthermore, the mechanism of post-transcriptional regulation by microRNAs is one of the key components in regulation of gene expression. Bioinformatics studies have identified several pre-miRNA genomic sequences embedded in CNV sites that may vary in a polymorphic manner within a population. In this work we aimed to screen five putative polymorphic miR-CNV loci by real-time quantitative PCR in an admixed sample of the Brazilian population (n=90), with genetic ancestry previously estimated. The miR-499b region was determined to be the most polymorphic, ranging from zero to 6 copies, with 47% of individuals showing 2 copies. The region with the lowest variation was that of miR-650, where 2 copies were observed in 98% of subjects, and 3 copies in 2%. An interesting phenomenon was observed in the region of miR-126, where the most frequent copy number was 3 (66.7%), whereas in a diploid organism, the most frequent expected number would be 2. No correlation with genetic ancestry was observed. In summary, we have validated these miR-CNV sites as polymorphic in the Brazilian population and will further quantify the expression of these microRNAs, a factor which may be impacted by the number of genomic copies found.

3273T

Ascertainment Bias in Microsatellites: Impact on Estimates of Mutation Rates. *B. Li, M. Kimmel.* Depts. of Statistics and Bioengineering, Rice University, Houston, TX.

Microsatellite loci play an important role as markers for identification, disease gene mapping and evolutionary studies. Mutation rate, which is of fundamental importance, can be obtained from interspecies comparisons, which however are subject to ascertainment bias. This bias arises when a locus is selected based on its large allele size in one species (cognate species 1), in which it is first discovered. This bias is reflected in average allele length in any non-cognate species 2 being smaller than that in species 1. This phenomenon was observed in various pairs of species, including comparisons of allele sizes in human and chimpanzee. Various mechanisms were proposed to explain ascertainment bias. Here, we examine the framework of a single-step asymmetric and unrestricted stepwise mutation model with genetic drift. The mechanism of ascertainment bias in this model is a tighter correlation of allele sizes within a cognate species 1 than of allele sizes in two different species 1 and 2. This is a follow-up presentation based on the work presented at the previous meeting. The current study includes more realistic demography to consider the bottleneck of recent human population and discusses a wider range of reasonable settings of model parameters with sensitivity analysis of each parameter. We will show computations of the expected bias, given the mutation rates, effective population sizes of species 1 and 2 and their ancestral population, time of separation of species 1 and 2, the age of the allele and the ascertaining scheme. In particular, we show that when the past demographic history of the cognate and non-cognate taxa are different, the rate and directionality of mutations will impact the allele sizes in the two taxa differently than the simple effect of ascertainment bias. This effect may exaggerate or reverse the effect of difference in mutation rates. This effect may exaggerate or reverse the effect of difference in mutation rates. The theoretical results are cross-validated by simulations using simuPOP, a general purpose forward-time population genetics simulation framework. We re-analyze literature examples, and conclude that despite the bias, the microsatellite mutation rate in human exceeds that in chimpanzee.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3274F

Characterizing recent evolutionary changes on the human lineage using the high-coverage Denisovan genome. *F. Racimo, M. Kircher, J. Kelso, S. Pääbo, The Archaic Genome Analysis Consortium.* Max Planck Institute for Evolutionary Anthropology, D04103 Leipzig, Germany.

We have increased the sequence coverage of the previously published genome of an archaic human from Denisova cave in southern Siberia to approximately 30-fold coverage. This high-coverage genome provides the opportunity to identify the complete set of sequence changes that have risen to high frequency or fixation in modern humans since the split from the common ancestor with Denisovans. These recent changes may set anatomically modern humans apart from other extinct hominin forms. Here we present an updated catalog of all single nucleotide changes (SNCs) and insertion/deletion events (InDels) now identified using the high-coverage genome data. Using variation data for modern humans we are able to identify sites where the Denisova genome sequence is ancestral while the derived allele is either fixed or at high frequency (>90%) in modern human populations. We can confidently identify 260 fixed and 393 high-frequency amino acid substitutions in protein-coding genes. A further set of nucleotide changes and InDels in 3' and 5' UTRs, splice sites, miRNAs and motif patterns in regulatory regions were also identified. An analysis of the functional impact of specific amino acid changes, as well as Gene Ontology and Human Phenotype enrichment analyses, allow us to prioritize substitutions that might have important phenotypic effects.

3275W

Recurrent tissue-specific mtDNA mutations are common in humans. *S.M. Williams[1,2], D.C. Samuels[1], B. Li[1], Z. Song[1], E. Torsteson[1], A. Rokas[3], T.A. Thornton-Wells[1], J.H. Moore[2], T.M. Hughes[1], R.D. Hoffman[4], J.L. Haines[1], D.P. Mortlock[1], C. Li[1].* 1) Ctr Human Gen Res, Vanderbilt Univ Med Ctr, Nashville, TN; 2) Genetics, Dartmouth College, Hanover, NH; 3) Biological Sciences, Vadnerbilt University, Nashville, TN; 4) Pathology, Vanderbilt University, Nashville, TN.

Mitochondrial DNA (mtDNA) variation can affect cell function and phenotypes; therefore, knowing its distribution is of importance to understanding many human diseases. Tissue-specific differences are also an important source of variation in disease presentation. Inter-tissue mtDNA variation (heteroplasmy) within an individual has been generally assumed to be random. However, we discovered that certain tissues, notably kidney, liver and skeletal muscle, display identical recurrent tissue-specific mutations in different individuals that are absent from other tissues in the same individuals. This was done using massively parallel sequencing to assess heteroplasmy across 23 samples from two subjects taken at autopsy. Eight tissue-specific, recurrent mutations were identified in both our subjects, and they all are tightly clustered between nucleotide positions 60 and 408. All of these recurrent mutations are within or in very close proximity to sites known to regulate mtDNA replication, strongly implying that these variations alter the replication dynamics of the mutated mtDNA genome. No other tissue-specific mutations were found elsewhere in the mtDNA. The most parsimonious explanation of the data is that these frequently repeated mutations experience tissue-specific positive selection, probably through replication advantage.

3276T

Associations of Human Leukocyte Antigen G with Resistance and Susceptibility to HIV-1 Infection in the Pumwani Sexworker Cohort. *M. Luo[1,2], W. Turk[1,2], J. Kimani[2,3], C. Wachihi[3], T. Bielawny[1], T. Ball[1,2], F. Plummer[1,2].* 1) HIV and Human Genetics, Public Health Agency of Canada, Winnipeg, Manitoba, Canada; 2) Department of Medical Microbiology, University of Manitoba; 3) University of Nairobi, Kenya.

Objective: To determine the association between HLA-G genotypes and resistance or susceptibility to HIV-1. **Design:** A group of sex trade workers in Pumwani, Kenya can be epidemiologically defined as resistant to HIV-1 infection despite frequent exposure and provide an example of natural protective immunity. Human leukocyte antigens (HLA) class I and II molecules have been shown to be associated with resistance/susceptibility to infection in this cohort. HLA-G is a non-classical class I allele that is primarily involved in mucosal and inflammatory response which is of interest in HIV-1 resistance. **Methods:** In this study, we used a sequence-based typing method to genotype HLA-G in this cohort and examined the influence of HLA-G genotypes on resistance to HIV-1 infection. **Results:** The G*01:01:01 genotype was significantly enriched in the HIV-1 resistant women (p=0.002, OR:2.11, 95%CI:0.259–0.976), whereas the G*01:04:04 genotype was significantly associated with susceptibility to HIV-1 infection (p=0.039, OR:0.502, 95%CI:0.259–0.976). Kaplan-Meier survival analysis correlated with these results. G*01:01:01 genotype was associated with significantly lower rate of seroconversion (p=0.001). Whereas, G*01:04:04 genotype was significantly associated with an increased rate of seroconversion (p= 0.013). The associations of these HLA-G alleles are independent of other HLA class I and II alleles identified in this population. **Conclusions:** Our study showed that specific HLA-G alleles might play an important role in resistance/susceptibility to HIV-1 acquisition in this high-risk population. Since HLA-G is important in mucosal and inflammatory responses further studies are needed to better understand it's functional significance in HIV-1 transmission.

3277F

Alleles at the rs3212368 SNP at the 3'UTR of the MC1R gene are associated with human pigmentation by possible microRNA influence. *L.A. Marano[1], A.L. Simoes[1], E.A. Donadi[2], C.T. Mendes-Junior[3].* 1) Departamento de Genetica, Faculdade de Medicina de Ribeirao Preto, Universidade de Sao Paulo,14049-900, Ribeirao Preto-SP, Brazil; 2) Divisao de Imunologia Clínica, Departamento de Clinica Medica, Faculdade de Medicina de RibeiraoPreto, Universidade de Sao Paulo, 14048-900, Ribeirao Preto-SP, Brazil; 3) Departamento de Quimica, Faculdade de Filosofia, Ciencias e Letras de Ribeirao Preto, Universidade deSao Paulo, 14040-901, Ribeirao Preto-SP, Brazil.

Some human genetic markers have been associated to prediction of externally visible traits such as skin, hair and eye color. Prediction of such traits is potentially useful for forensic genetics casework in several situations. Among the known genes influencing these traits, melanocortin-1 receptor (MC1R) encodes a protein in melanocytes responsible for melanin synthesis regulation. Many MC1R alleles have already been associated to red hair, freckles and fair skin in Europeans but none of the studies done so far considered the MC1R regulatory regions or expression levels. Considering that some 3' untranslated region (UTR) mutations may create potential microRNA targets on the mRNA and alter its stability, the present study aims to genotype the 3'UTR SNP (rs3212368; +2346 A/G) of MC1R gene on a Brazilian population sample and to analyze its functional influence on human pigmentation by in silico methods. When the sample was stratified according to pigmentation features, the A allele was associated with hyperpigmentation (black hair, dark skin and dark brown eyes), while the G allele was associated with a lighter pigmentation (brown hair, fair skin and green eyes) of the evaluated traits. No association concerning the presence or absence of freckles was identified. The in silico analysis revealed the existence of several miRNAs with stronger affinity for the G allele, which may result in a diminished stability of MC1R mRNA carrying the +2346 G allele and lower MC1R expression levels and eumelanin production, thus resulting in lighter phenotypes such as green eyes and fair skin. Experimental studies must be carried out in order to corroborate the presented hypothesis.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3278W

Targeted Re-Sequencing of 328 Inherited Disease-Associated Genes in a Family Trio and HapMap Populations using the Ion Ampliseq™ Inherited Disease Panel and Ion PGMTM Semiconductor Sequencing. *I. Casuga, B. Kong, D. Joun, S-M. Chen, C-Y. Li, D. Ruff, R. Bennet, M. Shannon.* Ion Torrent part of Life Technologies 7000 Sierra Point Way South San Francisco, CA.

The Ion Ampliseq™ Inherited Disease Panel (IDP) provides ready-access to hundreds of genes involved in some of the most common inherited diseases in humans, making it ideal for broad targeted re-sequencing studies of individuals and populations. The panel employs more than 10,000 primer pairs to amplify the coding exons of 328 genes that are associated with neuromuscular, cardiovascular, developmental, metabolic and other types of disorders. This panel is part of the transformative Ion AmpliSeq™ technology suite that enables the selective amplification of 10s to 1000s of target sequences in a single multiplexed PCR. The IDP is comprised of only 3 pools of primer pairs, requires just 30ng of sample DNA, and can generate comprehensive data from a sample on a single 316TM chip using Ion PGMTM semi-conductor sequencing. In this study, the panel was used to examine the set of genes in a family trio from the CEPH population as well as in 24 unrelated HapMap samples from 4 ethnic populations (CEPH, YRI, CHB and JPN). We demonstrated Mendelian inheritance within the trio by analyzing known variants passed from parent to child. We also demonstrated highly similar performance of IDP with samples from all 4 ethnic populations tested as well as the ability to identify genetic variants between populations, which highlight the benefits of using the panel for studies of genetically diverse populations. Taken together, the results of this study demonstrated that the IDP can be used to detect multiple variants in more than 300 genes in families and across populations and is ideally suited for a broad range of investigations, including genetic risk and susceptibility studies in families and disease-association studies in ethnic populations.

3279T

Sequencing of an extended pedigree in Western chimpanzees. *O. Venn[1], I. Turner[2], Z. Iqbal[1], I. Mathieson[1], N. de Groot[3], G. McVean[1,2].* 1) Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, Oxfordshire, United Kingdom; 2) Department of Statistics, University of Oxford, Oxford, Oxfordshire, United Kingdom; 3) Department of Comparative Genetics and Refinement, Biomedical Primate Research Centre, Rijswijk, The Netherlands.

A wide spectrum of processes including crossover, gene conversion, mutation and DNA repair impact genomes from the level of single nucleotide changes to multiple megabases. Although many of these processes have been studied biochemically, their relative genome-wide distributions are not fully characterized. The analysis of genetic variation estimated from high-throughput population sequencing is a potentially powerful approach to study such processes. However, a major challenge in interpreting this information is the detection of erroneous, mis-assembled, or incorrectly genotyped genetic variants. Currently these errors are identified using ad hoc filters, which have limited application for the analysis of *de novo* mutation events.

We present an approach to study rare events in large genomes through the whole genome sequencing of a Western chimpanzee, *Pan troglodytes verus*, pedigree. We sequenced three founder, three F1, and three F2 individuals to a combined mapped coverage of 315x. This combined coverage is utilized effectively because 7/8 of the founder genomes are transmitted at least once, thereby improving precision to identify and validate genomic changes of different types. To detect events we called variants and Indels across the samples using both alignment-based and *de novo* assembly algorithms. The data enabled 99.2% of sites that are polymorphic across all individuals to be phased unambiguously. By resolving transmission vectors we were able to accurately observe a wide spectrum of genomic changes enabling insight into biological processes. For instance, we describe the impact of meiotic recombination on the chimpanzee genome at crossover, non-crossover, and gene conversion events across the two generations.

3280F

Allelic and genotypic associations of SNPs of the OCA2 and HERC2 genes with eye, hair and skin pigmentation and the presence of freckles in Brazil. *E.S. Andrade[1], N.C.A. Fracasso[1], P.S. Strazza Júnior[2], A.L. Simões[1], C.T. Mendes-Junior[3].* 1) Department of Genetics, School of Medicine of Ribeirao Preto, University of Sao Paulo, 14049-900, Ribeirao Preto, Sao Paulo, Brazil; 2) Department of Biology, School of Philosophy, Sciences and Letters of Ribeirao Preto, University of Sao Paulo, 14040-901, Ribeirao Preto, Sao Paulo, Brazil; 3) Department of Chemistry, School of Philosophy, Sciences and Letters of Ribeirao Preto, University of Sao Paulo, 14040-901, Ribeirao Preto, Sao Paulo, Brazil.

Panels composed of Single Nucleotide Polymorphisms (SNPs) in genes related to pigmentation, when associated with different phenotypes such as skin, eyes and hair color can be very useful in criminal cases assisting in predicting the physical appearance of an individual from an unknown biological sample found in crime scenes. The OCA2 (oculocutaneous albinism II) gene (15q11.2-12) encodes the protein P, a transmembrana protein involved in small molecule transport, specifically tyrosine, a precursor of melanin (Branicki et al. 2008). Mutations in this gene result in type 2 oculocutaneous albinism. Its neighboring gene, HERC2 (HECT and RLD domain containing E3 ubiquitin protein ligase) (15q13), encodes for functional protein domains involved in intracellular transport (Yonggang et al. 2000) and has been reported to influence OCA2 expression (Eiberg et al., 2008; Kayser et al., 2008; Sturm et al., 2008). Both genes have been associated with skin/hair/eye pigmentation variability (Eiberg et al., 2008; Branicki et al. 2009). The present study aimed at evaluating the influence of OCA2 and HERC2 alleles and genotypes in the determination of pigmentation traits in a highly admixed Brazilian population sample. Seven SNPs (four from OCA2 gene and three from the HERC2 gene) were evaluated in 286 unrelated individuals from the Ribeirao Preto area, located at the Northwestern region of the Sao Paulo State, Southeastern Brazil. SNPs were genotyped by PCR-RFLP, followed by Polyacrylamide Gel Electrophoresis. The associations' tests (Fisher's Exact Test and Odds Ratio) were performed using the GraphPad InStat 3.06 software. All the seven SNPs presented one allele associated with phenotypes from at least two pigmentation features and the alternative allele associated with the opposite phenotypes from the same trait. The A allele of the rs1800404 SNP (OCA2), for instance, was associated with the presence of blond/red hair, pale skin and freckles and blue/green eyes, whereas the G allele was associated with dark eyes, hair and skin and the absence of freckles. The genotypic associations followed the same pattern for all seven SNPs. The SNPs panel showed to have predictive value for phenotypic characteristics. However, further studies, including haplotype analysis, are needed to determine the effects of SNPs in the OCA2 and HERC2 genes in human pigmentation and their forensic application. FINANCIAL SUPPORT: CNPq/Brazil (Grants 478843/2009-7 and 305493/2011-6) and CAPES.

3281W

AKT3, ANGPTL4, eNOS3, and VEGFA associations with high altitude sickness in Han and Tibetan Chinese at the Qinghai-Tibetan Plateau. *N. Buroker[1], X.-H. Ning[2], Z.-N. Zhou[3], K. Li[4], W.-J. Cen[4], X.-F. Wu[3], W.-Z. Zhu[3], C. R. Scott[1], S.-H. Chen[1].* 1) Pediatrics, 356320, University of Washington, Seattle, WA; 2) Division of Cardiology, Seattle Children's Hospital. Institute. Foundation, Seattle, WA; 3) Laboratory of Hypoxia Physiology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Beijing, China; 4) Lhasa People Hospital, Lhasa, Tibet, China.

Mountain sickness (MS) occurs among humans visiting or inhabiting high altitude environments. Genetic analyses of the AKT3, ANGPTL4, eNOS3 and VEGFA genes was made between lowland (Han) and highland (Tibetan) Chinese. Ten SNPs within these genes were evaluated in Han and Tibetan patients with Acute (A) and Chronic (C) MS. We compared 74 patients with AMS to 79 Han unaffected with MS as well as 48 CMS patients to 31 Tibetan unaffected. The ten SNPs studied are AKT3 (rs4590656, rs2291409), ANGPTL4 (rs1044250), eNOS3 (rs1007311, rs1799983) and VEGFA (rs79469752, rs13207351, rs28357093, rs1570360, rs3025039). Direct sequencing was used to identify individual genotypes for these SNPs. Hb, Hct and RBC were found to be significantly associated with the AKT3 SNP (rs4590656), Hb was found to be associated with the eNOS3 SNP (rs1007311) and RBC was found to be significantly associated with the VEGFA SNP (rs1570360) in Tibetan patients with CMS. CMS patients were found to diverge significantly for both eNOS3 SNPs as measured by genetic distance (0.042, 0.047) and for the VEGFA SNP (rs28357093) with a genetic distance of 0.078 compared to their Tibetan control group. HR was found to be significantly associated with the eNOS3 SNP (rs1799983) and SaO2 was found to be significantly associated with the VEGFA SNPs (rs13207351, rs1570360) in Han patients with AMS. The Han and Tibetan control groups were found to diverge significantly for the ANGPTL4 SNP and VEGFA SNP (rs28357093) as measured by genetic distance (0.049 & 0.073), respectively. Seven of the SNPs from non-coding regions are found in the transcriptional factor response elements and their possible role in gene regulation was evaluated with regard to HAS. AMS and CMS were found to be significantly associated with the four genes compared to their Han and Tibetan control groups, respectively, indicating these nucleotide alterations have a physiological effect for the development of high altitude sickness.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3282T

A population genetic measure of the de novo mutation rate using identity-by-descent estimates. *C.W.K. Chiang[1], J. Li[2], M.G. Ehm[3], M.R. Nelson[3], J. Novembre[1].* 1) Department of Ecology and Evolutionary Biology, UCLA, Los Angeles, CA; 2) Department of Human Genetics, University of Michigan, Ann Arbor, MI; 3) Quantitative Sciences, GlaxoSmithKline, RTP, NC, USA; Upper Merion, PA, USA; and Stevenage, UK.

The *de novo* mutation rate is an important parameter that underlies much of population genetics theory and is critical for understanding the genetic architecture of human traits and disease. Estimates of mutation rate have been derived using family-based approaches based on direct sequencing in pedigrees, or using evolution-based approaches by modeling the demographic history of a population and analyzing its impact on the site frequency spectrum. However, family-based approaches provide a mutation rate estimate based on only a few families and evolution-based approaches are sensitive to the assumed demographic model. Therefore, a parallel method for estimating mutation rates would be useful to complement the current approaches. Here we present a novel method to estimate an average mutation rate using identity-by-descent (IBD) tracts coupled with sequence data. Using genome-wide genotyping data at over 716,000 SNPs for 7,401 individuals of European descent, we inferred pair-wise IBD tracts by Beagle. In total we detected 4.96 million tracts of IBD over 1 cM in length, averaging 1.79 cM of IBD per 10,000 pairs of individuals, consistent with the observed average in Europeans. By cross-referencing with deep sequencing data in the exons of 202 genes distributed genome-wide for these individuals, we identify putatively novel mutations since the IBD event, and use these to infer a preliminary estimate of $\sim 8\times10^{-8}$ mutations per generation. While this estimate is higher than previous estimates, we anticipate that by developing a more sophisticated framework to take into account the uncertainty of imputation errors, haplotype phasing errors, and IBD tract lengths we will produce a more refined estimate of the average *de novo* mutation rate. These measures provide context to aid in the interpretation of how variants are related to phenotypes.

3283F

Determination of PON1 and P2RY12 polymorphisms distribution in Hungary living population samples. *B. Melegh, I. Janicsek, Cs. Sipeky, B. Duga, B.I. Melegh, L. Jaromi, L. Magyari, J. Bene.* Department of Medical Genetics, University of Pecs, Pecs, Hungary.

Objectives: Clinical efficacy of clopidogrel in the antiplatelet therapy has recently been in the focus of clinical investigations. Recent studies suggest that, among others, genetic factors like SNPs in genes of specific receptors and enzymes have a large contribution to the high interindividual variability observed in clopidogrel response. The pharmacogenomic investigation of these factors might also have clinical benefit. The purpose of our study was to determine the distribution of variants of paraoxonase 1 (PON1) and purinergic receptor P2Y, G-protein coupled, 12 (P2RY12) genes in 477 average Hungarian and 491 Roma (gipsy) samples. For the PON1 gene rs662 (Q192R) and rs854560 (L55M), for the P2RY12 gene 3 SNPs, rs2046934, rs6798347 and rs6801273 were analyzed as the most frequently investigated naturally occurring variants. Methods: The different genotypes were determined by PCR-RFLP. Results: For the Q192R variant the frequency of the GG genotype was found more than 2.5 times higher in the Hungarian group compared to the Romas (11.3% vs 4.3%, $p<0.001$). In the G allele frequencies similar, significant difference could be observed (24.8% vs 31.7%, $p<0.001$). For the L55M variant the frequency of the TT genotype was more than 2.5 times higher in the Roma group than in the Hungarian one (10.0% vs 3.8%, $p<0.001$). For the 3 P2RY12 variants significant differences were found only in rs2046934: the frequency of the CC genotype is 7 times higher in Hungarians than in Romas (1.4% vs 0.2%, $p<0.05$). Conclusion: The data presented here confirm major differences between the distribution of PON1 and P2RY12 variants in Hungarian and Roma patients that might have clinical implications.

3284W

Mapping the genetic diversity of HLA haplotype in Asia populations. *W.Y Saw[1], R.T.H Ong[2], C.C Khor[3], N. Kato[4], Y.Y Teo[1,2,3,5].* 1) Life Science Institute, National University of Singapore, Singapore; 2) Saw Swee Hock School of Public Health, National University of Singapore, Singapore; 3) Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore, Singapore; 4) Department of Gene Diagnostics and Therapeutics, Research Institute, International Medical Center of Japan, Japan; 5) Department of Statistics and Applied Probability, National University of Singapore, Singapore.

Numerous studies have identified disease associations with genetic variants located in the HLA regions, and recent population genetics studies have also discovered genetic evidence of population differences at HLA loci between seemingly homogeneous populations in North and South China. Here we report a systematic survey of the entire major histocompatibility complex (MHC) across 3,499 individuals from eighteen populations from seven countries in Asia, including China, Japan, India, Sri Lanka, Thailand, Vietnam and Singapore. In the study, we are looking at the MHC region which locates on chromosome 6, 25Mb to 35Mb in humans, consisting of approximately 1600 common SNPs across the eighteen populations. In this study, we have investigated different strategies to perform haplotype phasing of the genotype data from these samples in order to minimize artifactual diversity introduced as a result of the phasing. By comparing the haplotype diversity across the samples from these populations, we observed a geographical cline in terms of haplotype variation across populations in East and Southeast Asia, and even between seemingly homogeneous populations within Japan and China separately, we observe differences in the genetic architecture of the MHC region.

3285T

Long runs of homozygosity contain a higher fraction of all genome-wide deleterious homozygotes relative to the fraction of all genome-wide non-damaging homozygotes. *Z.A. Szpiech[1,2], J. Xu[3], T.J. Pemberton[1], W. Peng[3], S. Zöllner[4], N.A. Rosenberg[1], J.Z. Li[3].* 1) Department of Biology, Stanford University, Stanford, CA; 2) Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI; 3) Department of Human Genetics, University of Michigan, Ann Arbor, MI; 4) Department of Biostatistics, University of Michigan, Ann Arbor, MI.

Exome sequencing offers the potential to study the genomic variables that underlie patterns of deleterious variants. Here we consider how runs of homozygosity (ROH) correlate with these variants. In general, short ROH represent homozygosity for ancient haplotypes; ROH of intermediate length represent homozygosity arising from background relatedness in a population; and long ROH result from recent parental relatedness. Because many deleterious variants act recessively, they can be more effectively purged by selection in ROH-rich regions than in non-ROH regions. Thus, ROH could contain a lower fraction of all genome-wide damaging homozygotes relative to the fraction of all genome-wide non-damaging homozygotes. However, inbreeding can present low-frequency variants in homozygous form, and many of these are more likely to be deleterious than common variants. When the homozygous deleterious variants are not lethal, one might expect ROH to contain a higher fraction of genome-wide deleterious homozygotes compared to the fraction of all genome-wide neutral homozygotes, and longer (newer) ROH might have a higher fraction than shorter (older) ROH. We sequenced the exomes of 27 individuals from 6 HGDP populations, representing a wide range of genomic ROH coverage (4%–46%, Pemberton et al. 2012). We classified non-reference variants as damaging or non-damaging using PolyPhen2. Consistent with the work of Lohmueller et al. (2008), the proportion of variants private to a given population that are damaging increases with the distance of the population from Africa. For the individual genomes, the fraction of all genome-wide homozygotes lying in ROH is positively correlated with the total length of ROH in the genome. Damaging homozygotes show a significantly higher fraction falling into any size ROH than non-damaging homozygotes ($p < 0.001$). This trend is also significant for long ROH only ($p < 0.05$) but not for intermediate or short ROH only. These results can be explained by ROH containing more rare and damaging variants in homozygous form as a result of inbreeding. Our study provides a demonstration that long ROH harbor disproportionately more deleterious homozygotes than would be predicted solely by the ROH coverage of the genome.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3286F

The genetic structure of Western Balkan populations based on autosomal and haploid markers. *K. Tambets[1], L. Kovacevic[1,2,3], D. Primorac[4], G. Lauc[5], A. Leskovac[6], Z. Jakovski[7], K. Drobnic[8], S. Kovacevic[9], T. Bego[10], E. Metspalu[11], D. Marjanovic[2,5], R. Villems[1,11].* 1) Estonian Biocentre, Tartu, Estonia; 2) Institute for Genetic Engineering and Biotechnology, Sarajevo, Bosnia and Herzegovina; 3) Faculty of Pharmacy, University of Sarajevo, Sarajevo, Bosnia and Herzegovina; 4) University Center of Forensic Science, Split, Croatia; 5) Genos doo, Zagreb, Croatia; 6) Vinca Institute of Nuclear Sciences, Belgrade, Serbia; 7) Institute of forensic medicine, criminology and medical deontology and Medical Faculty, University of "St. Cyril and Methodious", Skopje, Republic of Macedonia; 8) Faculty for Justice and Security, Ljubljana, Slovenia; 9) Forensic Center, Bozova glavica, Danilovgrad, Montenegro; 10) Faculty of Farmacy, University of Sarajevo, Sarajevo, Bosnia and Herzegovina; 11) University of Tartu, Tartu, Estonia.

Contemporary inhabitants of the Balkan Peninsula belong to several ethnic groups of diverse cultural backgrounds. In this study, three ethnic groups from Bosnia and Herzegovina - Bosniacs, Bosnian Croats and Bosnian Serbs - and four other Slavic-speaking Western Balkan populations: Serbians, Croatians, Macedonians from Republic of Macedonia, Montenegrins, and Albanian-speaking Kosovars have been characterized for the genetic variation of 660 000 autosomal single nucleotide polymorphisms. Genetic structuring of Western Balkan populations has been analyzed in a global context. Comparison of the variation within autosomal and haploid data sets of studied Western Balkan populations revealed their genetic closeness regardless of a genetic system inspected, in particular among the Slavic speakers. Hence, culturally diverse Western Balkan populations are genetically very similar to each other. Only the Kosovars show slight differences both in the variance of autosomal and uniparentally inherited markers from the other populations of the region, possibly also due to their historically strict patrilineality. In a more general perspective, our results reveal clear genetic continuity between the Near Eastern and European populations, lending further credence to extensive, likely multiple and possibly bidirectional ancient gene flows between the Near East and Europe, cutting through the Balkans.

3287W

Analysis of TLR4 SNPs 299 and 399 in a population of full term Wisconsin infants. *D. Pillers[1], J. DeValk[1], M. Baker[1], S. Schrodi[2], S. Tokarz[1].* 1) University of Wisconsin-Department of Pediatrics, Madison, WI; 2) Center for Human Genetics, Marshfield Clinic Research Foundation, Marshfield, WI.

Background: Toll-like receptors (TLRs) are present in many cell types and serve as the first point of defense in the innate immune system by initiating the inflammatory cascade in response to infectious microorganisms. Single nucleotide polymorphisms (SNPs) are present in many TLR genes and have been associated with many disorders of infection and inflamation. TLR4 229 and 399 SNPs are associated with increased susceptibility to infection from various pathogens. As infection can play a role in the development of preterm labor, TLR SNPs that alter the response to infection may have a different frequency in the pre-term population as compared to individuals of full-term birth status. Most current population data on SNP frequency does not exclude individuals who were born preterm. To this end, we screened a cohort of full-term infants, selected at random, for a subset of TLR4 SNPs to determine their population frequencies. **Objective**: Determine the allele and genotype frequencies of TLR4 299 and 399 in a population of full-term infants **Design/Methods:** We have analyzed the genotypes of these SNPs in a population of 487 full-term Wisconsin newborns between 2010 and February 2012. Anonymized DNA samples from 487 (≥37 weeks of gestational age) infants were obtained from the Wisconsin State Laboratory of Hygiene. SNP assays were purchased from Applied Biosystems Inc. (ABI) and performed on the ABI StepOnePlus system. Data were analyzed using ABI StepOne Software. **Results:** The allele frequencies for TLR4 299 found in this population are 0.94 (A) and 0.06 (G) and 0.95 (C) and 0.05 (T) for TLR4 399. For TLR4 299 the observed genotype frequency was 0.88 (AA), 0.13 (AG), and 0.0 (GG). In TLR4 399 we observed genotype frequencies of 0.90 (CC), 0.10(CT), and 0.0 (TT). **Conclusions:** As expected we found a higher allele frequency for the TLR4 major alleles (A and C). However, although TLR4 299 and 399 are tightly linked, TLR4 299 had a slightly lower frequency of homozygous major allele (AA) than TLR 399 homozygous major allele (CC). TLR4 SNP 299 and 399 have been linked to alteration in the innate immune response and correlated to preterm birth. Published population data that correct for the factor of prematurity are few, making our work a beginning in filling in this gap in population information.

3288T

Using Time to Most Recent Common Ancestor to detect selection in population samples of whole-genome sequences. *H. Hunter-Zinck[1], A.G. Clark[2].* 1) Department of Biological Statistics and Computational Biology, Cornell University, Ithaca, NY; 2) Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY.

Previous studies have investigated the use of Identity-by-Descent (IBD) measures to detect the past action of positive selection (Albrechtsen, et al. 2010, PMID 20592267). However, these studies used IBD inference methods that define IBD as a binary statistic and apply only to genotype data. Inferring Time to Most Recent Common Ancestor (TMRCA) in whole-genome sequence provides a richer metric for the characterization of shared ancestry and, potentially, greater power for applications of IBD in population genetics. We used the pairwise sequentially Markovian coalescent (PSMC) model, developed by Li and Durbin (2012, PMID 21753753), to infer TMRCA for chromosomes pairs both within and between individuals in the 1000 genomes data (1000 Genomes Project Consortium, 2010). Using the distribution of TMRCA at each genomic locus, we then investigated the presence of natural selection genome-wide. A recent complete sweep would result in an unusually recent TMRCA, while a partial sweep may appear as a strongly skewed or even bimodal TMRCA distribution. Balancing selection may manifest as an exceptionally deep TMRCA. We also compared the distribution of pairwise TMRCA for chromosomes within individuals with the distribution of TMRCA for chromosomes between individuals at respective loci. Differences between these distributions may arise due to departures from random mating or due to selection, and these cases may be distinguished when nonrandom mating results in genome-wide effects. Several loci show significant differences in the TMRCA distributions for chromosome pairs within and between individuals. Inferring TMRCA in sequence data from population samples provides a powerful approach to infer forces that impact the ancestral history of the sample.

3289F

Estimating inbreeding coefficients from NGS data: impact on genotype calling and allele frequency estimation. *F.G. Vieira[1], M. Fumagalli[1], A. Albrechtsen[2], R. Nielsen[1,2].* 1) Integrative Biology, UC Berkeley, Berkeley, CA, USA; 2) Department of Biology, University of Copenhagen, Copenhagen, Denmark.

Next-generation sequencing (NGS) methods provide fast, cheap and reliable large-scale DNA sequencing data. They are used in de-novo sequencing, disease mapping, gene expression and in population genetic studies, providing rapid and complete sequencing of candidate genes, exomes, transcriptomes or even whole genomes. Current NGS technologies produce short read sequences that are later de-novo assembled or mapped to a reference genome and used for SNP and genotype calling. However, this data typically has high error rates due to multiple factors from random sampling of homologous base pairs in heterozygotes to sequencing or alignment errors. Furthermore, many NGS studies rely on low coverage sequence data (<5x), making SNP and genotype calling difficult, and often associated with considerable uncertainty. It is, therefore, crucial to quantify and account for this uncertainty, as it will influence downstream analyses. One way of reducing this uncertainty is to sequence target regions deeply (>20x)but,the ever-increasing demand for larger samples suggests that medium- (5–20x) or low-coverage sequencing will be the most common and cost-effective study design in many applications of NGS data for years to come. Recent genotype calling methods rely on probabilistic frameworks to accurately call SNPs and genotypes, even at lower depths.These methods integrate base quality scores with other error sources to calculate an overall genotype likelihood which, in conjunction with a genotype prior, can be used to calculate a posterior probability. This way it is possible to get higher accuracy levels, an associated measure of statistical uncertainty and a natural framework for incorporating prior information. Various prior sources can be used but one of the most common is based on genotype frequencies under Hardy-Weinberg Equilibrium (HWE). However, HWE assumes random mating and, while this assumption might hold for most species, it is clearly violated in others like self-polinating plants, domesticated species (inbreeding) as well as species with asexual life cycles. This violation can result in the under-calling of homozygotes, biasing downstream analysis. Here, we've developed an algorithm to estimate HWE deviations from NGS data. These estimates can then be incorporated into the prior to get an improved genotype posterior probability. We demonstrate the accuracy of our method using simulations and report the increased accuracy in genotype calling and SFS estimation.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3290W

Maternal genetic variation near *PRKAA1* and *EDNRA* is associated with birth weight among residents of high altitude. *A.W. Bigham[1], M.J. Wilson[2], V.A. Browne[3], C.G. Julian[2], E. Vargas[4], C. Rodriquez[5], M.D. Shriver[6], L.G. Moore[7].* 1) Department of Anthropology, University of Michigan, Ann Arbor, MI; 2) Altitude Research Center, University of Colorado Denver, Aurora, CO; 3) Department of Health and Behavioral Science, University of Colorado Denver, Denver, CO; 4) Instituto Boliviano de Biología de Altura, La Paz, Bolivia; 5) Clinica Siraní, Santa Cruz, Bolivia; 6) Department of Anthropology, Pennsylvania State University, State College, PA; 7) Graduate School of Arts and Sciences, Wake Forest University, Winston-Salem, NC.

The hypoxic environment of high altitude restricts fetal growth. However, long-term residents of high altitude are protected from altitude-associated intrauterine growth restriction (IUGR) compared to recent migrants. To identify the genes protecting against IUGR and therefore affecting infant birth weight, we performed a quantitative association study of birth weight in a sample of 178 self-identified Andean and European women living at low and high altitude in Bolivia who were longitudinally studied during pregnancy. We tested 63 SNPs from 16 candidate gene regions with previously identified evidence of natural selection in Andeans. Significant associations between birth weight and maternal SNPs near endothelin receptor type A (*EDNRA*) and protein kinase, AMP-activated, alpha 1 catalytic subunit (*PRKAA1* also known as [aka] *AMPKa1*) were identified. Our results suggest that maternal variation in these genes may influence maternal physiological responses to pregnancy that are instrumental in determining fetal growth and indicate that altitude-selected genes in the hypoxia inducible transcription factor (HIF) pathway play a role in the determination of birth weight at high altitude.

3291T

Selective turnover in human regulatory regions: out with the old, in with the new. *L.D. Ward[1,2], M. Kellis[1,2], The ENCODE Project Consortium.* 1) CSAIL, Massachusetts Institute of Technology, Cambridge, MA; 2) Broad Institute of MIT and Harvard, Cambridge, MA.

The ENCODE project is revealing active transcription and chromatin marks spanning ~80% of the human genome, while comparative genomics of multiple mammals suggests that only ~5% of the human genome is constrained across those species. Analyses of the pilot ENCODE data were statistically underpowered to quantify human constraint for active unconserved elements, leaving open the question of whether they are functional and associated with negative selection in the human lineage. Here, we revisit this question using three new resources: (1) a much higher-resolution map of mammalian constraint using 29 species which pinpoints specific constrained elements covering 4.5% of the human genome; (2) genome-wide experimental coverage from the scaled-up ENCODE project enabling a nearly comprehensive annotation of functional elements; and (3) variation data from the 1000 Genomes Project, enabling us to estimate constraint within the human lineage. We also developed a bootstrapping method to detect human constraint over genomic intervals which accounts for the regional patterns in diversity caused by variable mutation rate and LD. We find strong evidence of human-specific selective constraint in active elements that are not constrained across mammalian genomes. This spans a broad range of unconserved noncoding elements, including: TF binding sites, accessible chromatin regions detected by DNase and FAIRE, transcribed regions outside previously-annotated genes, and candidate promoter and enhancer regions showing specific combinations of histone marks. On aggregate, these show a 16% reduction in diversity in active regions compared to inactive regions, despite not showing constraint across mammals. We confirm that this signal of negative selection is due to recently-gained selection acting in the human lineage: it is also reflected in a depressed frequency of derived alleles, and is robust to controls for background selection, CpG content, biased gene conversion, and read mapping bias. Conversely, we find a recent loss of purifying selection in the human lineage for those elements that are conserved in mammals but lack evidence of activity in ENCODE datasets. Lastly, we study the likely functional implication of recent human-specific selection. We find that enhancers that lack mammalian constraint but show the strongest evidence of human-specific selection are associated with color vision and nervous system development genes, consistent with recent human evolution.

3292F

Signs of high populational differentiation on the MHC (Major Histocompatibility Complex) region on Native South American populations. *K. Nunes[1], E.J.M. Santos[2], J.F. Guerreiro[2], D. Meyer[1].* 1) Universidade de Sao Paulo, Sao Paulo, Brazil; 2) Universidade Federal do Para, Belem, Brazil.

The MHC (Major Histocompatibility Complex) is a 4Mb region of human chromosome 6 (6p21.31) in which is the highest concentration of immune genes. Many of these genes, especially HLA (Human Leukocyte Antigens), present a pattern of genetic variation in their coding regions and/or regulators consistent with that expected for regions that are (or have been) on balancing selection. The balancing selection is characterized by maintaining highest levels of population variability than expected under neutrality. In the case of HLA genes, this pattern is explained by the co-evolution between hosts and pathogens. Also, there is a positive correlation between the abundance of pathogens in a geographic region and the diversity of HLA genes. Based on this, the American continent is an interesting region of the world for studies of selection in the HLA genes. America was the last continent peopled by modern human and has a large variation on latitude, biomes, climates and pathogens, many of which are endemic and could have exercised selective pressure in the MHC genes. In the present study we are looking for signs of selection in the region of the MHC genes in Native populations of South America. We genotyped 16 STRs in the MHC region and compared the genetic profile with 61 STRs spread over genome (demographic control) in 11 Amazon Native populations (Arara do Laranjal, Arara do Iriri, Araweté, Assurini, Awa-Guajá, Kayapó Krokaimoro, Xikrin Kayapó, Parakanã, Urubu-Kapoor, Tiryó and Zoe). We found that the population differentiation (FST) of MHC STRs is greater than the STRs throughout the genome (FST MHC = 0.16 and FST Genomic = 0.11). Based on the finite islands model and exploring the relationship between population differentiation (FST) and the heterozygosity rate, implemented in the LOSITAN program, we found that the 7 of the 16 MHC STRs are outside the null distribution obtained by 61 genomic STRs (C.I. 95%). The highest differentiated MHC STRs are located near the HLA-B gene (D6S2926, D6S2927, D6S2928 e MICA) and DQB1 gene (D6S2876). This result suggest that there are signs of selection in the HLA genes of the Native American populations and indicates that selection favors different sets of alleles in different populations, resulting in a high genetic diversity and differentiation pattern among the populations.

3293W

High diversity of *ADH1B* gene among the Tibetans. *L. Kang[1,2], L. Yan[2], K. Hu[1], F. Chen[1], H. Li[1,2].* 1) School of Medicine, Tibet University for Nationalities, Xianyang, Shaanxi, China; 2) MOE Key Laboratory of Contemporary Anthropology, Fudan University, Shanghai, China.

ADH1B is one of the most studied human genes with many polymorphic sites. One of the single nucleotide polymorphisms (SNP), rs1229984, coding for a variant of Arg47His, has been proved to be associated to many serious diseases including alcoholism and cancers of digestive system. The derived allele, ADH1B*47His, reaches high frequency only in East Asia and Southwest Asia, and is highly associated to agriculture. Micro-evolutionary study defined seven haplogroups for ADH1B. Haplogroups H5, H6, and H7 contain the ADH1B*47His allele. H5 is in Southwest Asia and the other two are found in East Asia. H7 has an additional derived allele of rs3811801 than H6 and was proved to have undergone significant positive selection in Han Chinese, Hmong, Koreans, Japanese, Khazak, Mongols, etc. In this study, we typed 19 SNPs in region covering ADH1B gene among 1175 individuals of 12 Tibetan populations from all districts of the Tibet Autonomous Region. The major populations of the Tibetans (Qamdo, Lhasa, Nagqu, Nyingchi, Shannan, and Shigatse), most of whom are farmers, have around 12% of H7 and 2% of H6. The minor populations, living on hunting or switched to farming recently, have lower H7 frequencies (Tingri 9%, Gongbo 8%, Monba and Sherpa 6%). Luoba (2%) and Deng (0%) have even lower frequencies. Long range haplotype analyses revealed very weak positive selection signals for H7 among the Tibetans. Interestingly, the haplotype diversity of H7 is the highest in the Tibetans than any other populations studied, indicating a longest diversification history of the haplogroup in the Tibetans. Network analysis on the long range haplotypes revealed that H7 in the Han Chinese did not come from the Tibetans but from a common ancestor of the two populations. In conclusion, we argued that H7 of ADH1B originated in the ancestor of Sino-Tibetan populations and flowed to the Tibetans very early. However, as the Tibetans have only a relative short history of agriculture, selection has only laid weak effects, and the frequency of H7 has not risen to high, whereas the diversity of the gene has accumulated to very high level.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3294T

Natural selection increases mutational robustness in complex diseases: Mendelian evidence from early versus late onset common diseases. *B.E. Baysal.* Pathology, Roswell Park Cancer Institute, Buffalo, NY.

Complex diseases are common phenotypic variations that are influenced by genetic factors. Therefore complex diseases should be subject to the process of natural selection, the defining principle of evolutionary change. However, the empirical impact of natural selection on the genetic architecture of complex diseases is poorly understood. It is fundamentally important that we better understand whether and how natural selection affects the genetics of complex diseases. Here, I hypothesize that negative selection of diseased individuals leads to systemic genetic differences between common diseases that primarily occur before or during the reproductive years (early onset) and those that occur after the reproductive years (late onset).

To test this hypothesis, a comprehensive literature survey of nonpleiotropic, nonsyndromic susceptibility genes with a penetrance of at least 80% (Mendelian phenocopies) was completed for early (average age at onset less than 30 years) versus late onset (average age at onset 30 years or more) common diseases, organized using the World Health Organization (WHO) ICD-10 disease classification scheme. Mendelian phenocopies were identified for 16 primarily late onset common diseases from 9 distinct WHO categories including papillary renal carcinoma, obesity, Alzheimer disease, Parkinson disease, frontotemporal dementia, amyotrophic lateral sclerosis, primary open angle glaucoma, age-related hearing loss, coronary artery disease, small vessel stroke, pancreatitis, thrombotic thrombocytopenic purpura, systemic lupus erythematosus, Paget's disease of bone, inclusion body myositis and focal segmental glomerulosclerosis. In contrast, no Mendelian phenocopy was found for any primarily early onset common disease ($p<5.8\times10^{-4}$; Fisher's exact test; two-sided). Thus, rare coding variants are highly predictive for a subset of late onset common diseases, but not for early onset common diseases.

These results suggest that the genetic architecture of early onset common diseases is more robust against the phenotypic expression of highly penetrant predisposing mutations than is the case for late onset common diseases, implying that clinically significant rare variants are less likely to be discovered for early onset common diseases. The primary candidate for increased genetic robustness in early onset common diseases is proposed to be natural selection.

3295F

Evolutionary role of the human skull: comparing neutral and selective markers. *D.V. Bernardo[1], T.F. Almeida[2].* 1) Department of Genetic and Evolutionary Biology, University of São Paulo, São Paulo, São Paulo, Brazil; 2) Unity of Genetics. Children's Institute. HCFMUSP, São Paulo, Brazil.

The microevolutionary processes involved in the human skull evolution are researched in many different ways. Recent studies compared the biological distances obtained for human skulls and microsatellites, attempting to test the correlation between them. Most of them found a positive and significative correlation between morphology and neutral molecular markers and concluded that the evolution of human skull as a whole followed stochastic events such as genetic drift and gene flow. Some found a correlation with climatic events and explained the morphological patterns based on this. Following this line of thought we compared the biological distances of genes known to be under positive selection with microsatellites and with craniometrics to investigate if the morphology preferably obeys a neutral pattern, or if it is associated with selective genes. To accomplish the test we used ten populations representative of Africa, Europe, Asia and Oceania in which molecular and craniometric data were available for the same populations. We constructed distances matrices for all the variables and then calculated the correlation and significancy between them. The results obtained were significative and positive between morphology and microsatellite between morphology and two selective genes and also between three selective genes and microsatellites. They showed that it is possible to obtain positive correlations between features under selection and neutral molecular markers possibly because of the shared demographic history and that this correlation is not always synonymous of an exclusively neutral pattern of evolution. Besides, the EDAR, described previously related to the thickness of hair in Asians, had a significant correlation with morphology and not with microsatellite, suggesting a non-neutral configuration of the skull. Just as this gene, the human skull suffered a possible selective pressure in the asian populations (appearing as the mongoloid morphology), this can be due either by climate or by other environmental pressures that separated these groups in a recent human evolutionary history. This findings favors the interpretation that selective forces are more important for the evolution of human skulls, than previously assumed, because the genetic neutral markers are not exclusively related to the neutral features and a known selective gene is correlated only with morphology, excluding the background role that genetic drift might have in their evolution.

3296W

Characterization of human-specific duplicated neural genes suggests a model of functional antagonism. *M.Y. Dennis[1], F. Antonacci[1], C. Golzio[2], J. Huddleston[1], T.A. Graves[3], P.H. Sudmant[1], D.W. Raible[4], N. Katsanis[2], R.K. Wilson[3], E.E. Eichler[1,5].* 1) Department of Genome Sciences, University of Washington School of Medicine, Seattle, WA, USA; 2) The Center for Human Disease Modeling, Duke University, Durham, NC, USA; 3) The Genome Institute at Washington University School of Medicine, Washington University School of Medicine, Saint Louis, MO, USA; 4) Department of Biological Structure, University of Washington, Seattle, WA, USA; 5) Howard Hughes Medical Institute, Seattle, WA, USA.

Gene duplication and divergence is an important source for phenotypic change and adaptive evolution. We previously identified 23 genes that are duplicated exclusively in the human lineage, many of which have been implicated in neuronal function and brain development. Using a haploid clone resource to resolve the duplicate loci that are missing or misassembled within the human reference genome, we are systematically cloning and sequencing these recent duplicate genes in order to understand their evolution and potential role in human-specific neurocognitive traits. Our previous analysis of *SRGAP2* suggests that the duplicate genes may function to antagonize or modulate the function of an ancestral gene. Here, we assess this paradigm in two other gene families, *ARHGAP11A* and *CHRNA7*, implicated in cortex development and synaptogenesis as well as autism, schizophrenia, and epilepsy. We show that each gene duplicated once in the last six million years producing truncated paralogs found in nearly all humans today (1000 Genomes Project, N = 661). Specifically, a 41 kbp segment containing the putative promoter and first seven exons of *ARHGAP11A* duplicated to create *ARHGAP11B* ~5.5 million years ago (mya), shortly after the human-chimpanzee split. For *CHRNA7*, a larger region (~150 kbp), containing the terminal six exons of the gene, duplicated and created a novel fusion gene, *CHRFAM7A*, less than 1 mya. We show that all paralogs are expressed in both fetal and adult brain and are predicted to encode proteins. Interestingly, though these genes are fixed in most humans, their locations and potential cis-regulatory machinery are varied due to polymorphic inversions of the 15q13.3 locus. Preliminary functional experiments show that knockdown of *CHRNA7* and *ARHGAP11A* orthologs in zebrafish leads to distinct neuronal defects suggesting that dosage is important during development and a potential role of the duplicate genes as functional antagonists. We are simultaneously screening these recent gene innovations for disruptive mutations among patients with idiopathic intellectual disability and neuropsychiatric disease. Our preliminary data suggest that these and other duplicate genes have contributed to unique aspects of human brain evolution and function.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3297T

A genetic mechanism for Tibetan high-altitude adaptation. *C.D. Huff[1,2], F.R. Lorenzo[3], M. Myllymäki[4], S. Swierczek[3], M.E. Salama[5], G.L. Semenza[6], V. Gordeuk[6], J. Xing[2,7], T.S. Simonson[2,8], L.B. Jorde[2], P. Koivunen[4], J.T. Prchal[2,3].* 1) Department of Epidemiology, The University of Texas M.D. Anderson Cancer Center, Houston, TX; 2) Eccles Institute of Human Genetics, University of Utah School of Medicine, Salt Lake City, UT; 3) Division of Hematology, University of Utah School of Medicine, Salt Lake City, UT; 4) Biocenter Oulu, Department of Medical Biochemistry and Molecular Biology, Oulu Center for Cell-Matrix Research, University of Oulu, Oulu, Finland; 5) Department of Pathology, University of Utah and ARUP Laboratories, Salt Lake City, UT; 6) Vascular Program, Institute for Cell Engineering, Johns Hopkins University School of Medicine, Baltimore, MD; 7) Department of Genetics, Rutgers, the State University of New Jersey, Piscataway, NJ; 8) Department of Genetics, University of California in San Diego, San Diego, CA.

Many physiological responses to hypoxia, including erythropoiesis, metabolism, respiration, angiogenesis, and development, are mediated by hypoxia-inducible factors (HIFs). The prolyl 4-hydroxylase PHD2 (encoded by *EGLN1*) is a negative regulator of HIFs. Several haplotypes have undergone positive selection in Tibetans, including a haplotype at the *EGLN1* locus. Chronic exposure to hypoxia leads to an elevation of hematocrit due to an increased number of erythrocytes (polycythemia). While increased hemoglobin concentration may be considered a beneficial adaptation to hypoxia, excessive erythrocytosis results in high blood viscosity, which impairs tissue blood flow and leads to impaired tissue oxygen delivery. In contrast to non-adapted populations, the majority of Tibetan highlanders maintain hematocrit levels comparable to populations living at sea level. In our previous work, we demonstrated that the advantageous *EGLN1* haplotype is associated with protection against polycythemia in Tibetan highlanders. We now describe two high-frequency Tibetan variants, $PHD2^{D4E}$ and $PHD2^{C127S}$. We observed complete linkage disequilibrium between the advantageous haplotype and the two mutations, each with an observed allele frequency of 0.92 among Tibetans. In closely related Asian populations (Korean, Han Chinese, Japanese, and Mongolian), we observed allele frequencies of 0.3 for $PHD2^{C127S}$ and 0.02 for $PHD2^{D4E}$ (*Fst*=0.76, empirical p-value=6.3×10^{-6}). We estimate that the $PHD2^{D4E}$ mutation originated ~6,000 years ago on a $PHD2^{C127S}$ haplotype. Individually, we observed no measurable phenotype for either variant. However, together, these variants demonstrated a lower K_m value for O_2 than wild-type PHD2, suggesting that they downregulate HIFs more efficiently under hypoxic conditions. This was experimentally verified in native erythroid progenitors homozygous for $PHD2^{D4E,C127S}$ that had decreased proliferation under hypoxic conditions, whereas wild-type progenitors have increased proliferation. Thus, $PHD2^{D4E,C127S}$ interferes with the hypoxic augmentation of erythropoiesis that is normally mediated through HIF activation. Our results suggest that the both the $PHD2^{D4E}$ and $PHD2^{C127S}$ mutations are required to produce a gain of function in PHD2 and that the two mutations were favored by selection as a coadaptation to high-altitude beginning approximately 6,000 years ago in Tibetan ancestors.

3298F

Selection and migration in spatially structured populations. *I. Mathieson[1], G. McVean[1,2].* 1) Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom; 2) Department of Statistics, University of Oxford, Oxford, United Kingdom.

Identifying which regions of a genome are, or have been, under selection is important in many settings. In humans, it can provide clues to human population history and the genetic origin of differential disease risk across populations. In ecology, selection can tell us about species' response to changing environmental pressures. In cancer, identifying selection within a tumor can help understand why different tumors respond differently to intervention, and thus guide treatment. However, in all these cases, the effects of selection can be confounded by structure, demography and spatially varying selection. We demonstrate that incorporating spatial structure is important for modeling selection in realistic settings, and that it is possible to make robust inference about relatively strong selection, with selection coefficients on the order of 0.01 or greater.

First we consider the situation where we have time series data of samples of allele frequencies. Working in a simple lattice model of spatial structure, we develop an estimator for selection coefficients and migration rates in a spatially structured population. By treating the true allele frequency as a hidden variable, we can pose the problem as a hidden Markov model, which we can then solve to obtain maximum likelihood estimates. We illustrate this approach by estimating selection coefficients from classical datasets of frequencies of different morph frequencies for the moth species Panaxia dominula and Biston betularia, collected across England over the 20th century. We estimate that the selection coefficient against the medionigra allele of P. dominula is 0.063, and that against the carbonaria allele of B. betularia varies from 0.06 to 0.12 across England.

Finally, we consider the case, relevant to humans and most cancer studies where, rather than time series information, we have only a single observation. We present a structured coalescent model incorporating recombination. In this framework, we simulate haplotypes with structure and selection. We demonstrate inference in this model using human samples from the POPRES dataset. We also use this model to make qualitative predictions about the effect of changing selection and migration on population variation.

3299W

Evidence of recent positive selection in Africans at known and novel BMI loci. *T. Edwards[1], G.J. Papanicolaou[4], K. North[3], D.R. Velez Edwards[2], The. African American BMI Consort.[3].* 1) Center for Human Genetics Research, Department of Medicine, Vanderbilt University, Nashville, TN; 2) Center for Human Genetics Research, Department of Obstetrics and Gynecology, Vanderbilt University, Nashville, TN; 3) University of North Carolina, Department of Epidemiology, Chapel Hill, NC; 4) Division of Prevention and Population Sciences National Heart, Lung, & Blood Institute, Bethesda, MD.

The obesity epidemic in the US is a major public health concern with notable disparity between Europeans (EA) and persons of recent African ancestry (AA). AAs have higher rates of obesity, higher risk of sequelae, and more acute sequelae of obesity compared with EAs. The genetics of obesity have been studied extensively in EAs and many genes have been discovered, although similar work has not been done for AAs. The interplay of environmental influences and genetics has not been as carefully studied as the direct effects of genetic variation. We participated in a meta-analysis of 39,144 AA adults with GWAS data and imputation from 19 studies, and subsequent analysis in 10,000 additional participants to discover the genetic determinants of body mass index (BMI) in AAs. We evaluated the novel and previously discovered gene regions for evidence of recent positive selection using Treeselect, iHS, XP-EHH, Tajima's D, and Fst in the Human Genome Diversity Panel, HapMap, Perlegen, and the BioVU GWAS data. The region near the novel BMI gene region containing KLHL32 had corroborating evidence of recent selection in Africans from all statistics (iHS = 4.05 [p-value = 0.04], XP-EHH = 1.3, Tajima's D = 2.7, Treeselect p = 4×10-4, Max Fst = 0.30), and has been previously associated with cardiac traits and mitochondrial complex 1 deficiency. The gene ADCY3 also lies in a region of recent selection in Africans (iHS = 5.2 [p-value = 9×10-4], XP-EHH = 1.8, Tajima's D = 2.8, Treeselect p = 2.6×10-3, Max Fst = 0.27), and is associated with obesity in humans and mice. ADCY3 -/- mice also lack olfaction, research has shown that ADCY3 functions in olfaction in higher mammals, and previous olfaction research in humans has described statistically significant differences in olfaction between AAs and EAs. These results suggest that selective pressure at BMI loci may determine in part the phenotypic disparity in BMI between AAs and EAs.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3300T

Characterizing the adaptation to high altitude in Tibetans. *E. Huerta-Sanchez[1], X. Jin[2], B. Peter[1], Y. Liang[2], M. He[2], X. Yi[2], A. Asan[2], Y. Shan[2], P. Ni[2], J. Wang[2], R. Nielsen[1], J. Wang[2].* 1) Statistics and Integrative Biology, UC Berkeley, Berkeley, CA; 2) BGI, Beijing genomics institute, BGI-Shenzhen, China.

High altitude environments have exerted a strong selective pressure on some human populations leading to one of the most convincing examples of human adaptation. One gene (EPAS1) in particular has been identified in multiple genome scans of high-altitude residents and harbors some of the largest frequency differences observed between two closely related populations (the Han and the Tibetans). Re-sequencing the gene, however, reveals a very distinctive pattern of genetic variation. Both haplotype and frequency based summary statistics do not support the case of selection on a de novo mutation. Preliminary results using the within-patterns of genetic variation in the Tibetan population suggest that another model, selection on standing variation is a more likely model, but the probability of selection on standing variation is not large enough (about 0.70) to be convincing. Therefore, we modify a Bayesian computation framework to include between population differences (between Han and Tibetans) to gauge whether both the strong signature of differentiation between populations and the level of within population variation are consistent with a model of selection on standing variation. In addition, we estimate both the timing and the strength of selection.

3301F

Recent positive selection of *HLA-DPB1*04:01* in Japanese population. *M. Kawashima[1], J. Ohashi[2], N. Nishida[3], K. Tokunaga[1].* 1) Human Genetics, Graduate School of Medicine, The Univ. of Tokyo, Bunkyo-ku, Japan; 2) Molecular and Genetic Epidemiology, Graduate School of Comprehensive Human Sciences, University of Tsukuba, Ibaraki, Japan; 3) Research Center for Hepatitis and Immunology, International Medical Center of Japan Konodai Hospital, Ichikawa, Chiba, Japan.

The human leukocyte antigen (HLA) genes exhibit the highest degree of polymorphism in the human genome. This high degree of variation at classical HLA class I and class II loci has been maintained by balancing selection for a long evolutionary time. However, little is known about recent positive selection acting on specific HLA alleles. To detect the signature of recent positive selection, we genotyped six HLA loci, *HLA-A*, *HLA-B*, *HLA-C*, *HLA-DRB1*, *HLA-DQB1*, and *HLA-DPB1* in 418 Japanese subjects, and then assessed the haplotype homozygosity (HH) of each HLA allele. There were 120 HLA alleles across the six loci. Among the 80 HLA alleles with frequencies of more than 1%, one allele, *DPB1*04:01*, which had a frequency of 6.1%, showed exceptionally high HH (0.53). The *DPB1*04:01* allele, which was present in the most common 6-locus HLA haplotype (4.4%), *A*33:03-C*14:03-B*44:03-DRB1*13:02-DQB1*06:04-DPB1*04:01*, seems to have flowed from the Korean peninsula to the Japanese archipelago in the Yayoi period. A theoretical calculation indicated that the strong linkage disequilibrium between *DQB1*06:04* and *DPB1*04:01* observed in Japanese cannot be explained without positive selection favoring *DPB1*04:01*. A computer simulation further revealed that the selection coefficient of *DPB1*04:01* was 0.0365 (95% credible interval 0.0254–0.0550). These results suggest that *DPB1*04:01* has recently undergone strong positive selection in Japanese population.

3302W

Haplotype signatures of negative selection. *D. Ortega Del Vecchyo[1], J. Novembre[1,2].* 1) Interdepartmental Program in Bioinformatics, University of California, Los Angeles, CA; 2) Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA.

In line with the nearly neutral theory, recent sequencing projects have identified a high proportion of functional variants at a low frequency in human populations. If negative selection is acting on those functional variants, the genetic variation surrounding them should be different than that of a neutral allele. That implies that the haplotype diversity patterns around a deleterious allele should be informative about the strength of negative selection. Here we propose a simulation framework to study the neutral genetic variation that surrounds a deleterious allele with a specific frequency and negative selection coefficient. We simulated a set of haplotypes using an importance sampling method and the coalescent model. The simulated haplotypes were used to study the haplotypic diversity patterns expected using a collection of different summary statistics such as iHS. The frequency and the selective coefficient of the deleterious allele impacted the neutral variation of the haplotypes, as assessed by the summary statistics. We explain how these findings will be used to analyze the strength of negative selection by estimating the distribution of selective coefficients in alleles at different frequencies.

3303T

Genome-wide signatures of natural selection in diverse African populations. *L.B. Scheinfeldt[1], S. Soi[1], C. Lambert[2], D. Hu[1], A. Coulibaly[1], H. Hutton[1], C. Elbers[1], W. Ko[1], W. Beggs[1], A. Ranciaro[1], S. Thompson[1], J. Hirbo[1], J. Bodo[3], O. Doumbo[3], M. Ibrahim[4], A. Froment[5], G. Lema[6], T. Nyambo[6], S. Omar[7], C. Wambebe[8], D. Meskel[9], G. Belay[9], S.A. Tishkoff[1,10].* 1) Department of Genetics, Univ Pennsylvania, Philadelphia, PA, USA; 2) Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, USA; 3) Malaria Research and Training Center, University of Bamako, Bamako, Mali; 4) Department of Molecular Biology, Institute of Endemic Diseases, University of Khartoum, 15-13 Khartoum, Sudan; 5) UMR 208, IRD-MNHN, Musée de l'Homme, Paris, France; 6) Department of Biochemistry, Muhimbili University of Health and Allied Sciences, Dar es Salaam, Tanzania; 7) Kenya Medical Research Institute, Center for Biotechnology Research and Development, 54840-00200 Nairobi, Kenya; 8) International Biomedical Research in Africa, Abuja, Nigeria; 9) Department of Biology, Addis Ababa University, Addis Ababa, Ethiopia; 10) Department of Biology, University of Pennsylvania, Philadelphia, PA, USA.

Over the past decade, several genome-wide scans for selection have been developed and applied to worldwide human genetic data; however, these studies did not take into account the extensive continental variation within Africa. Here we interrogate genome-wide data from a geographically and ethno-linguistically diverse set of African population samples for signatures of recent positive selection. Our study includes 589 individuals from 22 populations living in distinct environmental conditions, encompassing regions with a range of disease exposures and regions at a range of altitudes. In addition, our study includes populations that practice a wide range of subsistence, such as agriculture, pastoralism, fishing, and hunting and gathering. We collected over 1 million SNP genotypes with the Illumina 1M Duo array and utilized genome-wide scans of positive selection, e.g. F_{ST}, iHS, and XP-CLR. Our results demonstrate several strong signals of recent, regionally restricted positive selection in genes involved in immune response, insulin resistance, and fertility, many of which, have not been highlighted in previous genome-wide scans of selection.

3304F

Contrasting selective forces shaping the type-II C-type lectin receptor family. *H. Quach, S. Fornarino, G. Laval, L. Quintana-Murci.* Human Evolutionary Genetics, Institut Pasteur, Paris, France.

The C-type lectin receptors (CLRs), which are crucial proteins in immunity, recognize carbohydrate structures of self and non-self origin. By binding to endogenous carbohydrate moieties, CLRs support specific functions that contribute to immune homeostasis, such as cell-cell or cell-matrix adhesion. By detecting pathogen associated molecular patterns, they are responsible of antigen recognition and internalization for presentation on MHC molecules. Evolutionary studies have already showed that some human CLRs have been subject to selective forces following host-pathogen co-adaptation. Here, we investigated how natural selection has targeted the type- II CLR family, by resequencing the 15 genes of this family in a panel of 186 individuals from various human populations worldwide. Our results revealed that natural selection has influenced the patterns of diversity of these receptors to different extents, going from strong selective constraints to high levels of amino-acid variation. As an instance, we found that the two genes encoding the Ashwell receptor (ASGR1 and ASGR2), which is functionally active in the liver as a hetero-oligomer, have been both subject to strong purifying selection, preventing the accumulation of amino-acid changes. This is consistent with an essential role of the Ashwell receptor in host survival. In mice, for example, it has been shown to be critic during Streptococcus pneumonia sepsis. By contrast, Langerin (CD207) presents an opposite evolutionary pattern. This receptor recognizes HIV-1 gp120 and structures specific to Candida albicans and Mycobacterium leprae, thus assuring a key protection against a broad range of pathogens.We detected an excess of non-synonymous polymorphisms at Langerin in humans, with respect to divergence, suggesting that some form of diversifying selection has been targeting the protein, making it prone to adapt to environmental changes. Our studies aim to deepen our understanding of the evolutionary relevance of the different type-II CLRs in host defense, to gain knowledge on how variation at these genes may be involved in differential susceptibility to human infectious diseases.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3305W

Balancing selection in the human genome. *M. DeGiorgio[1], K.E. Lohmueller[1], R. Nielsen[1,2,3,4].* 1) Department of Integrative Biology, University of California, Berkeley, CA; 2) Department of Statistics, University of California, Berkeley, CA; 3) Department of Biology, University of Copenhagen, Copenhagen, Denmark; 4) Beijing Genomics Institute, Shenzhen, China.

Balancing selection is an evolutionary process that maintains variability within a population. Two major mechanisms of balancing selection are heterozygote advantage and frequency-dependent selection. A common characteristic of balancing selection is the maintenance of polymorphism for extended periods of time, which is reflected in deep coalescence histories at linked neutral sites. Previous scans for balancing selection have focused on the use of summary statistics, such as the Hudson-Kreitman-Aguade (HKA) test and Tajima's D. Here, we develop two coalescent-based composite likelihood ratio methods that utilize between-species substitution and within-species polymorphism data to detect balancing selection. Our methods take advantage of the effect of balancing selection on genealogies at linked neutral loci and use background patterns of substitution and polymorphism to adjust for demographic history. Through simulations, we show that both methods outperform the HKA test and Tajima's D and are robust to demographic history. Further, we apply our methods to whole-genome sequences from human individuals, using chimpanzee as an outgroup. We find strong support for balancing selection on a number of genes across the genome, including genes previously hypothesized to be under balancing selection, such as the HLA loci and PRIM2, validating our method. We also identified several novel candidate genes. In particular, through gene ontology analysis, we find enrichment for genes that are involved in immunity. Our two methods have been implemented in C and incorporated into the existing software package SweepFinder.

3306T

Reproduction and immunity driven natural selection in the hominid *WFDC* locus. *Z. Ferreira[1, 2, 3], S. Seixas[2], A. Andres[4], W. Kretzschmar[5], J. Mullikin[1, 6], W. Swanson[7], M.K. Gonder[8], S. Tishkoff[9], A. Stone[10], A.G. Clark[11], E. Green[1, 6], B. Hurle[1], NIH Intramural Sequencing Center, Bethesda, MD.* 1) NHGRI, NIH, Bethesda, MD; 2) IPATIMUP, Porto, Portugal; 3) Department of Zoology and Anthropology, Faculty of Sciences, University of Porto, Porto, Portugal; 4) Genetic Diversity and Selection, Department of Evolutionary Genetics - Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany; 5) Genomic Medicine and Statistics, Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom; 6) NIH Intramural Sequencing Center, Bethesda, MD; 7) University of Washington Seattle, WA; 8) Department of Biological Sciences University at Albany, State University of New York, Albany NY; 9) Departments of Genetics and Biology, University of Pennsylvania, Philadelphia, PA; 10) Department of Anthropology, Arizona State University, Tempe, AZ; 11) Department of Biology of Molecular Biology and Genetics, Cornell University, Ithaca, NY.

The *WFDC* and *SEMG* genes warrant a concerted evolutionary analysis because of their strikingly high K_A/K_S, indicative of response to adaptive pressures during vertebrate evolution. *WFDC* is the whey acidic protein (WAP) four-disulfide core domain locus located on human chr20q13. This locus spans 19 genes with WAP and/or Kunitz domains. These genes participate in antimicrobial, immune, fertility and tissue homoeostasis activities. WFDC-related genes include nearby genes encoding seminal proteins Semenogelin 1 and 2 (SEMG1 and SEMG2). To understand better the selection pressures acting on *WFDC* genes in human populations, we sequenced 17 genes and 54 non-coding tags in 71 European (CEU), African (YRI) and Asian (CHB+JPT) individuals. We identified 484 Single Nucleotide Polymorphisms (SNPs), including 65 coding mutations, of which 49 were non-synonymous substitutions. Using classic neutrality tests we confirmed a signature of short-term balancing selection on *WFDC8* in Europeans; and a signature of positive selection spanning genes *PI3*, *SLPI*, *SEMG1* and *SEMG2*. Associated with the latter signal, we identified an unusually homogeneous derived haplotype with a frequency of 88% in Asians. A putative candidate variant targeted by selection is Thr56Ser in SEMG1, which may alter the proteolytic profile of SEMGI and antimicrobial activities of semen. Ancestral Thr56 was traced to the last common ancestor of SEMG1 in Old World Monkeys and hominoids ~25 million years ago (MYA) - with derived allele Ser56 arising less than 0.25 MYA. These results are all consistent with patterns of variation seen in the 1000 genomes data as well. To examine if the adaptive pressures and patterns of genetic variation within the *WFDC* locus differ among hominoids, we sequenced 68 chimpanzees (*P. t. troglodytes*, *P. t. verus* and *P. t. ellioti*), the most closely related species with a different mating system and exposure to pathogens. For that, we generated ~13 Mb of high-quality sequence data, identifying 847 SNPs, and we calculated summary statistics for each gene. Both the chimpanzee data and the human-chimp joint analysis indicate selective signals in the *EPPIN* and *WFDC6* genes, which have been shown to have active roles in primate fertility and immune response. This study provides further evidence that the *WFDC*s and *SEMG*s have been under strong adaptive pressures within hominid evolution, improving our knowledge of biological dynamics of rapidly evolving genomic regions in primates.

3307F

Identification of Hypoxia-Tolerance Regions in High-Altitude Populations. *R. Ronen[1], N. Udpa[1], D. Zhou[2], T. Stobdan[2], O. Appenzeller[3], K. Frazer[1,2], J. Liang[4], Y. Li[4], V. Bafna[1,5], G. Haddad[2].* 1) Bioinformatics and Systems Biology, University of California at San Diego, La Jolla, CA; 2) Department of Pediatrics, University of California at San Diego, La Jolla, CA; 3) New Mexico Health Enhancement and Marathon Clinics Research Foundation, Albuquerque, New Mexico, United States of America; 4) BGI-Shenzhen, Shenzhen 518083, China; 5) Department of Computer Science and Engineering, University of California at San Diego, La Jolla, CA.

The body's response to hypoxic stress is known to play a crucial role in many cardiovascular, neonatal, and oncological conditions (Zhou et al., PLoS Genetics, 2007). However, the mechanisms underlying the basic pathobiology are not well understood. One way to determine the genetic elements of this response is to identify regions undergoing long-term selective sweeps in (healthy) individuals living at high-altitude, low-oxygen environments. There are three major high-altitude environments in which individuals are known to reside permanently - the Ethiopian highlands, the Himalayan plateau, and the Andes mountains (Beall, Integr. Comp. Biol., 2006). Previous genetic studies of these populations have focused on either sparse genotyping of a number of locations across the genome or sequencing of just the protein-coding exome (Yi et al., Science, 2010; Bigham et al., PLoS Genetics, 2010, Scheinfeldt et al., Genome Biology, 2012). Both of these designs only sample a subset of the genome, and thus, can potentially bias the data and result in a loss of statistical power. In this work, we performed high (15–43x) coverage, whole genome sequencing of 37 individuals from the Ethiopian (~3500m altitude) and Andean (~4300m altitude) highlands. To gauge the impact of time under selection, we sampled two different populations from Ethiopia - one has been in the highlands for thousands of years, whereas the other is a recent (~700 years) migrant to the area. In contrast to the Ethiopians, the Andean highlanders display a variety of physiological responses to hypoxia, ranging from perfect fitness to a condition known as chronic mountain sickness. We sequenced Andean individuals from the extremes of this distribution. In all these populations, we identify a number of regions based on changes in the SNP frequency spectrum expected under a model of classical positive selection. In this model, beneficial mutations not only increase in frequency, but also raise the frequencies of “hitchhiking” alleles on the same initial haplotype. We prioritize regions that have sharp SNP frequency differences in our adapted populations, compared to the individuals with chronic mountain sickness and multiple lowlander populations. We then select the strongest gene candidates and characterize differences and similarities in the response mechanisms between populations. Finally, using model organism systems, we show that these genes are important for survival in low oxygen environments.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3308W

Positive selection on Base Excision Repair Pathway in Yoruba and British Populations. *M. Wang[1], R. Li[1], Y. He[1], L. Jin[1,2].* 1) CAS-MPG Partner Institute for Computational Biology, Shanghai, Shanghai, China; 2) State Key Laboratory of Genetic Engineering and Ministry of Education Key Laboratory of Contemporary Anthropology, School of Life Sciences and Institutes of Biomedical Sciences, Fudan University, Shanghai, China.

Base Excision Repair (BER) pathway is the primary repair system involved in the removal of endogenous DNA base damage throughout the cell cycle that could otherwise cause mutations by mispairing or lead to breaks in DNA during replication. Therefore, it is crucial to understand evolutionary role of the BER system when genetics and biochemistry of this system have being clarified. Using the last full-genome sequencing data from the 1000 genome project, we examined 152 genes of the BER pathway to identify potential genetic signature of positive selection in Yoruba (YRI) and British (GBR) populations. For each gene, all genetic variants located in the gene and its' surrounding 200Kb genome region were examined by inspecting pattern of linkage disequilibrium (LD) between and within populations using XP-EHH and iHS methods. A proposed summary statistic was applied on initial results to combine p-values of each 50 neighboring SNPs. All summary statistics with corrected p-value less than 0.01 were declared to be significant after Bonferroni corrections. Nine genes that might experience positive selection were discovered in between-population comparison using XP-EHH. Meanwhile, eight and six genes were selected as candidates within YRI or GBR population using iHS, respectively. PTEN (correct p-value < 4.24E-22) and ATMIN (correct p-value < 4.26E-20) gene discovered by XP-EHH method were also identified by iHS method in YRI population. The sharing of selected genes indicated that the two genes might experience positive selection in the YRI population. No selected gene was shared by the between-population analysis and within-population investigation for GBR population. However, the results for seven out of the nine selected genes were still reliable because power of the XP-EHH method was believed to be superior to that of the iHS method in many genetic scenarios. XPC gene was highlighted by within-population investigations in both YRI and GBR populations with corrected p-value less than 1.28E-23 and 6.88E-17, respectively. The sharing results for XPC gene suggested positive selection might act on this gene both before and after the Out-of-Africa. Our discovery about positive selection on BER pathway suggested evolutionary activities of the critical pathway were still promptly going on in human populations. The discovery supported the hypothesis that biological mutation rate was largely determined by mutation-selection balance.

3309T

Expression quantitative trait loci (eQTLs) are adaptive to multiple environmental factors. *K. Ye, Z. Gu.* Division of Nutritional Sciences, Cornell University, Ithaca, NY, 14850.

The relative importance of regulatory SNPs and coding SNPs in adaptation to local environment has been controversial. eQTLs have been suggested to be targets of natural selection and to be enriched in trait-associated SNPs. However, whether eQTLs are adaptive to different types of environmental factors and its relative importance compared with non-synonymous SNPs have not been systemically explored. Recently a Bayesian linear model method was developed to identify local adaptation signals by detecting a correlation between the frequency of an allele and an environmental factor. This method has been applied in three genome-wide analyses involving 42 environmental variables, for each of which more than 500,000 SNPs were tested for evidence of correlation. Utilizing these data, we performed a systemic examination of the adaptive patterns of eQTLs to different environmental factors. Compared with intergenic neutral SNPs, eQTLs tend to be enriched in adaptive signals to the majority of 42 environmental factors. The significant enrichment is independent of recombination rate, although recombination rate is reversely correlated with the degree of enrichment. The enrichment of eQTLs in adaptive signals is due to their functional effect as supported by the discovery that groups of eQTLs with higher true positive rate tend to have higher degree of enrichment. Over all 42 environmental factors, the enrichment of eQTLs in adaptive signals is similar to that of non-synonymous SNPs, while both of them are higher than that of genic SNPs. At the individual level, some environmental factors, including climate, latitude, and virus diversity, have an enrichment of eQTLs in adaptive signals significantly higher than that of non-synonymous SNPs. Further pathway enrichment analysis identified a number of pathways enriched with adaptive signals from eQTLs but not from other genic SNPs. These pathways are mostly related with immune system, cellular signaling and metabolism, such as peroxisomal lipid metabolism with adaptive signals to short wave radiation flux (Summer), long-term depression to climate, and systemic lupus erythematosus to latitude. Our analyses demonstrate that eQTLs are in general adaptive to local environment while for some environmental factors and pathways, eQTLs have an adaptive effect stronger than non-synonymous SNPs.

3310F

Balancing selection on a non-coding region with regulatory function. *Q. Zhu[1,2], O. Gokcumen[1,2], L. Mulder[3], R. Iskow[1,2], C. Austermann[3], M. Steward[3], C. Scharer[4], T. Raj[2,5], J. Boss[4], S. Sunyaev[2,5], A. Price[6,7], B. Stranger[2,5], V. Simon[3,8,9], C. Lee[1,2].* 1) Department of Pathology, Brigham and Women's Hospital, Boston, MA; 2) Harvard medical school, Boston, MA; 3) Department of Microbiology, Mount Sinai School of Medicine, New York, NY; 4) Department of Microbiology, Emory University, GA; 5) Division of Genetics Department of Medicine, Brigham and Women's Hospital, Boston, MA; 6) Department of Epidemiology, Harvard School of Public Health, MA; 7) Department of Biostatistics, Harvard School of Public Health, MA; 8) Global Health and Emerging Pathogens Institute, New York, NY; 9) Division of Infectious Diseases, Department of Medicine, New York, NY.

Most functionally important loci in modern humans, including the majority of exons, are under negative (purifying) selection and consequently show little, if any, genetic variation. In contrast, other forms of selection, such as balancing or directional positive selection, occur less frequently and identifying genomic loci under such selection entails the detection of variants that show unexpectedly high population differentiation or deviation from the prevalent haplotype structure. In this study, we have identified a ~36 kb DNA segment at the upstream of Apobec3 locus that has an ancient substructure, predating Human-Neandertal sequence divergence, and examined SNP data around this locus from 629 individuals originating from 12 human populations. The variation at this DNA segment exists primarily as two highly divergent haplogroups. One of the haplogroups, which includes the Neandertal haplotype, contains a 4.6 kb deletion polymorphism in perfect linkage disequilibrium with 12 SNPs across diverse populations, while the other more ancestral haplogroup does not contain this 4.6 kb deletion. Population genetic analyses of these SNPs within each of these haplogroups revealed high population differentiation and signatures of balancing selection. This DNA segment harbors no protein-coding genes, but the presence of balancing selection, bioinformatic predictions of transcription factor binding, in vitro experiments and expression quantitative trait loci analyses indicate putative regulatory function.

3311W

Characterizing selection signatures associated with high-altitude adaptation in Tibetans. *A. Ferrer Admetlla[1], Y. Wang[1], B. Peter[1], N. Vinckenbosch[1], E. Sanchez-Huerta[1], X. Jin[2], A. Asan[2], J. Wang[2], J. Wang[2], R. Nielsen[1].* 1) Department of Integrative Biology, UC-Berkeley, Berkeley, CA, US; 2) Bejing Genomics Institute, Shenzhen, China.

The oxygen concentration (21% at sea level) decreases with altitude (~13% at 4000m). Some human populations live at very high elevations (>4000 meters), where the reduction of oxygen concentration represents a challenge for the body. Residents of the Tibetan Plateau have adapted to this extreme condition. To date, several studies have looked for the genetic basis of high-altitude adaptation in the Tibetan population and found signatures of selection in several genes, but the time and strength of selection in these regions have been not yet characterized. In this work we gather 23 candidate genes that have been associated with altitude adaptation in Tibetans to determine the selective regime that has acted on them as well as the time and the strength of selection. To this purpose, we have analyzed targeted next -generation sequencing data for these 23 regions in 148 Tibetan samples from 5 villages ranging from 3000 to 4800 meters. We have also correlated the genotypic information of these samples to phenotypic information related to oxygen transportation, such as hematocrit and hemoglobin. Our results indicate a subset of the genes were selected during the same time range and, that this time range overlaps with previous estimates of the colonization of the Tibetan Plateu.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3312T

Detecting population structure and genetic differentiation from high-throughput re-sequencing data. *M. Fumagalli[1], T. Korneliussen[2], T. Linderoth[1], F. Vieira[1], A. Albrechtsen[2], R. Nielsen[1,2].* 1) University of California, Berkeley, CA; 2) University of Copenhagen, Copenhagen.

New, high-throughput, DNA re-sequencing technologies provide a cost-effective means of obtaining large-scale genetic data, and are now a primary resource for population genetics studies. Some drawbacks for being able to fully benefit from this technology, however, is the computational challenge associated with large datasets and high sequencing error rate. Moreover, many re-sequencing studies employ medium to low sequencing coverage. Under such circumstances, SNP and genotype calling can be unreliable and lead to a biased allele frequency distribution. Accurate estimation of the site frequency spectrum is vital to population genetic inference of demography, natural selection, and population subdivision, as well as the statistics used to summarize these processes. Many previously used methods for estimating allele frequencies that are primarily based on direct counting of sequencing reads, lead to inaccurate estimates of local nucleotide diversity indices, and FST, a measure of genetic population differentiation. Recently, more sophisticated algorithms for SNP and genotype calling have been proposed. Most promising are those that use a Bayesian framework to incorporate base quality scores and statistical uncertainty in order to obtain posterior probabilities associated with each genotype. Accordingly, I propose several new strategies for computing estimates of FST from estimation of allele frequencies per site. I will demonstrate the high performance of these novel methods for detecting population subdivision, especially when facing high error rates and low sequencing coverage. I will also discuss how population structure can be investigated with principle component analysis within this probabilistic framework. In addition, I will perform sliding-window analyses on previously published low-coverage re-sequencing data from multiple populations to demonstrate an efficient and reliable strategy for genome-wide scans of natural selection. This analysis will highlight genomic regions displaying exceptionally high genetic differentiation along with reduced levels of nucleotide diversity among pairs of populations. Methods herein presented are powerful and reliable tools for investigating population genetic variation on a large scale directly from high-throughput re-sequencing data.

3313F

Different haplotypes in East Asia and Europe both show positive selection in 1q24. *C. Heffelfinger[1,2], A.J. Pakstis[2], W.C. Speed[2], M.P. Snyder[3], K.K. Kidd[2].* 1) Dept MCDB, Yale University, CT; 2) Genetics Dept, Yale University, CT; 3) Genetics Dept, Stanford Univerity, CA.

We have identified a region of chr1q24 under positive selection in East Asia and Europe, with smaller signatures of selection present in Southwest Asia, Central Asia, and South Central Asia. Tests for selection were conducted using normalized Haplosimilarity (nHS) and Relative Extended Haplotype Homozygosity (REHH) on a set of 140 SNPs from chr1:168653857 to chr1:169640109 (Hg19) in the 52 HGDP populations. Heterozygosity and linkage disequilibrium were assessed using the same panel for all populations using HAPLOT. Amino-acid changing SNPs and major structural variants were also tested in the HapMap and 1000Genomes panels to identify variants in linkage disequilibrium with the selected haplotype. The strongest signatures of selection occur for two different haplotypes. One shows evidence of selection in East Asia where it is present at approximately 65% frequency. The other shows evidence of selection in Europe where it is present at approximately 35% frequency. Most SNP alleles differ between these two haplotypes, but the highest REHH scores for both haplotypes occur near the 7th exon of the NME7 gene. In East Asia this core haplotype is GTA (rs2157597, rs12121994, rs12118611) whereas in Europe it is ATG. While the haplotypes with the highest REHH scores are in NME7, significant REHH and nHS scores extend across a region containing ATP1B1, NME7, BLZF1, C1orf114, and SCL19A2. Possible candidates for selection were also examined. A deletion from ~169229000-39000 was tested for linkage disequilibrium with the core selected haplotype in the HapMap dataset. The deletion was consistently present on the European selected haplotype but absent on the East Asian Haplotype. The derived allele of rs1028180, a missense mutation in BLZF1, has frequencies matching the East Asian haplotype under selection but is virtually absent in Europe. Finally the derived allele of rs3820059, a missense mutation in C1orf114, has frequencies matching the European selected haplotype but is at much lower frequency in East Asia. Funding for this research was provided by NIH grants 5T32GM007223 and 5P50HG00235711.

3314W

Early Origins of Southeast Asian Ovalocytosis during the Austronesian Expansion. *J.A. Wilder, A.M. Paquette, M.K. Thompson.* Biological Sciences, Northern Arizona University, Flagstaff, AZ.

Southeast Asian Ovalocytosis (SAO) is a recessive lethal trait that confers significant protection against malaria-related mortality in heterozygous carriers. SAO is caused by a 27-bp exonic deletion in *SLC4A1*, which encodes the Band 3 protein in red blood cells. Here we describe the first analysis of haplotypic variation among SAO chromosomes from across its range in Southeast Asia. Our survey of 60 SAO-bearing chromosomes reveals substantial short-range haplotype diversity, caused primarily by the action of a nearby recombination hotspot. Previous studies have found that SAO is uniformly associated with the rare nonsynonymous “Memphis” allele at a nearby SNP; the “Memphis” allele is known to cause altered Band 3-mediated anion transport across the erythrocyte membrane. We find that a fraction of SAO chromosomes do not share a haplotype with the “Memphis” allele, suggesting that there may be functional variation among SAO-carriers with respect to erythrocyte properties. We used variability among SAO chromosomes to estimate that the causal mutation occurred at least 3,500 years bp, making it the oldest malaria-resistance mutation found in Southeast Asia to date. Previous studies have noted an association between the SAO mutation and speakers of Austronesian languages. Our result suggests that the SAO mutation likely occurred during the early stages of the Austronesian expansion, and then accompanied this movement of people as they migrated throughout Island Southeast Asia. Despite being a recessive lethal, SAO occurs at elevated frequencies in many malarial populations, with an observed allele frequency as high as 0.176. This is congruent with the action of exceptionally strong balancing selection favoring heterozygotes (*i.e.*, wild-type homozygotes have a fitness as low as 79% that of heterozygotes). Our study suggests that the Austronesian expansion into malarial regions may have been facilitated by the early appearance of SAO and its associated fitness benefits relative to non-SAO carriers.

3315T

Local Genome Diversity Studies in Puerto Rico: uncovering distribution of ancestry proportions and disease allele frequencies across the island. *Y.M. Afanador[1], I. Rivera[2], E.P. Tascón-Peñaranda[2], J.L. Rodriguez-Florez[3], C.A. Winkler[4], A.V. Washington[2], E. Suárez[5], J.C. Martinez-Cruzado[2], T.K. Oleksyk[2], Local Genome Diversity Studies Consortium.* 1) RISE 2 BEST Program University of Puerto Rico, Mayaguez Campus, Mayaguez, PR; 2) Biology, University of Puerto Rico, Mayaguez Campus, Mayaguez, PR; 3) Department of Genetic Medicine, Weill Cornell Medical College, New York NY, USA; 4) Molecular Genetic Epidemiology Studies Section, Frederick National Lab, Frederick, MD; 5) Biology, University of Puerto Rico, Ponce Campus, Ponce, PR.

Describing local population diversity in the context of recent evolution can enhance large-scale studies emerging from the concerted effort of the international scientific community. With whole genomes of many Puerto Rican (PR) individuals publicly available from the 1,000 Genome Project, the ancestral contributions and evolutionary history of the PR population across the island can now be examined in detail to uncover patterns of migration, extent and character of admixture, resistance to the local infectious diseases, and recent adaptations to climate and food. We develop, evaluate and implement a new instructional strategy built into an educational curriculum. Local Genome Diversity Studies (LGDS) involved 42 undergraduate students that carried out sample collection, DNA extraction and genotyping in a series of individual projects within the LGDS cycle. A geographically representative DNA samples from more than 2,000 individuals across 40 out of 78 municipalities of PR have been collected. Using our sample set, we identify patterns of European, African and Amerindian admixture proportions across different parts of the island using a panel of 93 autosomal, 24 maternally inherited, and 11 paternally inherited Ancestry Informative markers (AIMs). This information is used to develop a map of local genome diversity and to uncover the complexity of the islands' rich genetic background. In the second phase of the cycle, called Adopt-a-Gene Project, students use 1,000 Genome data to identify evolutionarily relevant, locally polymorphic genetic variants, related to disease, local environment, or other adaptations. Candidate loci are analyzed in the LGDS samples to document their actual geographic distributions. Preliminary island-wide analysis revealed non-random distribution of some disease-relevant variants (CCR5 Δ32). Polymorphisms involved in kidney disease (e.g. APOL1, MYH9, NPHS2) and other chronic diseases (e.g. obesity, diabetes, and asthma), as well as physical phenotype determinants (EDAR, OCA2, MC1R, ABCC11) are genotyped, while a panel of cancer candidate genes is re-sequenced for the entire sample using next generation technology. Our project is an example of a successful local community outreach and integration of minority undergraduate education into the research process that shows a new way to deepen our knowledge about the local genome diversity of human populations worldwide following the effort led by the 1,000 Genomes Project.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3316F

Indian genome architecture reveals high diversity and suggests ancient African ancestry. *S. Kumar[1], R. Singh[1], N. Banerjee[2], H. Gowda[3], M. Kramer[4], S. Kamalakaran[2], V. Varadan[2], A. Janevski[2], B. Muthusamy[3], B. Chakrabarti[1], SM. Srikanth[3], S. Keerthikumar[3], TS. Keshava Prasad[3], S. Vivekandandan[1], V. Makkapati[1], A. Panigrahi[1], P. Kumar[3], P. van Hooft[5], E. Ghiban[4], L. Gelley[4], S. Muller[4], G. Atwal[4], WR. McCombie[4], A. Pandey[3,6,7], A. Chakravarti[6], N. Dimitrova[2].* 1) Philips Research, Philips Innovation Campus, Bangalore-560045, Karnataka, India; 2) Philips Research NA, 345 Scarborough Road, Briarcliff Manor, NY 10510, USA; 3) Institute of Bioinformatics, International Technology Park, Bangalore 560066, India; 4) Cold Spring Harbor Laboratory, 1 Bungtown Rd, Cold Spring Harbor, NY 11724, USA; 5) Philips Research, eScience Support, HTC36-03, High Tech Campus 36, 5656 AE Eindhoven, The Netherlands; 6) McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore 21205, Maryland, US; 7) Departments of Biological Chemistry, Oncology and Pathology, Johns Hopkins University School of Medicine, Baltimore 21205, Maryland, US.

Present day Indians are thought to contain huge genetic diversity which is attributed to multiple waves of migration and gene flow that took place since the early human dispersal out of Africa. Recent studies on population genomics are starting to map the route of migration of modern humans into Indian subcontinent, however at a limited resolution. We report the diploid genome sequences of two South Indian individuals and co-analyze their genomes with samples from Africa, Europe and Asia. The two Indian individuals exhibit highest nucleotide diversity next only to individuals of African ancestry. At the protein level, diversity varies by 10-fold across proteins and 100-fold across domains at levels that are characteristic of all humans. We identified multiple pathways that were significantly enriched for non-synonymous variants in the Indian individuals suggesting local adaptation. Our analysis of population histories suggests that Indian genomes have had contributions from African ancestry in multiple waves: from 200,000–600,000 years ago and the more recent `Out of Africa' ancestry (100,000 years). These results, consistent with previous marker data, are the first genome-wide evidence of a Southern coastal migration and colonization of India by early humans. Consequently, Indian genome sequences not only contribute new features of human genome variation but suggest that additional sequencing of South Indian samples will be critical for human history, biology and medicine.

3317W

Genetic Evidence of Paleolithic Colonization and Neolithic Expansion of Modern Humans on the Tibetan Plateau. *B. Su[1], X. Qi[1], C. Cui[2], Y. Peng[1], X. Zhang[1,5], Z. Yang[1,5], H. Zhong[1], H. Zhang[1], K. Xiang[1,5], X. Cao[1], Y. Wang[1,5], O. Ouzhuluobu[2], B. Basang[4], C. Ciwangsangbu[4], B. Bianba[2], G. Gonggalanzi[2], T. Wu[3], H. Chen[6], H. Shi[1].* 1) State Key Laboratory of Genetic Resources and Evolution, Kunming Inst Zoology, Chinese Academy Sci, Kunming, China; 2) High Altitude Medical Research Center, School of Medicine, Tibetan University, Lhasa, China; 3) National Key Laboratory of High Altitude Medicine, High Altitude Medical Research Institute, Xining, China; 4) People's Hospital of Dangxiong County, Dangxiong, China; 5) Graduate University of Chinese Academy of Sciences, Beijing, China; 6) Department of Epidemiology and Biostatistics, Harvard School of Public Health, Boston MA, USA.

Sitting 4,000 m above sea level, at the foot of the northern Himalayas of western China, the Tibetan Plateau is a remarkable area for cultural and biological studies of human population histories. However, the chronological profile of the Tibetan Plateau's colonization remains an unsolved question of human prehistory. By studying genetic markers from both paternal (Y-chromosome) and maternal (mitochondrial DNA) lineages of 6,109 individuals from 41 Tibetan populations sampled across the Tibetan Plateau, we found there have been two distinct, major prehistoric migrations of modern humans into the Tibetan Plateau. Among these populations, we identified ancient Tibetan-specific genetic signatures dated to around 30,000 years ago, indicating that the initial peopling of the Tibetan Plateau by modern humans occurred during the Paleolithic rather than Neolithic. We also found evidences for relatively young (only 7–10 thousand years old) shared Y chromosome and mitochondrial DNA haplotypes between Tibetans and Han Chinese, suggesting a second wave of migration during the Neolithic. Collectively, the genetic data indicate that Tibetans have been adapted to high altitude environments since initial colonization of the Tibetan Plateau in the early Upper Paleolithic, before the Last Glacial Maximum, followed by a rapid population expansion triggered by the establishment of farming and yak pastoralism in the early Neolithic.

3318T

Genomic ancestry and structure of the African-European admixed population of the Cape Verde islands. *S. Beleza[1,2], N.A. Johnson[3], S.I. Candille[1], D.M. Absher[4], I.I. Araújo[5], A. Correia e Silva[5], M.D. Shriver[6], J. Rocha[2,7,8], G.S. Barsh[1,4], H. Tang[1].* 1) Dep of Genetics, Stanford University School of Medicine, Stanford, CA., USA; 2) Instituto de Patologia e Imunologia Molecular da Universidade do Porto (IPATIMUP), Porto, Portugal; 3) Department of Statistics, Stanford University, Stanford, California, USA; 4) HudsonAlpha Institute for Biotechnology, Huntsville, Alabama, USA; 5) Universidade de Cabo Verde (Uni-CV), Praia, Santiago, Cabo Verde; 6) Department of Anthropology, The Pennsylvania State University, University Park, Pennsylvania, USA; 7) Centro de Investigação em Biodiversidade e Recursos Genéticos (CIBIO),Vairão, Portugal; 8) Departamento de Biologia, Faculdade de Ciências da Universidade do Porto, Porto, Portugal.

Over the past few years, recognition of the unique utility of admixed populations in studies of complex traits and diseases that differ in risk between populations has increased. However, given the cultural and genetic heterogeneity of admixed groups, it is essential that multiple admixed populations are studied to fully appreciate the relationships among the genetic, historical, and environmental determinants of those traits. Cape Verde is an archipelago of ten islands located 300 miles off the coast of West Africa. The islands were uninhabited until discovery by the Portuguese in the 15th Century and were initially peopled by European immigrants and enslaved Africans. We performed an analysis of the genetic structure and admixture of Cape Verde based on high-density genotype data on 697 individuals living in the six most populated islands. Principal component analysis is reminiscent of the geographic map of Cape Verde. Clustering analysis reveals extensive genetic admixture between European and African ancestral populations (median African ancestry=0.58, sd=0.12), with individual African ancestry proportions varying considerably among the islands. The European ancestors of these individuals genetically resemble Portuguese. The African ancestry of Cape Verdeans is largely derived from West African populations, as opposed to African Americans who derive their ancestry from West-Central African groups. The parsing of the European and African components of the Cape Verdean genomes allows for a more detailed characterization of the genetic diversity and differentiation across the islands. Kinship coefficients corrected for the bias introduced by admixture and IBD analysis indicate that individuals are genetically more related within an island than between islands, as expected, and that the magnitude of relatedness varies across the archipelago, reflecting island-specific demographic histories. The observed genetic heterogeneity can be explained by historical events that have taken place during and since the colonization period. Cape Verde genetic structure is a dramatic example of the complex interplay between demographic events influencing the genetic diversity and structure of human populations: founder and drift effects along with admixture. This work provides a foundation for admixture-based genetic studies in Cape Verde and highlights the great potential for research on the genetic basis of complex traits of this population.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3319F

Mitochondrial DNA in Myanmar: Complete mitochondrial genome sequencing revealed several new lineages within macro-haplogroup M. *J. Horst[1], M. Summerer[2], G. Erhart[2], D. Horst[3], B. Horst[4], S. Sanguansermsri[5], A. Manhart[2], F. Kronenberg[2], A. Kloss-Brandstätter[2].* 1) Institut für Humangenetik, Universität Münster, Germany; 2) Division of Genetic Epidemiology, Innsbruck Medical University, Innsbruck, Austria; 3) Pathologisches Institut, Ludwig-Maximilians-Universität, Münster, Germany; 4) Department of Dermatology, Section of Dermatopathology, Columbia University, New York, NY; 5) Department of Pediatrics, Chiang Mai University, Chiang Mai, Thailand.

Background: Myanmar (Burma) is the largest country in Southeast Asia, with a population of 56 million people subdivided into more than 100 ethnic groups. The Bamar represent the largest group amongst them. Ruled by changing kingdoms and dynasties and lying on the trade route between India and China, Burma was influenced by a variety of cultures. Since Burma's independence after the British occupation, minorities suffer from government's repression and especially the Karen people struggle against the domination of the Bamar culture. We analyzed the mtDNA control region of 327 unrelated donors from Myanmar according to highest quality standards. To refine haplogroup (HG) information, especially in the complex haplogroup M, 44 selected individuals were subjected to complete mitochondrial genome sequencing. Results: The distribution of the macro-haplogroups M and N and the geographic assignment of the haplogroups were typical for Southeast Asia, however, the frequency of individual haplotypes was very specific in Myanmar. In general, the Myanmar sample exhibited pronounced mtDNA diversity, with the ethnic group of the Bamar being the most diverse. The 44 complete mitochondrial genome sequences revealed 10 new so far un-described mtDNA lineages, represented by 15 haplotypes, all lying within macro-haplogroup M. One lineage, comprising 3 haplotypes, clustered within HG G2, two lineages represented subgroups of HG M49 and two lineages with in total 4 haplotypes clustered within the M13'46'61 group. In addition, we found 5 yet undefined basal M lineages. Their exact phylogenetic position can only be defined with deeper knowledge of macro-haplogroup M. Conclusions: The multi-ethnic population and the complex history of Myanmar are well reflected in its distinct mtDNA heterogeneity. In this region with its long history of human settlement, plenty of mitochondrial haplogroups, especially in the complex haplogroup M, await to be newly described.

3320W

Asian Expansion of Modern Human out of Africa. *H. Li.* MOE Key Laboratory of Contemporary Anthropology, School of Life Sciences, Fudan University, Shanghai 200433, China.

Global population of the modern human originated in Africa around 200 thousand years ago, which has been proved by various evidences, especially by genetic evidences. Recent studies revealed another important event during the natural history of modern human, i.e., the Asian Expansion. The Asian Expansion was supported by both linguistic and genetic evidences. In 2011, Atkinson reported a study on the phonemic diversity of world languages, and found an African Origin of the present languages. However, more detailed analyses on the phonemic diversity reflected an Asian Expansion rather than an African Origin. All the languages, except for the Khoisan languages, derived from the expansion from western Asia around Iran. The Asian Expansion was also supported by genetic evidence. Y chromosome haplogroup phylogenetic tree has an African root, but only haplogroups A and B were African aboriginal types, i.e., these two haplogroups have never leaven Africa. All the other haplogroups, all within CF and DE, derived from one ancient type with a marker M168 out of Africa around 50–70 thousand years ago. They expanded from western Asia around 3–4 thousand years ago, and gave birth to all the haplogroups from C to T. Therefore, the haplogroup with highest frequency in Africa, haplogroup E, was also from the Asian Expansion. In Africa, Haplogroup A is mostly in Khoisan and Saharan populations, and haplogroup B is mostly in Pygmies and other populations around Congo. The major populations of Africa, sometimes called Bantu or Niger-Congo people, were then most probably came back from Asia. The Asian Expansion was firstly hypothesized by German biologist Ernst Haeckel in his map of human origin and migration in 1863. Now, we are accumulating more evidences.

3321T

Using Y-Chromosome Haplotypes to Improve Inferred Ancestral Origins in European Populations. *N. Myres[1], P. Underhill[2], R. King[2], J. Byrnes[1], K. Noto[1], S. Woodward[1], N. Angerhofer[1], C. Ball[1], K. Chahine[1].* 1) AncestryDNA, Provo, UT; 2) Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, CA.

Studies analyzing datasets of autosomal single nucleotide polymorphisms (SNPs) have shown that significant amounts of information can be inferred about an individual's ancestral origins. Methods for assigning geographic origins according to continental boundaries can be achieved with high accuracy, but more localized geographic assignments have proven more challenging. Identifying ancestry informative markers (AIMs) has been shown to significantly improve geographic assignments for populations of close proximities, but for certain regions within Europe accuracy rates are highly variable, ranging from 20%–74%. Such variability likely arises from Europe's complex population history as well as limitations of current autosomal systems for assessing the temporal components of underlying SNP variants. In contrast, robust and well resolved haploid phylogenies have been developed which delineate the sequential branching order and coalescent times of genetic variants within global populations. In particular, Y-chromosome diversity within Europe is well characterized and has been used extensively for elucidating ancient and recent contributions to extant European genetic diversity, including both external (intrusive) and local (in situ) elements. Thus, although haploid systems capture only a narrow component of a population's or an individual's ancestry, Y-chromosome analysis offers a relatively uncomplicated framework for disentangling complex patterns of gene flow. Herein, we assess the extent to which augmenting autosomal population data with Y-chromosome information improves the accuracy of ancestral origin assignments in regions of Europe. Ancestral origins are inferred from autosomal SNPs for a dataset of >1000 individuals with well documented and long standing histories in Europe using principal component analysis (PCA) and clustering algorithms. Leave-one-out analysis is used to determine the accuracy of ancestry inferences. Male samples are additionally analyzed with Y-SNP markers to determine their membership in haplogroups known to be informative for differentiating European populations, thereby providing a phylogeographic framework for interpreting autosomal patterns of diversity that result in imprecise inferences of ancestral origins. Additionally we incorporate highly informative Y-SNP markers into the inference calculations and report changes to accuracy rates compared to those based on autosomal SNPs alone.

3322F

Ethiopian genome project. *L. Pagani[1,2], T. Kivisild[2], A. Tarekegn[3], R. Ekong[4], C. Plaster[4], I. Gallego Romero[1], T. Oljira[3], E. Mekonnen[3], Q. Ayub[1], S.Q. Mehdi[5], M.G. Thomas[6], D. Luiselli[7], E. Bekele[3], N. Bradman[4], D.J. Balding[8], C. Tyler-Smith[1].* 1) The Wellcome Trust Sanger Institute, The Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, United Kingdom; 2) Division of Biological Anthropology, University of Cambridge, CB2 1QH, UK; 3) University of Addis Ababa, and Center of Human Genetic Diversity, P.O. Box 1176, Ethiopia; 4) The Centre for Genetic Anthropology, University College London, WC1E 6BT, UK; 5) Centre for Human Genetics, Sindh Institute of Urology and Transplantation, Karachi, 74200, Pakistan; 6) Molecular and cultural evolution lab, University College London, WC1E 6BT, UK; 7) Anthropology Unit, Department of Experimental Evolutionary Biology of the University of Bologna, 40126, Italy; 8) University College Genetics Institute, University College London, WC1E 6BT, UK.

Humans and their ancestors have traversed the Ethiopian landscape for millions of years and present-day Ethiopians show great cultural, linguistic and historic diversity, which makes them essential for understanding African variability and human origins. Following the genotyping of 235 individuals from 10 Ethiopian and two neighboring (South Sudanese and Somali) populations we confirmed substantial genetic diversity both within and between populations, and revealed a match between genetic data and linguistic affiliation (Pagani et al. 2012, AJHG). Using comparisons with African and non-African reference samples in 40-SNP genomic windows, we identified “African” and “non-African” haplotypic components within each individual Ethiopian genome and found evidence for substantial gene flow into East Africa to ~3 KYA, as expected from historical records. To refine the genomic information and overcome the ascertainment bias of the genotyping procedure, we sequenced the whole genomes of 25 individuals from each of five representative Ethiopian populations on the Illumina HiSeq Platform (24 samples at ~10X and 1 sample at ~30X from each population). These data provide an unbiased description of Ethiopian genomic diversity and place the Ethiopian populations among the most diverse worldwide. Our ongoing analyses aim to further refine the “African” and “non-African” genetic partitioning to elucidate the admixture dynamics and identify regions with “African” or “non-African” ancestries. These regions will be independently analyzed in search of signatures of selection and for the information they can provide about the genetic source of the out-of-Africa migration.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3323W

Analysis of whole mitochondrial genomes from Iñupiat populations of the Alaskan North Slope. *J. Raff[1], M. Rzhetskaya[1], J. Tackney[3], M.G. Hayes[1, 2].* 1) Division of Endocrinology, Metabolism, and Molecular Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL; 2) Department of Anthropology, Northwestern University, Chicago, IL; 3) Department of Anthropology, University of Utah, Salt Lake City, UT.

Much has been inferred about the population history and colonization of the North American Arctic from the analysis of mitochondrial genomes of Siberian, Canadian, and Greenlandic peoples. However, inhabitants of the Alaskan North Slope—the presumed geographic origin of the Thule culture, and the hypothesized staging point for at least two trans-Arctic colonizations eastward (at 4000 YBP and 800 YBP)—have not been genetically characterized. While the ancestors of modern Iñupiat/Inuit peoples have been placed squarely within the more recent Thule cultural expansion, the extent to which earlier Holocene migrations of extinct Paleo-Eskimo contributed genetically to present day Arctic gene pools is unclear. Our analysis of mitochondrial HVSI and II from 139 North Slope individuals had previously revealed unexpected genetic diversity in this region. We therefore sequenced complete mitochondrial genomes from these individuals and have found new subtypes of the expected haplogroups A2a, A2b and D3, common among modern Inuit populations. We have also sequenced haplotypes of D2a, common in North America amongst the Aleutian Island chain, and C4, a recently established pan-American founding C lineage currently represented by only a handful of complete mtDNA genomes. Our findings of the expected subset of Beringian-specific mtDNA haplotypes, seen throughout the eastern Arctic, supports the hypothesis that the Alaskan North Slope was the source of the Paleo- and Neo- Eskimo expansions. However, the addition of haplotypes D2a and C4 in these northern-most populations raises interesting questions concerning Holocene gene flow in the North Slope. We present revised phylogenies for these haplogroups and the test various models for the origin and expansion of Arctic populations.

3324T

Genetic ancestrality estimative for Brasilia population using CODIS autosomal markers. *A.E. Svidzinski, G.C. Dalton, R.C. Toledo, A.C. Arcanjo, S.F. Oliveira.* Laboratório de Genética, Departamento de Genética e Morfologia, Instituto de Ciências Biológicas, Universidade de Brasília, Brasília, Distrito Federal, Brazil.

The Brazilian settlement process was carried out by three major population groups: the Amerindian land inhabitants, the European colonizers and the African slave workers. Considering its continental size, different regions of the country had unique populating and migration characteristics, resulting in many particular realities inside the same country. The Midwest region has a relatively recent occupation history, due to a large settlement process when the new capital of the country, Brasilia, was built far from the coastal region. Because of the intensive migration from all the parts of the country, the Brazilian Midwest region is considered to be a portrait of the whole Brazilian population. The objective of this study was to evaluate 12 STRs CODIS autosomal markers in a sample of Brasilia's population, compare with other Brazilian population groups and analyze its relation with the three parental population groups - Amerindian, European and African - aimed to extract genetic ancestrality information from these STR markers. The markers D8S1179, D21S11, D7S820, CSF1PO, D3S1358, TH01, D13S317, D16S539, vWA, TPOX, D5S818 and FGA were genotyped in a sample of 1715 individuals. The genetic profiles were analyzed using the STUCTURE software, OMNIPOP and MPGO/EHSTRAFD spreadsheets to seek for the genetic contribution of the parental groups and population structure in the sample. The results showed that, compared with other local groups in Brazil, the sample from Brasilia had no significant difference in genetic constitution. Regarding to ancestrality information, the estimated parental contributors were 69% Europeans, 20% Africans and 11% Amerindians. This result is very similar to the one obtained from AIMs specific ancestrality markers. Especially in five markers (TPOX, D5S818, TH01, D13S317 and FGA), there are significant allele frequency differences that allow an association between the individuals and the parental groups. Therefore, when analyzing an autosomal STRs genetic profile from an individual from Brasilia's population, even with markers that were not developed for this purpose, it is possible to retrieve ancestrality information regarding the Amerindian, European and African parental groups.

3325F

The population genetics of native Peruvian populations: evolutionary inferences and biomedical implications. *E. Tarazona-Santos[1], L. Pereira[1], M. Scliar[1], R. Zamudio[1], G.B. Soares-Souza[1], L.W. Zuccherato[1], M. Gouveia[1], F. Soares[1], F. Kehdy[1], W.C.S. Magalhaes[1], M.R. Rodrigues[1], E. Hollox[2], S.J. Chanock[3], R.H Gilman[4].* 1) Dept Biol Geral, CP 486, Univ Federal Minas Gerais, Belo Horizonte, Brazil; 2) Department of Genetics, University of Leicester, UK; 3) Laboratory of Translational Genomics, National Cancer Institute, NIH, US; 4) Johns Hopkins School of Public Health, MD US; Universidad Peruana Cayetano Heredia, Peru; Asociacion Benefica PRISMA, Peru.

Peru has the largest Native American (NA) population in South America. In the Andes, populations derive from a unique complex of societies that evolved during the last 10000 years along the Andes and its adjacent coast. During Pre-Columbian times, they reached a level of socioeconomic development unmatched elsewhere in South America. Smaller populations are settled in the Amazonian region, and in the hereafter called transition area between the Andes and the Amazonas. A mestizo population of millions is also settled in current Peruvian cities. We are studying the population genetics of Peruvians, addressing evolutionary inferences and biomedical implications of their diversity. We summarize three results: 1. Using 20kb resequenced from neutral inter-genic regions, we made inferences about the history of Andeans Quechua and a population from the transition area (Shimaa). These populations derived from an ancestral population with an effective size (Ne) of ~6000 individuals that split about 1025 years ago. At this time the Ne of Quechuas (Ne=5000) was five times larger than the Ne of the Shimaas, the former showing an evidence of a posterior demographic expansion, and the latter of a bottleneck. Thus, Andean and transition-area populations, even if they speak languages from different families and being culturally differentiated, share very recent ancestral populations. 2. Populations considered as mestizo have very high NA ancestry. A sample (n=297) from Lima, genotyped for 106 AIMs, has 78.4% of NA ancestry, suggesting that the Peruvian population is a potential target for large epidemiological studies in NA. 3. We identified genes with variants that are common in NA and rare elsewhere. The FCGR3B CNV locus has a high frequency of deletions in our native samples. This variant has been elsewhere associated with lupus erythematosus and rheumatoid arthritis, two diseases associated with NA ancestry. On the basis of the genotyping of 1442 SNPs on 411 genes, we determined that the following genes, that on the basis of recent GWAS are associated with complex diseases, show a differentiated haplotype structure in Native Americans: CASR, CAV1, CYP19A1, DRD2, FUT2, IL13, IL15, IL6R, IL7R, KRT23 and SCARB1 are highly differentiated from Africans/Europeans; ADH1C, APOB, CASR, CCND1, CDKN2A, CDK5, CYP19A1, ESR1, IGF2, IL1RN, LEPR and SCARB1 are highly differentiated from Asians. We are performing enrichment and functional analyses to validate these results.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3326W

Coalescent Inference of Evolutionary Parameters Using Serially Sampled Genomic Sequence Data: Theory and Applications in Human Population Genetics. *D.A. Vasco[1], M. Kato[2], Z. Ye[3], D. Lee[4], T. Carter[1], S.J. Hebbring[1], A. Rodrgio[5], S. Schrodi[1], S. Lin[3].* 1) Center for Human Genetics, Marshfield Clinic Research Foundation, Marshfield, WI.,; 2) Cold Spring Harbor Laboratory, Cold Spring Harbor, NY; 3) Biomedical Informatics, Marshfield Clinic Research Foundation Marshfield, WI; 4) Department of Animal Life and Environment Sciences, Hankyong National University Republic of Korea; 5) Dept of Biology, Duke University, Durham, NC.

Estimating ancestral and current effective population sizes using genomic information in expanding populations is a fundamental problem in human population genetics. Also, patterns of genomic variation may be used to estimate divergence times between human populations that split apart in the recent past such as Africans and non-Africans. This is possible because current patterns of genomic information in a sample of sequences are a reflection of past evolutionary processes of demographic change, mutation, migration, genetic drift and selection. These historical processes can be efficiently studied using a random genetic process, the coalescent, which mimics the evolutionary biology of genealogies underlying these complex processes going backwards in time. It has been one of the main statistical tools to infer parameters using sequence data by population geneticists. Coalescent inference in molecular population genetics has often been based upon full-likelihood methods which potentially offer a powerful statistical framework for inferring population genetic parameters. Recently, next generation sequencing methods have allowed developing evolutionary parameter inference pipelines for very large data sets, with large numbers of whole genomes of individuals sampled through space and time. However, for large and complex data sets computationally intensive methods based upon full-likelihood estimates may encounter difficulties. First, the methods may be slow or difficult to implement. Second estimation bias may markedly affect the accuracy and reliability of parameters. This has led to methods that tackle some of these problems by reducing the dimensionality in the full data set by adequately summarizing the information in the data. For example, historical population growth rates for putative population bottlenecks and expansions have recently been inferred by summary statistic coalescent methods using genomic data. We build upon this recent work using sampled ancient human DNA sequences. Using simulations and empirical data analysis we show how to extend current coalescent summary statistic estimation methods for serially sampled genomic data. A second problem to be discussed involves the evolutionary analysis of within-patient somatic cell line data obtained from flow cytometry using a next generation sequencer, we present a method to infer population divergence times at which punctuated clonal expansions occur with evolved population changes in ploidy.

3327T

Ancestry Perception Bias in Latin American Admixed Populations. *K. Adhikari[1], A. Ruiz-Linares[1], G. Bedoya[2], C. Bortolini[3], S. Canizales[4], C. Gallo[5], S. Gibbon[1], R. Gonzalez[6], F. Rothhammer[7], The CANDELA consortium.* 1) University College London, London, United Kingdom; 2) Universidad de Antioquia, Colombia; 3) Universidad Federal do Rio Grande do Sul, Brazil; 4) Universidad Nacional Autonoma de Mexico, Mexico; 5) Universidad Peruana Cayetano Heredia, Peru; 6) Centro Nacional Patagónico-CONICET, Argentina; 7) Universidad de Chile, Chile.

CANDELA (Consortium for the Analysis of the Diversity and Evolution of Latin America) is an international collaboration involving researchers from the UK and several Latin American countries (Argentina, Brazil, Chile, Colombia, Mexico and Peru). With its history of extensive population mixture, Latin America in many ways represents a form of natural experiment providing an advantageous opportunity to explore genetic and social aspects related to human biological diversity. CANDELA aims to perform a multidisciplinary study of a wide range of phenotypic, social and genetic data for about 8,000 Latin American individuals. An initial analysis including 40 Ancestry Informative Markers enabled estimation of individual Amerindian, European and African ancestry proportions (using STRUCTURE software) and their contrast with self-reported ancestry as well as with a range of phenotypic and social variables. As expected, we find a significant positive correlation of the genetic and self-perceived ancestries. However, across the region, we observe strong biases in self-perceived relative to genetic ancestry. For example, individuals with low to moderate (0–40%) European genetic ancestry tend to self-perceive their European ancestry as even lower; while individuals with high (>60%) African or Native American genetic ancestries tend to estimate these ancestries as even higher. There is also significant variation across countries in these biases, as much as 20–30% difference in the genetic v. self-perceived ancestries - in the Mexican or Chilean samples, individuals underestimate their Native American ancestry by about 20%, while in the Brazilian or Colombian samples individuals overestimate their Native American ancestry by a similar amount. Several factors could underlie these biases in self-perception of ancestry. For instance, we find that individuals with darker skin color significantly overestimate their African ancestry. We also find a strong effect of gender - relative to men, women underestimating their European ancestry and overestimating their African ancestry. Overall, we find evidence of a complex interplay of social and phenotypic factors in self-perception of ancestry: while visible traits such as skin color have an important impact, perception of ancestry is also intricately linked with social factors, as evidenced by variation across the countries examined.

3328F

OriginMiner : show me your genes and I'll tell you where you come from. *M. de Tayrac[1,2], M.C. Babron[3,4], E. Génin[3,4].* 1) CNRS, UMR 6061, Institut Génétique et Développement de Rennes, F-35043 Rennes, France; 2) Université Rennes 1, UEB, IFR 140, Faculté de Médecine, F-35043 Rennes, France; 3) INSERM U946, Univ Paris-Diderot, Sorbonne Paris Cité, F-75010 Paris, France; 4) Institut Universitaire d'Hématologie, Univ Paris Diderot UMRS946, F-75010 Paris, France.

Many genetic markers show significant differences in allele frequencies between human populations. These differences can lead to false positives in genetic association studies if not taken into account. A solution to limit this problem is to match patients and controls according to their population of origin. However, the population of origin is not always known or reliable. Fortunately, dense genotype information makes possible to determine it beforehand. The objective of our work was to develop a user-friendly tool, OriginMineR, to infer the population ancestry of individuals based on their genotypes. We used worldwide reference panels for which both population and genotype information are available: the HAPMAP3 panel consisting of 1,301 individuals from 11 populations and the Human Genome Diversity Panel (HGDP-CEPH), which includes 940 individuals from 51 populations. The genotypic information of these panels was restricted to subsets of SNP markers shared by the most commonly used genotyping arrays (Affymetrix Array 5.0 and Illumina Human660W-Quad). From these data, we identified the genetic signatures of the reference populations by Principal Component Analysis (PCA). For a set of unknown individuals for which only genotypic data are available, OriginMineR determines their geographic origin, allows their visualization by projection on the reference PCA maps and provides their posterior probabilities of geographic membership. According to the reference panels used, it is possible to refine the determination of the geographical origin of an individual within the major world regions. Our tool was validated on several samples of individuals recruited as part of genome-wide association studies. OriginMineR is a valuable tool both in the fields of genome wide association studies (GWAS) and of population genetics. For GWAS, it includes the possibility for each patient to extract a list of closely related controls from the reference panels. It is also possible to match cases and controls when both are provided. OriginMineR allows locating unreferenced populations on the worldwide map, thus providing clues for population genetics studies. The OriginMineR tool has been implemented in an R package.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3329W

Genetic ancestry and admixture analysis in a Bermudian population reveals evidence of Native American origins consistent with oral histories and genealogies. *J.B. Gaieski[1], E. Elhaik[2, 3], A.C. Owings[1], M.G. Vilar[1], A.T. Walia[1], D.F. Gaieski[4], R.S. Wells[5], T.G. Schurr[1], The Genographic Consortium.* 1) Anthropology, University of Pennsylvania, Philadelphia, PA; 2) Department of Mental Health, Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD; 3) McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD; 4) Department of Emergency Medicine, Hospital of the University of Pennsylvania, Philadelphia, PA; 5) Mission Programs, National Geographic Society, Washington, D.C.

Background: Shortly after its colonization in the early 17th century, Bermuda became the first English speaking dependency to forcibly import its labor by trafficking in enslaved Africans, European ethnic minorities, and indigenous Americans. Unlike the many ethnic groups that now call the island home, Bermuda's St. David's Islanders claim to be linked to Native American ancestors. In particular, their use of oral traditions and complex genealogies helps to reinforce their Native American identity. To elucidate the influence of historical events on genetic ancestry and native cultural identity among St. David's Islanders, we examined mtDNA and Y-chromosomal variation in over 100 individuals. We found that the majority of their mtDNA and Y-chromosome haplotypes (greater than 98%) were African and West Eurasian in origin. However, due to the limitations of this approach in reconstructing the genetic history of admixed populations, and because most participants were interested in learning more about their genetic genealogies, we expanded our analysis to include autosomal markers using a novel genotyping platform. Methods: To identify genetic contributions of putative indigenous American ancestors among the St. David's Islanders, we used the GenoChip to genotype Bermudians along with 200 samples from ~20 worldwide populations. Developed by Genographic Project scientists, the GenoChip is a SNP array ascertained from over 450 worldwide populations, and is dedicated to enhancing our knowledge of genetic anthropology. Results: Principal component analysis of the autosomal SNP data separated our participants into three discrete clusters. An admixture analysis identified up to 9% ancestry associated with Native Americans overall. The two largest clusters overlapped with African Americans and Puerto Ricans, and distributed evenly amongst the two main clusters (mean of 3% each). Samples from the third cluster averaged an unusually high Native American ancestry (mean of 6%). Conclusions: The GenoChip enabled us to detect otherwise elusory Native American ancestry among the Bermudians of St. David's Island. We speculate that the uneven distribution of this ancestry is due to admixture of Africans, Europeans, and Native Americans in varying degrees in the different source populations for modern-day St. David's Islanders. Application of this novel genotyping platform has provided new insights into the complex history of the Bermudian population.

3330T

The MHC and HLA alleleic diversity of the 1000 Genome samples. *P. GOURRAUD[1], N. Cereb[2], P. KhanKhanian[1], M. Maiers[5], S. Yang[2], R. Dunivin[3], M. Feolo[3], J. Rioux[4], S. Hauser[1], J. Oksenberg[1].* 1) UCSF, San Francisco, CA; 2) Histogenetics Ossining, NY; 3) National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD; 4) 4-University of Montreal Quebec Montreal Canada; 5) National Marrow Donor Programm, Minneapolis MN.

The 1000 Genomes (KG) Project aims to provide a deep characterization of human genome sequence variation. Specifically, the goal is to characterize over 95% of variants that are in genomic regions accessible to current high-throughput sequencing technologies in each of the five major population groups (about 100 individuals sampled in 14 populations from Europe, East Asia, South Asia, West Africa and the Americas). The KG project currently plans to sequence each sample to an average 4X coverage, which should allow the genome-wide detection of most variants with frequencies as low as 1% (the classical threshold for definition of polymorphisms). However, in the MHC, only the top 10 frequent haplotypes are in the 1% frequency range whereas thousands of haplotypes are present at lower frequencies. Because of the limitation of the sequencing techniques used by the KG project and due to the known genetic diversity of HLA genes in the MHC region, we have used sequence based techniques to type the HLA-A, HLA-B, HLA-C, HLA-DRB1 and HLA-DQB1 genes of 932 KG samples and combine the results with the 103,310 variants in the MHC region genotype by the KG project. Using pairwise Identical By Descent distances between individuals and Multiple Dimension Scaling, we establish the relationship between ancestry and the genetic diversity in MHC region and determine the correlation between the MHC based IBD distances (class I and/or Class II) to the genome-wide based corresponding estimations (r=0.53, p<1e-16). We use the data to reassess the existence and the robustness of MHC region variants tagging specific HLA alleles and assess the limitations of SNP-based HLA imputations. Further, we present plan to make this HLA information publically available on the DbMHC portal.

3331F

Identification of Sickle Cell Disease Alleles in the Eastern Caribbean Population. *C. Logronio, M. Graham, K. Wiitala, P. Gibbons, E. Bashover, C. Headland, I. McIntosh.* American University of the Caribbean School of Medicine, Cupecoy, St. Maarten.

Sickle Cell Disease (SCD) results from homozygosity or compound heterozygosity for mutations at codon 6 of β-globin and is commonly seen in populations of African origin. SCD has been studied extensively in N. America, Africa and Jamaica, however studies of Eastern Caribbean populations are limited. We developed a noninvasive method to detect the sickle cell mutations through extraction of DNA from buccal cells, PCR amplification and identification of the p.Glu6Val (Hb-S) and p.Glu6Lys (Hb-C) alleles through restriction endonuclease digestion and gel electrophoresis. We substantiated the efficacy of the method by testing 38 volunteers of African ancestry in St. Maarten and identified 5 Hb-S and 4 Hb-C heterozygotes. Amongst 26 subjects self-identified as Afro-Caribbean the frequencies of Hb-S and Hb-C were 0.074 and 0.056, respectively. The ratio of Hb-S:Hb-C alleles among Afro-Caribbeans is lower than previously reported for African-Americans suggesting diversity in founder populations. This non-invasive method could be used for future newborn screening and carrier detection on St. Maarten and broaden our understanding of SCD in the Caribbean population.

3332W

Khoesan genomic signatures in contemporary populations from southern Africa. *D.C. Petersen[1], O. Libiger[2], R. Hardie[3,4], L.I. Hannick[1], R. Wilkinson[5], R.H. Glashoff[6], M. Mukerji[7], P. Fernandez[8], N.J. Schork[2], V.M. Hayes[1], Indian Genome Variation Consortium.* 1) J. Craig Venter Institute, San Diego, CA; 2) The Scripps Translational Science Institute, Scripps Health and The Scripps Research Institute, La Jolla, CA; 3) Faculty of Medicine, University of New South Wales, Sydney, New South Wales, Australia; 4) The Garvan Institute of Medical Research, Sydney, Australia; 5) Namibian Blood Transfusion Services, Windhoek, Namibia; 6) Division of Medical Virology, Department of Pathology, Faculty of Health Sciences, University of Stellenbosch, Tygerberg, South Africa; 7) Insitute of Genomics and Integrative Biology (CSIR), Delhi, India; 8) Division of Urology, Department of Surgical Sciences, Faculty of Health Sciences, University of Stellenbosch, Tygerberg, South Africa.

The discovery of novel genomic variation in the Khoesan southern African population has broadened our perception of the extent of global genetic diversity. The term Khoesan (or Khoisan) collectively represents small isolated groups of forager peoples who communicate using clicking languages, but many of these groups have gone extinct. Global human migrations within and to Africa have led to extensive admixture and the emergence of contemporary populations that may carry genomic signatures of lost Khoesan lineages. The earliest arrivals into present day South Africa were southward migrating Bantu, including the amaXhosa. Centuries later, the arrival of European colonists together with slaves resulted in the founding of two admixed southern African populations, the South African Coloured and the Namibian Basters. We performed genome profiling of 70 ancestrally self-identified individuals (15 amaXhosa, 25 Coloured and 30 Basters) using Illumina HumanOmni1-Quad arrays and combined the results with data previously published for individuals from historically relevant ancestral populations including, indigenous Khoesan (Ju/'hoan), continental African (Yoruba), migrant colonists (European) and likely contributors to the Dutch East-Indian slave trade (Asian). We defined specific ancestral contributions to the amaXhosa and Basters, both showing significant Khoesan ancestry. Within our three geographically defined Coloured subgroups, we found that Khoesan admixture fractions only reach significance in subgroups located further away from the center of colonization, previously known as the Cape of Good Hope in South Africa. The most dramatic gender-specific contributions were found in the Basters, presenting with 91.7% Khoesan-derived mitochondrial and 93.3% non-African-derived Y-chromosomal DNA. The overall Khoesan contribution based on autosomal markers was most consistent within the amaXhosa (range 30–41%), then the Basters (17–40%), while most variable within the Coloured (6–57%). Contemporary populations with ancient Khoesan genomic signatures provide a unique model for studying complex global admixture patterns and will play an important role in defining the divergence between human populations.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3333T

Complete mitochondrial DNA genome sequences from the first New Zealanders: ancient DNA and the settlement of East Polynesia. *E.A. Matisoo-Smith[1,2], M. Knapp[1,2], S. Prost[1,2,3], K.A. Horsburgh[1,2,4], J. Stanton[1], H. Buckley[1].* 1) Anatomy, University of Otago, Dunedin, New Zealand; 2) Allan Wilson Centre for Molecular Ecology and Evolution, University of Otago; 3) Department of Integrative Biology, University of California, Berkeley; 4) School of Geography, Archaeology and Environmental Studies, University of Witwatersrand.

The issue of Polynesian origins has been debated since Europeans first arrived in the Pacific and the genetic evidence for their origins and migration pathway continues to be debated today. Identification of a combination of mtDNA mutations initially referred to as the “Polynesian motif” found at high frequency in Polynesian populations has provided a marker for tracking the origins of this last major human migration event. Despite the interest in identifying Polynesian origins, there has been little research on mtDNA variation within Polynesia. This is perhaps due to the reported near ubiquitous distribution of the motif within the Polynesian Triangle and the past focus on HVR sequences only. The supposed lack of mtDNA diversity in Polynesia is often presumed to be the combined result of population bottlenecks that occurred during colonising voyages and the speed and recency of colonization. We present here complete mitochondrial genome sequences of the likely founding population of Aotearoa/New Zealand recovered from the archaeological site of Wairau Bar. Of four individuals sequenced, at least three possess unique haplotypes indicating significantly more variation in the founding populations of East Polynesia that expected based on evidence from HVR data alone. These data represent the first complete mitochondrial genome sequences from ancient Polynesian voyagers and provide new insights into the genetic diversity of human populations in the Pacific at the time of the settlement of East Polynesia.

3334F

A Panel of Ancestry Informative Markers for Estimating and Correcting Potential Effects of Population Stratification in Han Chinese. *P. Qin[1], WF. Jin[1], DS. Lu[1], HY. Lou[1], JC. Wang[2], HJ. Xu[3], L. Jin[1,2], SH. Xu[1].* 1) Chinese Academy of Sciences Key Laboratory of Computational Biology, Chinese Academy of Sciences and Max Planck Society (CAS-MPG) Partner Institute for Computational Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shangh; 2) Ministry of Education (MOE) Key Laboratory of Contemporary Anthropology, School of Life Sciences and Institutes of Biomedical Sciences, Fudan University, Shanghai 200433, China; 3) Department of Rheumatology and Immunology, Shanghai, Changzheng Hospital, The Second Military Medical University Hospital, Shanghai 200003, China.

Population stratification can act as a confounding factor in genetic association studies and lead to false-positive or false-negative findings. Previous studies have revealed genetic substructures in Han Chinese population, the largest ethnic group in China and composing 20% of the entire global human population. In this study, we examined 1,610 individuals with about 1 million single nucleotide polymorphisms (SNPs) and attempted to screen a panel of ancestry informative markers (AIMs) to facilitate discerning and controlling population structure in the future association studies on Han Chinese. Our analysis of the genome-wide data confirmed the north-south differentiation of Han Chinese population, accordingly, we developed a panel of 150 validated SNP AIMs which can determine northern or southern origin of each Han Chinese sample. We further evaluated the performance of the AIMs panel in association studies using both empirical and simulated data. Our results showed that this AIMs panel could have enough power for discerning and controlling population stratification in candidate gene association studies which are much cost-effective in the post-GWAS era.

3335W

Genomic demographics of the genetically isolated Amish populations of Ohio and Indiana. *L.N. D'Aoust[1], A.C. Cummings[1], L. Jiang[1], R. Laux[1], D. Fuzzell[1], L. Caywood[2], L. Reinhart-Mercer[2], M. Courtenay[2], W.K. Scott[2], M.A. Pericak-Vance[2], J.L. Haines[1].* 1) Center for Human Genetics Rsearch, Vanderbilt University Medical Center, Nashville, TN; 2) Hussman Institute for Human Genomics, Miller School of Medicine, University of Miami, Miami, FL.

Founder populations, such as the Amish of Ohio and Indiana, are advantageous for genetic studies. These communities are genetically isolated because members marry primarily within their faith and have a strict lifestyle, leading to a more culturally, environmentally and genetically homogenous population. Therefore, it is hypothesized that the genetic etiology of a disease in the Amish is more likely simpler and restricted to fewer genes contributing risk. We characterized the genomic demographics of these populations to understand the genetic architecture of such isolated populations. We used data for 600,000 SNPs from 900 individuals to evaluate inbreeding coefficients, kinship coefficients, fixation index, and the number of paternal and maternal lineages present. The inbreeding coefficients calculated from the markers (average of 0.006 ± 0.011) are not similar to the coefficients determined from the pedigree structure (average of 0.014 ± 0.010). When we evaluated the kinship coefficients for each relative pair using the SNP data, they did not always match the coefficients based upon the pedigree. To determine why the estimates deviated, we evaluated the average coefficient for different types of pairs. The coefficients from the genotypes and the pedigree are respectively 0.250 ± 0.030 and 0.260 ± 0.006 for siblings, 0.250 ± 0.003 and 0.260 ± 0.004 for parent-offspring, and 0.128 ± 0.016 and 0.135 ± 0.004 for half-siblings. The average F_{ST} across the genome is 0.116 ± 0.359 for the Amish compared to the CEU HapMap samples, suggesting there may be multiple levels of inbreeding present. We identified paternal and maternal lineages from the combined pedigree of 5,437 individuals. In total, 71 paternal and 75 maternal lineages were identified, of which 37 and 54 were examined, respectively. Examination of Y chromosome markers identified three paternal lineages that include males whose alleles suggest the lineage includes unrelated males. The male members who carry the minor alleles can be related via a single male ancestor within one of these pedigrees. The mitochondrial genotypes of five maternal lineages suggest these lineages include individuals who are not maternally related. These data suggest that inbreeding is complex, and that a single figure may not accurately describe the entire population. Additionally, we can use these results to adjust analysis to properly reflect true relationships by correcting for errors in reported pedigree structure.

3336T

Jethro's descendants: a journey of 1000 years. *E. Friedman[1], J. Zidan[2], D. Ben Avraham[3], T. Maray[4], G. Atzmon[3,5].* 1) Susanne Levy Gertner Oncogenetics Unit, The Danek Gertner Institute of Human Genetics, Chaim Sheba Medical Center, 52621 Tel-Hashomer, Israel;; 2) The Oncology Department Ziv Medical Center, Zefat, Israel; 3) Departments of Medicine, Albert Einstein College of Medicine, Bronx, NY 10461; 4) Golan for Development Madjal Shams The Golan Heights; 5) Genetics, Albert Einstein College of Medicine, Bronx, NY 10461.

The Druze people enumerate ~1000000 people worldwide, share common beliefs and social practices, rarely marry outside of their faith and have a high rate of consanguinity. These features likely make the Druze people a genetic isolate. Previous attempts to define the genetic structure of the Druze people were based on unrelated individuals, primarily from Israeli residing Druze and using Y chromosome markers and mitochondrial DNA. The current study attempted to comprehensively decipher the genetic structure of geographically diverse residing Druze, based on parents-offspring trios. Forty parent-offspring trios from Beit Jan and the Golan Heights (20 each) were recruited, each trio representing a seemingly distinct geographical origin (=Hamulas). Genotyping was done using Affymetrix 6.0 platform and PCA, IBD sharing and Structure analysis as well as haplotype sharing were subsequently assessed by JMP Genomics, GERMLINE and STRUCTURE. Druze clustered with other Middle Eastern groups, and when compared with Bedouins, and Palestinians genotyped by the HGDP, the Druze formed its own distinct cluster, indicating the shared ancestry and relative isolation of each of those groups. The relative positions of the population clusters observed by PCA were confirmed by structure and IBD analysis. The closest genetic distance was noted between Golan Heights residing Druze and the Druze, Palestinian and Bedouin from HGDP panel. Haplotype and phasing are ongoing to accurately define common ancestor. This is the first study targeting Druze individuals that attempted to define common ancestor by geographical origin and facilitate future studies aimed at gene-disease interaction using the powerful tool of trios structure.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3337F

Overview of the genetic variations in the Y chromosome in the Japanese population. *Y. Sato[1,2], T. Shinka[2], A. Yamauchi[1], Y. Nakahori[2].* 1) Dept Pharmaceutical Information Science, Tokushima Univ, Tokushima, Japan; 2) Dept Human Genetics and Public Health, Tokushima Univ, Tokushima, Japan.

The modern Japanese people are widely believed to be the descendants of the Jomon and Yayoi people. There is one model to explain the formation of Japanese population called the dual structure model which hypothesizes that the indigenous people the Jomonese migrated in the northern island of Hokkaido and the Ryukyu Islands, and resided on the islands of Japan, before the Yayoi started to arrive in Kyushu (southern Japan) from the Korean Peninsula. As the Y chromosome is unable to recombine, the same type of Y chromosome is passed from father to son. Therefore, the Y chromosome is a useful tool for genealogical research. According to the haplogrouping and nomenclature system of the Y Chromosome Consortium, Japanese males mainly belong to haplogroups C, D, and O, which can be further sub-classified. Haplogroups C1 and D2 are postulated to be indigenous Japanese haplogroups, and the distributions of O2b1 and O2b* are limited to Japan and Korea. Hammer et al. have proposed that the Jomon males had the Y chromosomes with either haplogroups C or D, while Yayoi males had ones with haplogroup O from the analysis of the Ainu and Ryukyuans. Here, we hypothesized that if the abovementioned model is correct, the Y chromosome haplogroups C and D might display high frequencies on Hokkaido, and haplogroup O might display a high frequency in northern Kyusyu. In addition, the frequencies of these haplogroups might exhibit a geographical gradient in Japan. To investigate this hypothesis, we recruited 2,390 male volunteers from Hokkaido (Sapporo), northern Kyusyu (Nagasaki and Fukuoka), Shikoku (Tokushima), and Honshu (Osaka, Kanazawa, and Kanagawa) and typed their Y chromosome haplogroups. We also gathered data regarding the frequency distributions of Y chromosome haplogroups in East and Southeast Asia from the published literature. We did not detect any marked variability in the frequencies of these haplogroups among the Japanese populations. However, our results showed that haplogroup O3a4 is likely to be one of indigenous Japanese haplogroup. In conclusion, the Japanese population possesses several indigenous Y chromosome haplogroups, and their frequencies seem to have reached equilibrium.

3338W

Genome-wide SNP variation in sub-Saharan Africa is influenced by cultural and ethno-linguistic affiliation. *S. Soi[1], L. Scheinfeldt[1], C. Lambert[2], J. Hirbo[1], A. Ranciaro[1], S. Thompson[1], J. Marie Bodo[3], M. Ibrahim[4], G. Lema[5], T. Nyambo[5], S. Omar[6], C. Wambebe[7], D. Meskel[8], G. Belay[8], A. Froment[9], S.A. Tishkoff[1].* 1) Dept. Genetics and Biology, University of Pennsylvania, Philadelphia, PA, USA; 2) Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, USA; 3) Malaria Research and Training Center, University of Bamako, Bamako, Mali; 4) Department of Molecular Biology, Institute of Endemic Diseases, University of Khartoum, 15-13 Khartoum, Sudan; 5) Department of Biochemistry, Muhimbili University of Health and Allied Sciences, Dar es Salaam, Tanzania; 6) Kenya Medical Research Institute, Center for Biotechnology Research and Development, 54840-00200 Nairobi, Kenya; 7) International Biomedical Research in Africa, Abuja, Nigeria; 8) Department of Biology, Addis Ababa University, Addis Ababa, Ethiopia; 9) UMR 208, IRD-MNHN, Musée de l'Homme, Paris, France.

Human genetic relationships often correspond to ethno-linguistic classification and geographic proximity. However, language shifts and historical gene flow can obscure such patterns. Because linguistic methodologies do not permit the reconstruction of relationships beyond the level of language family and archaeological evidence is often incomplete, genetic data provide an important complement in the study of modern human demographic history. To further disentangle the complex demographic questions in Africa, we have generated an expanded genome-wide set of 1.19 million single nucleotide polymorphims (SNPs) using the Illumina 1M duo array from 724 individuals from 46 diverse ethno-linguistic populations from across sub-Saharan Africa. Principal components analysis (PCA), analysis of identity by descent (IBD) tracts inferred with GERMLINE, and Bayesian cluster analysis using STRUCTURE revealed unexpectedly close relationships among linguistically and geographically disparate hunting and gathering population samples. In particular, we have identified evidence of recent common ancestry among the several East African hunter-gatherer populations, southern African San, San, eastern and western Pygmies. Our results demonstrate that recent Neolithic population movements and dynamics have not completely obscured more ancient African population structure. In addition, we are applying approximate Bayesian computation (ABC) methods using haplotype and linkage disequilibrium statistics to infer demographic histories.

3339T

Positive natural selection and recent evolutionary history in Malays. *X. Liu[1], T.H. Ong[2], L.P. Wong[2], W.T. Poh[2], N.E. Pillai[3], C.C. Khor[4], Y.Y. Teo[1,2,3,4,5].* 1) NUS Graduate School, National University of Singapore, Singapore, Singapore; 2) Saw Swee Hock School of Public Health, National University of Singapore, Singapore; 3) Life Sciences Institute, National University of Singapore, Singapore; 4) Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore; 5) Department of Statistics and Applied Probability, National University of Singapore, Singapore.

Recent human evolutionary history may be revealed through the identification of positive selection events in the human genome. Here we perform a genome-wide scan of positive natural selection in the Malay genome. While most of the signals are shown to be shared between Malay and other populations in Asia and Europe, there were some signals that appear to be exclusive to South-East Asia. The shared signals may either be due to convergent evolution, where the populations were under similar selection pressure; or more commonly, these signals could have originated from a single mutation event in the common ancestral population of these groups. To differentiate between these two situations, we used a method that locates the haplotype pattern that the advantageous allele resides on, and predicts the frequency of the positively selected allele. Haplotype forms that are remarkably consistent between the populations indicate that these populations likely carry the same functional allele that has arisen in a single mutation event prior to the time to the most recent common ancestor (TMRCA) of these populations. To perform these analyses, we estimate the TMRCA of the Malay populations along with other Southeast and East Asian populations. Subsequently, for the positive selection signals that are inferred to have originated from the same mutation events, we fine-map these signals to identify the functional variants using dense whole-genome sequence data from the 1000 Genome Projects and from the Singapore Sequencing Malay Project. The alleles identified may yield vital insights into the functional aspects that led to survival advantages in Malays and other populations.

3340F

Genetic evidence of multiple non-Asian migrations into the New World. *N. Angerhofer[1,2], S. Woodward[1,2], N. Myres[1,2], T. Steele[2], A. Nelson[2], U.A. Perego[2,3], K. Chahine[1].* 1) AncestryDNA, Provo, UT 84604, USA; 2) Sorenson Molecular Genealogy Foundation, Salt Lake City, UT 84115, USA; 3) Dipartimento di Biologia e Biotecnologie, Università di Pavia, Pavia 27100, Italy.

Ancient migrations into the New World (North and South America) have been tied to Asia through Y chromosome (Ychr) and mitochondrial (mtDNA) affinities, where estimates of divergence place the original migrations at 15–18 kya. Recent migrations of populations to the Americas are a matter of historical record, starting with the Spanish in 1492, and with subsequent waves of Europeans, Africans and Asians. With targeted DNA testing, the evidence of these separate and distinct migrations can be detected.

During a recent survey of the nearly 45,000 Ychr haplotypes in the AncestryDNA database, a number of non-Asian haplotypes have been observed in Native American collections. Among these, a 37-loci STR non-Asian modal haplotype similar to haplogroup I was revealed to be shared between approximately 30 Central and South American samples. While they clustered together, these samples did not have close matches to known Native American haplogroups. Their time to coalescence was determined to be approximately 2000 years ago. They were distributed from Chile to Mexico, and with SNP testing, were determined to be members of Ychr haplogroup I-M26, a European haplogroup with high frequency in Sardinia and at low but significant frequencies in northern Spain, southern France, England and Ireland. Ychr SNP testing confirmed the rare M26 SNP could be used as evidence of a possible Iberian connection. Network and cluster analyses with more distant Y-STR I-M26 individuals show a strong connection to Spain and Portugal, which includes Basques and possible connections to ancient relatives in England, Wales and Ireland. It is expected that with the addition of autosomal SNP testing and analysis of linkage disequilibrium, characteristic indicators of distinct populations of post-Asian migrations into the Americas will become better defined.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3341W

The CARTaGENE Genomics Project : Population Genetics Analysis of the Founding Population of Quebec. *P. Awadalla[1,2,3], J. Hussin[1,2], Y. Idaghdour[2], A. Hodgkinson[2], J-P. Goulet[2], J-C. Grenier[2], T. de Malliard[2], V. Bruat[2], E. Gbeha[2], M. Capredon[2], E. Hip-Ki[2], Y. Payette[3], C. Boileau[3].* 1) University of Montreal, Montreal, QC, Canada; 2) CHU Ste Justine, Montreal, QC, Canada; 3) CARTaGENE Project, Montreal, QC, Canada.

European colonization of the province of Quebec began four centuries ago, with approximately 8,500 settlers, mostly of French origin. In the early settlement of Quebec, Europeans and local First Nations peoples exchanged genetic material, with the most common pattern being marriage between an aboriginal mother and a French father. After the British Conquest, between 2,000 and 4,000 Acadians and American Loyalists came to Quebec, and the French-Canadian population expanded rapidly. Today, the Quebec population comprises 7.8 million residents, of which ~80% are French speaking. In 2010, the CARTaGENE project collected data from 20,000 participants recruited throughout the province of Quebec. Genotyping data (Illumina Omni2.5M) was generated for ~1000 French-Canadians participants with a known extended pedigree. Despite their short divergence time and relatively similar geographic and environmental range, it has been demonstrated that the regional populations of Quebec are stratified and differences in disease prevalence are observed. Therefore, sampling includes individuals from three distinct regions of Quebec : the Montreal area, the Quebec city area and the Saguenay region. Here, we use this genetic data in order to explore regional population variation and study the effect of recent demographic history, admixture and migration patterns within Quebec. We further study the variation in recombination patterns and in IBD sharing and signatures of differential natural selection. Specifically, we (i) infer historical relationships between subpopulations in Quebec from genetic data; (ii) identify patterns of admixture in French-Canadians based on Native American and European source populations; (iii) compare recent breakpoints of recombination, ancestral population recombination rates and recombination hotspot location and (iv) identify differential natural selection operating in specific subpopulations of Quebec since the founding event. This work will give insight to gain a better understanding of the evolutionary forces that shaped the genomic landscape of French-Canadians and their implications in explaining differences in disease prevalence in regions of Quebec.

3342T

The Kaiser Permanente/UCSF Genetic Epidemiology Research Study on Adult Health and Aging: Characterizing ancestry of the GERA cohort using genome-wide data. *Y. Banda[1], M. Kvale[1], T. Hoffmann[1], S. Hesselson[1], H. Tang[3], C. Sabatti[3], D. Ranatunga[2], C. Schaefer[2], P. Kwok[1], N. Risch[1].* 1) Institute for Human Genetics, University California San Francisco, San Francisco, CA; 2) Kaiser Permanente Northern California Division of Research, Oakland, CA; 3) Department of Statistics, Stanford University, Stanford, CA.

Understanding human genetic diversity has important ramifications for all genomic analyses, including genome-wide association studies, admixture analysis and traits associated with ancestry. Primarily using Principal Components Analyses (PCA), we characterized the genetic structure in a sample of over 100,000 multi-ethnic subjects from Northern California who are part of the Genetic Epidemiology Research on Aging (GERA) cohort using 144,000 single nucleotide polymorphisms. The analysis includes approximately 84,000 whites, 8,000 Asians, 7,400 Latinos and 3,500 African Americans. We also projected subjects from the Human Genome Diversity Project onto the PC results to facilitate geographic interpretation. As in previous studies, in the analysis of the whites, Ashkenazi Jews formed a clearly distinct cluster based on the first two principal components, and individuals who are part-Jewish were also easily discerned. After removing the Jews from the analysis, the two-dimensional representation of the genetic relationships from the first two PCs closely follows the geographic distribution of the different European populations, with the first PC corresponding to a northwest-southeast cline, and the second PC to a southwest-northeast cline. In this analysis there was a continuum rather than clusters due to extensive mixing among various European Americans in the San Francisco bay area. South Asians formed a clearly distinct cluster, with apparent subgrouping likely separating Indian versus Pakistani origins. We also observed close correspondence between genetic and geographic distances when analysing the different East Asian subjects, with an apparent clustering from north to south in East Asia, consistent with extensive endogamy. Individuals of mixed East Asian-European ancestry were easily identified, and of particular interest was the observed continuous nature of a modest amount of European ancestry in individuals self-identified as Filipinos. Analysis of African Americans and Latinos revealed extensive European and African ancestry for both, and Native American ancestry for the latter, consistent with their self-reported nationality. The results obtained here are fundamental to the study of complex genetic disease in this large, population based cohort, through association studies, admixture analysis and admixture mapping, and in particular for investigating observed ethnic variation in diseases and traits.

3343F

Population diversity and history of the Indian subcontinent: Uncovering the deeper mosaic of sub-structuring and the intricate network of dispersals. *A. Basu, N. Sarkar Roy, P. P Majumder.* National Institute of BioMedical Genomics, Kalyani, India.

Past genomic studies have provided a broad picture of the peopling of the India: proposing a southern exit route through western coastline and migration from west and central Asia, through north-western corridor. We undertook a genomewide study(Illumina 1M SNP chip) comprising 367 individuals, drawn from 20 ethnic groups of India, including, Jarawa(JR) and Onge(OG) of the Andaman & Nicobar (A&N) Islands. Populations represent all geographic regions, linguistic groups and all social hierarchies within geographical/linguistic classes. Our analysis corroborated with previous studies on the existence of deep rooted population structure in India, as evidenced by a high average pairwise FST ~0.04. The average pairwise FST was extremely high(>0.1) between JR/OG and the mainland Indian populations, indicating the high differentiation between people of A&N and mainland India. The first Principal Component (PC), which accounts for 14% of the genetic variation, separates the JR and OG from the 18 mainland Indian populations. On inclusion of the Human Genome Diversity Panel(HGDP) data, the PCA shows that the JR and OG are close to the Melanesian and Papuan populations. Unlike Europe, geography alone is a poor predictor of genetic diversity and/or ancestry, as indicated by PCA, ADMIXTURE and MANTEL test indicating complex local histories. There was poor correlation between pairwise genetic distance and geographical distance(0.33). However, this correlation was high (0.41–0.76) when populations belonging to same social stratum were considered. North to south clinal variation, as reported earlier, was found for caste populations, but not for populations belonging to other social strata. Our results of admixture estimation and hierarchical clustering reveal that genetically the populations of mainland India form four ancestry clusters (cross validation error minimum(0.52) compared to other competing models): (1)TibetoBurman speaking populations of the northeast region. (2)Upper Caste groups of both northern(IE) and southern(DR) regions. (3)Austro-Asiatic(AA) speakers of central and eastern region. (4)DR speaking lower caste and tribal groups of southern region. Population ancestry analysis revealed that the AA and DR speakers have a deep-rooted ancestry and are perhaps relics of a very ancient dispersal event, perhaps the migration that entered India in the southern exit route wave. This finding is also supported by the analysis of similarity of haplotype structures.

3344W

Mitochondrial origins of the Cape Coloureds of South Africa. *K.G. Beaty[1], D.L. Phillips[1], M. Henneberg[2], M.H. Crawford[1].* 1) Lab of Biological Anthropology, Department of Anthropology, University of Kansas, Lawrence, KS; 2) Department of Anatomical Sciences, University of Adelaide, Adelaide, Australia.

Previous studies of genetic diversity have suggested that the Cape Coloureds of South Africa are a highly admixed population with genetic roots from indigenous African groups including Khoisans and later arriving Bantu speaking Xhosa farmers. Further genetic contributions came during the European colonization of South Africa, which added Dutch, French, German, and British markers to the gene pool. Slaves from Indonesia, Malaysia, Madagascar and India are also thought to have contributed to the genetic makeup of this ethnic group. This study examines the maternal contribution of each of these groups to the genetic makeup of the Cape Coloreds through sequencing of hypervariable region I of the mitochondrial DNA, and through restriction fragment length polymorphisms for haplogroup assignment and confirmation. A total of 94 individuals from Klein Karoo, South Africa were examined for this study, displaying a high level of gene diversity at 0.9663 when compared to other African groups. High frequencies of African haplogroups L1 (42.6 percent) and L2 (26.6 percent) were found. African haplogroup L0 was also detected in low frequencies, including L0d (6.4 percent), a haplogroup associated with Khoisan groups. L0a, a haplogroup associated with the Bantu expansion, was also detected at 4.3 percent. While the majority of lineages indicate an African maternal origin of Cape Coloureds, other contributions are noted through European and Eurasian haplogroups H (3.2 percent), U (4.3 percent), and R (1.1 percent). Asian haplogroup M was also found at 6.4 percent, indicating a Melanesian or Indian contribution to the Cape Coloured gene pool. This study shows the diverse maternal origin of South African Cape Coloureds.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3345T

North African Jewish and non-Jewish populations form distinctive, orthogonal clusters. *C. Campbell[1], P.F. Palamara[2], M. Dubrovsky[3], L.R. Botigué[4], M. Fellous[5], G. Atzmon[1], C. Oddoux[1], A. Pearlman[1], L. Hao[6], B. Henn[7], E. Burns[1], C. Bustamante[7], D. Comas[4], E. Friedman[3], I. Pe'er[2], H. Ostrer[1].* 1) Albert Einstein College of Medicine, Bronx, NY; 2) Columbia University, New York, NY; 3) Tel Aviv University, Israel; 4) Universitat Pompeu Fabra, Barcelona, Catalonia, Spain; 5) Cochin Institute, France; 6) University of Medicine and Dentistry of New Jersey; 7) Stanford University, Stanford, CA.

North African Jews constitute the second largest Jewish Diaspora group. Yet, their relatedness to each other, to European, Middle Eastern, and other Jewish Diaspora groups and to their former North African non-Jewish neighbors has not been well-defined. Here, genome wide analysis of five Jewish groups (Moroccan, Algerian, Tunisian, Djerban and Libyan) and comparison with other Jewish and non-Jewish groups demonstrated distinctive North African Jewish population clusters with proximity to other Jewish populations and variable degrees of Middle Eastern, European and North African admixture. These populations showed a high-degree of endogamy and were part of a larger Ashkenazi and Sephardic Jewish group. Two major sub-groups were identified by principal component, neighbor joining tree, and identity by descent (IBD) analysis - Moroccan/Algerian and Djerban/Libyan that varied in their degree of European admixture. The Libyan, Tunisian and Djerban Jewish populations showed the highest degree of within-population and cross-population IBD sharing. STRUCTURE and Xplorigin analysis demonstrated an enrichment of European ancestry in Moroccan and Algerian Jews, most likely of recent origin. By principal component analysis, all North African Jewish groups were orthogonal to contemporary populations from North and South Morocco, Western Sahara, Tunisia, Libya and Egypt. Thus, this study is compatible with the history of North African Jews - founding during Classical Antiquity with proselytism of local populations, followed by genetic isolation with the rise of Christianity and then Islam, and admixture following the emigration of Sephardic Jews during the Inquisition.

3346F

Hyper-sharing individuals in founder populations: Theory and implications to imputation strategies. *S. Carmi[1], P. Palamara[1], V. Vacic[1], T. Lencz[2], A. Darvasi[3], I. Pe'er[1].* 1) Department of Computer Science, Columbia University, New York, NY 10025, USA; 2) Department of Psychiatry, Division of Research, The Zucker Hillside Hospital Division of the North Shore-Long Island Jewish Health System, Glen Oaks, NY, 11004, USA; 3) Department of Genetics, The Institute of Life Sciences, The Hebrew University of Jerusalem, Givat Ram, Jerusalem, 91904, Israel.

Widespread sharing of long, identical-by-descent (IBD) genetic segments is a hallmark of populations that have experienced recent genetic drift. Recent progress in genotyping and sequencing technologies, coupled with new algorithms, enable now the detection of IBD segments even in large cohorts, with applications ranging from phasing and imputation through demographic inference to detection of association and positive selection. We focused on the total fraction of the genome shared IBD between arbitrary pairs of individuals. Our observations of total sharing in real cohorts suggest that this total IBD has a surprisingly wide distribution. Specifically, when considering the cohort-averaged sharing— the average total sharing between each individual to everyone else in the cohort— we repeatedly observed small groups of individuals with much more average sharing compared to others. The existence of such groups of `hyper-sharing' individuals bears important consequences to the design of sequencing studies, since, if they are selected for whole-genome sequencing, a larger fraction of the cohort can be subsequently imputed. We therefore set out to provide theoretical grounds for our empirical findings. In the Wright-Fisher model, we calculated the moments of the total sharing, and provided an approximate form for its distribution. We then investigated the average sharing between each individual to everyone else in the cohort. We found that even for very large samples, the cohort-averaged sharing approaches a normal distribution with variance that is cohort size-independent, thus implying the existence of hyper- (and hypo-) sharing individuals. We worked out explicitly the expected gain in imputation (and gene-disease association) power in sequencing studies when prioritization of samples for sequencing is IBD sharing-aware.

3347W

Analytical inference of human demographic history using multiple individual genome sequences. *H. Chen[1], D. Reich[2,3].* 1) Dept Epidemiology, Havard Sch Public Health, Boston, MA; 2) Department of Genetics, Harvard Medical School, Boston, MA 02115; 3) Broad Institute of Harvard and MIT, Cambridge, MA 02142.

A central problem in human population genetics study is to infer the past demographic history using genetic polymorphism data from present-day humans. The whole-genomic sequence data generated by new sequencing technology provide an unprecedented opportunity for population genetic studies. We present a new approach that infers the joint ancient human demography of multiple populations using multiple genomes from divergent populations. We use coalescent theory to analytically derive the joint allele frequency spectrum of multiple populations that experienced different demographic histories. Compared with the recently developed approaches, such as diffusion process approximation (dadi), sequential Markov Coalescent method (PSMC), and coalescent-MCMC (Markov Chain Monte Carlo, GPhoCS) approaches, our method has several advantages: (1) the method is computationally efficient, as the coalescent likelihood for the joint pattern of multiple populations can be obtained analytically without relying on numerical approximation of diffusion (e.g., Gutenkunst et al 2009) or computationally intensive approaches (e.g. Gronau et al 2011); (2) the method jointly analyzes more than three populations, and multiple individuals from each population, whereas previous approaches were limit to three populations (Gutenkunst et al. 2009) or one individual from each population (Gronau et al 2011; Li and Durbin 2011); and (3) detailed demographic scenarios can be incorporated into the model, including temporally-varying population sizes. We apply the method to two data sets: the first data includes three Europeans, two Asians, three Yoruba and one San sequenced using Illumina technology; and the second data set includes eight East Asians, five Northern Europeans and seven Yoruba sequenced using Complete Genomics technology. We estimate that Eurasian populations split from ancient Africans at 58,000–120,800 years ago, and the divergence time of Europeans and Asians occurred at 35,750–70,500 years ago. The inferred demographic history is generally consistent with previous studies.

3348T

Predicting the geographic origin of individuals using genome-wide SNP data. *D.J.M. Crouch, M.E. Weale.* Deparment of Medical and Molecular Genetics, King's College London, London, United Kingdom.

A capacity to predict the geographic origin of human DNA samples would have applications in forensic science, anthropology and population genetics. Over short geographic ranges, for example sub-continental regions, genetic differentiation between populations tends to be low, and the distribution of variation is clinal. Consequently, the most appropriate predictive model is one which returns continuous longitude/latitude estimates along with associated measures of uncertainty. We use the genetic covariance between individuals at known sample locations to model the spatial dependencies between allele frequencies as a function of distance. Allele frequencies at unknown locations are then imputed via this model, an approach known as Kriging, and the probability of observing the target individual's genotypes under Hardy-Weinberg Equilibrium is maximised for longitude and latitude. It is possible to extend this model by deriving the between-individual genetic covariance via haplotype information. Performance is assessed using cross-validation with European individuals from the HGDP and POPRES datasets, and we compare with alternative methods from the literature.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3349F

Y Chromosome J Haplogroups trace post glacial period expansion from Turkey and Caucasus into the Middle East. *B. Douaihy[1], D. Platt[2], M. Haber[1,3], A. Salloum[1], F. Mouzaya[1], M. Bou Dagher-Kharrat[4], G. Khazen[1], E. Matisoo-Smith[5], R.S. Wells[6], C.T. Smith[7], P. Zalloua[1,8], The Genographic consortium.* 1) Lebanese American University, Beirut, Lebanon; 2) Computational Biology Centre, IBM TJ Watson Research Centre, Yorktown Hgts, NY, USA; 3) Institut de Biologia Evolutiva (CSIC-UPF), Departament de Ciències de la Salut i de la Vida, Universitat Pompeu Fabra, 08003 Barcelona, Spain; 4) Laboratoire `Caractérisation Génomique des Plantes, Faculté des Sciences, Université Saint-Joseph, Campus Sciences et Technologies, Mar Roukos, Mkalles, BP: 1514 Riad el Solh, Beirut 1107 2050, Lebanon; 5) Allan Wilson Centre for Molecular Ecology and Evolution and University of Otago, Dunedin 9054, New Zealand; 6) The Genographic Project, National Geographic Society, Washington, DC, USA; 7) The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, UK; 8) Harvard School of Public Health, Boston, MA, USA.

Paleoclimate reconstructions have shown that several glacial refugia formed around the Mediterranean and Black Sea during the last glacial period (LGP) that dramatically affected the distribution of the populations of Eurasia and the Middle East. Post-glacial warming, beginning around 12,000 years ago, resulted in population migrations out of those refugia, and drove the Neolithic revolution. The timing and routes of these migrations and their specific regions of expansion remain elusive. The genetic signals marking the initial settlement following the LPG are also unclear. On the Y chromosome, there is significant regional variation among subhaplogroups of J within the Middle East that are informative about these events, which were investigated in more detail. 2774 samples were analyzed, including 941 newly genotyped, to characterize populations where J haplogroups have expanded geographically. Haplogroup J diversity was measured by coancestry using Y-STR haplotype effective number. Reduced Median networks were calculated for each of the J subhaplogroups and SNP coalescence times were estimated by BATWING. The J subhaplogroups' frequencies show substantial variation across the Mediterranean Basin. J subhaplogroups show little organization of their haplotypes by geography, suggesting that diversity evolved primarily within a pool of ancestral populations for a larger part of its history, then post-glacial expansions carried this diversity throughout their modern geographical range. Coalescence time estimates indicate longer evolution of J haplogroups in northern populations, in agreement with co-ancestry diversity. Population divergence time estimates are recent compared to coalescence times, supporting long evolution times prior to post-glacial expansions. Our data provide evidence for the timing and differential routes of post glacial repopulation of the region. When combined with archaeological and linguistic evidence, these genetic data allow us to reconstruct the spread of agriculture and the origins of various Neolithic cultures of the Middle East. The Neolithic expansion has been marked by haplogroups characteristic of the Middle East, with J haplogroups showing geographically differential frequency distributions and haplotype diversities. These results are suggestive of evolution within the Anatolian Peninsula and the Black Sea basin during the LGP, followed by multiple expansions taking distinct routes at different times subsequent to the LGP.

3350W

Genotyping at the DAT1 Locus; Determination of VNTR Allele Frequencies in Descendents of the Founding Families of Southwest Louisiana. *J. Drost, A. Figueroa, J. Angle.* Biology and Health Sciences, McNeese State Univ, Lake Charles, LA.

The human population of Southwest Louisiana offers a rich opportunity for genetics study because it was founded by a small number of individuals arriving independently representing a variety of ethnic backgrounds. Moreover, due to geographical and cultural isolation from other populations many current residents are direct descendants of several of these founders. Louisiana's incidence of Attention-Deficit Hyperactivity Disorder (ADHD) diagnosis in children is 14.2%, near the highest in the nation. In a pilot project we genotyped natives of southwest Louisiana to identify common VNTR alleles, which have been associated with increased risk of ADHD, among this population. DAT1 encodes the Dopamine Active Transporter (SLC6A3, MIM: 126455), a carrier protein for the reuptake of dopamine from the synaptic cleft in certain regions of the brain and a primary mechanism of dopamine regulation. A common VNTR polymorphism in the 3' untranslated region apparently influences gene expression. Repeats of the 40 base pair sequence range from 3 to 11, but the most common alleles are 10R (71.9%) and 9R (23.4%). The presence of the 10R allele has been associated with increased risk of ADHD. We sampled 80 unrelated, informed, and consenting adults who are descended from one or more of the founders arriving before the beginning of the 20th century. DNA will be isolated from whole blood, cheek cells, or hair samples. Genotyping was performed using PCR amplification and gel electrophoresis. In addition, individuals completed a questionnaire concerning ADHD diagnosis or associated phenotypes and genealogical relationship(s) to founding families. Diagnosis of ADHD was not necessary for participation as our primary interest is in assessing allele frequencies among direct descendants of the population's founders. Amplified samples are retained for sequencing to determine whether these variants are identical by descent as is assumed by the classic founder's effect model or arrived independently via different founding individuals.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3351T

A two-sex model for the admixture history of a hybrid population. *A. Goldberg, P. Verdu, N.A. Rosenberg.* Stanford University, Stanford, CA.

Admixed human populations have often experienced sex-biased admixture, in which males and females contributing to the gene pool of the admixed population are drawn in different proportions from the various source populations. Genetic variation in autosomes and sex-specific chromosomes has often been examined in order to infer patterns of sex-biased admixture, typically using statistical approaches that do not fully model the complexity of a sex-specific history of admixture. Expanding on the model of Verdu & Rosenberg (2011), we develop a model that mechanistically considers sex-specific admixture histories, in which multiple source populations contribute to the admixed population, potentially with varying contributions from male and female lineages at each generation. In an admixed population descended from two source groups, we derive the moments of the distribution of the admixture fraction from a specific source population as a function of sex-specific introgression parameters and time, separately for autosomes and sex chromosomes. Considering admixture processes that are constant in time, we demonstrate that although the mean admixture fraction from a specific source population cannot discern a sex bias in the admixture history for autosomal markers, the variance is informative about sex bias. Specifically, the long-term variance decreases as the sex bias from a contributing source population increases. We highlight a case in which different constant admixture histories produce the same autosomal variance of admixture, but consideration of the X chromosome enables the scenarios to be distinguished. Our approach can contribute to methods for inference of the history of complex sex-biased admixture processes by contrasting the distribution of admixture fractions between autosomal and non-autosomal DNA.

3352F

Admixed human genomes reveal complex migration patterns from the early spread of modern humans to the 21st century. *S. Gravel[1], J.M. Kidd[1,2], J.K. Byrnes[1], A. Moreno-Estrada[1], F. Zakharia[1], S. Musharoff[1], F.M. De La Vega[1], C.D. Bustamante[1].* 1) Genetics, Stanford University, Stanford, CA; 2) Human Genetics, University of Michigan, Ann Arbor, MI.

Populations with the slightest geographical, ethnic, or cultural differentiation often exhibit differences in their patterns of genome-wide diversity due to assortative mating. Thus admixture, defined as gene flow across preferentially mating groups, occurs in most genealogies. Whether it is considered a confounding factor (as in standard GWAS) or is directly used for mapping (as in admixture mapping), better models for admixture and population structure will result in better outcomes for medical genetics studies.

Individuals deriving recent ancestry from multiple continents represent a substantial and increasing proportion of the human population. Because of the tens of thousands of years of divergence between continental populations, we can often infer the contributions from different continental ancestries for individual loci, facilitating the interpretation of diversity patterns. We first discuss a simple Markov model that relates the time-dependent migration history to the inferred patterns of local ancestry. We use this framework to infer the timing of admixture and to differentiate between punctual and continuous models of migration: using demographic models that are consistent with both historical records and genetic data, we find evidence for continuous migration patterns in both Mexican and African-American populations.

We also present a diversity of methods for the analysis of whole-genome sequence data from admixed individuals, and apply them to 50 genomes sequenced by Complete Genomics, including 4 Mexican-Americans, 4 African-Americans and 2 individuals from Puerto Rico, together with SNP genotype data from hundreds of additional samples, as well as data recently generated by the 1000 Genomes project. We consider the allele frequency distribution, pairwise TMRCA's, and a simple extension of the recently introduced Pairwise Sequencial Markov Chain approach for demographic inference. The inferred source population demographic histories are in broad agreement with previous results for European and West-African populations, and the inferred demography for the Native source population closely follows the European one until about 20,000 years ago. Taken together, whole genome sequencing and local ancestry assignment therefore permit inferences about long-term histories of unsampled ancestral populations and highlights recent historical demographic processes that altered patterns of variation observed in admixed populations.

3353W

Quantitative characterization of the ancestry of the Estonian population. *T. Haller[1], L. Leitsalu-Moynihan[1], P.C. Ng[1,2], K. Fischer[1], T. Esko[1], K. Läll[1], A. Metspalu[1].* 1) Estonian Genome Center, University of Tartu, Tartu, Estonia; 2) Genome Institute of Singapore, Singapore.

Ancestry-related questions are increasingly getting more attention owing to the genome-wide marker information becoming available for a growing number of populations. This information can be utilized in a variety of ways ranging from the re-construction of historical events to advancing public health. Separating genetically relatively distant populations has been carried out for years. Recently the methods have gained in sensitivity by adopting haplotype-based approach and now allow comparative studies of closely related populations. The Estonian Genome Center of the University of Tartu maintains a Biobank with over 50,000 DNA and plasma samples as well as a comprehensive phenotype database (objective measurement data, health data, information on lifestyle/habits, metabolite concentrations, gene expression data etc.) for these individuals. This population-based database represents 5% of Estonia's adult population. Currently over 10,000 self-reported Estonians representing all regions of Estonia have been genotyped with genome-wide chips, and the number is expected to grow to 50,000. This allows detailed studies of the Estonian population structure in terms of ancestry and correlating the ancestry information with the phenotype information. We adopted the latest tools (SHAPEIT, ChromoPainter) together with original in-house algorithms to develop a method for quantitatively studying the Estonian population structure. Our method allows to detect what fraction of the genome of each person is inherited from which of the 19 European reference populations we use. Performing a deconvolution of these closely related genetic influences has previously remained a challenge and is addressed by our method. For each person tested the algorithm assigns percent similarity with three of the closest reference populations by performing a best fit simulation. The goodness of fit is used to estimate how much of the genetic background cannot be accounted for by the reference populations. Dissecting how well the modern time ancestry structure reflects the key historical events indicates the geographical mobility of people and displays the genetic fingerprints of the populations that have influenced Estonia throughout the history. Estonia serves as a good model for detailed ancestry studies; it is a small country located in a historically turbulent location and at the same time is a typical representative of a European population.

3354T

Demographic inference and whole genome scan for natural selection in Biaka Pygmies from Central Africa. *P.H. Hsieh[1], K.R. Veeramah[2], J.D. Wall[3], R.N. Gutenkunst[1,4], M.F. Hammer[1,2].* 1) Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ; 2) Arizona Research Laboratories Division of Biotechnology, University of Arizona, Tucson, AZ; 3) Institute for Human Genetics, University of California, San Francisco, CA; 4) Department of Molecular and Cellular Biology, University of Arizona, Tucson, AZ.

African Pygmy groups reside mostly in Central African rainforests and are characterized by their short stature and predominantly hunter-gatherer lifestyle. Many Pygmy populations have been influenced by neighboring Niger-Kordofanian-speaking farmer populations through socio-economic contacts, especially since the extensive agriculture expansion from the Nigeria/Cameroon border to Southern and Eastern sub-Saharan Africa beginning ~3–5 thousand years ago. Two questions of particular interest concern 1) the population demographic dynamics of African Pygmies and the genetic consequences of the agriculture expansion, and 2) the genetic basis of adaptation to the harsh rainforest environment, in which they are exposed to not only hot and humid conditions, but also many parasites and diseases pathogens. We generated whole-genome sequences at >40X coverage for a sample of Biaka Pygmies. As a control, we also examined publicly available high-coverage whole-genome sequences data from two Niger-Kordofanian-speaking farmer populations of normal stature, the Yoruba from Nigeria and the Luhya from Kenya. To unravel the demographic history and to search for signatures of natural selection in these populations, we built a novel computational whole-genome-sequences pipeline integrating population genetic, functional, and statistical analyses. After quality control filtering, our data consist of >3.7 million SNPs that are fully called in all three populations. Through use of the model-based demographic inference tool, ∂a∂i, we found that demographic models involving recent admixture from neighboring non-Pygmy Niger-Kordofanian farmer populations to the Pygmies were significantly better than those that only featured low levels of unidirectional or bidirectional migration after divergence from a common ancestor. In light of recent findings, we also explored the likelihood of introgression between the ancestors of Biaka and archaic humans. Preliminary results from a whole-genome selection scan using multiple complementary tests, such as G2D and XP-CLR, reveal that genes involved in cell adhesion, cellular signaling, olfactory perception, and immunity were likely targeted by natural selection in the Biaka Pygmies or their recent ancestors. The present work not only demonstrates the usefulness of our computational pipeline, but also sheds new light on the history of one of the most enigmatic African populations and on the genetic basis of local adaptation.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3355F

Analysis of contributions of archaic genome and their functions in modern non-Africans. *Y. Hu[1], Q. Ding[1], Y. Wang[1], H. Zheng[1], L. Jin[1,2].* 1) MOE Key Laboratory of Contemporary Anthropology, School of Life Sciences and Institutes of Biomedical Sciences, Fudan University, Shanghai 200433, China; 2) Chinese Academy of Sciences and Max Planck Society (CAS-MPG) Partner Institute for Computational Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, People's Republic of China.

Archaic hominin admixture with modern non-Africans was detected by genome wide analysis of Neanderthal and Denisovan individuals. Compared with Africans, non-Africans share excess derived alleles with Neanderthals. Especially, Europeans and East Asians are similarly related to Neanderthals, suggesting Neanderthal gene flow to Eurasians before their split. Additional Denisovan admixture was detected in Melanesian individuals. Several models were subsequently proposed regarding time and location of admixture events. To gain better understanding in demographic and evolutionary significance of archaic hominin admixture, we implemented an algorithm to identify archaic segments, i.e., carrying locally maximized number of alleles that are only observed in non-Africans, by analyzing 1K Genome Phase I data. Totally, we identified 410,683 archaic segments in 909 non-African individuals with averaged segment length 83,460bp. In the genealogy of each archaic segment with Neanderthal, Denisovan, African and chimpanzee segments, 77~81% archaic segment coalesced first with Neanderthal, 4~8% coalesced first with Denisovan, and 14% coalesced first with neither, validating the algorithm. Interestingly, a large proportion of all the archaic segments identified shared 88.9% similarity with Neanderthal, suggesting a single major admixture with Neanderthal at 82~121kya, right after the Africa exodus of the ancestors of modern humans. Furthermore, we found that several disease associated sites with alleles specific to archaic segments, suggesting a possible contribution of Neanderthal and Denisovans to human diseases.

3356W

Anisotropic isolation by distance: the main orientations of human genetic differentiation. *F. Jay[1,2], P. Sjödin[3], M. Jakobsson[3,4], M.G.B. Blum[2].* 1) Laboratoire TIMC-IMAG UMR 5525, Université Joseph Fourier, Centre National de la Recherche Scientifique, Grenoble, France; 2) Department of Integrative Biology, UC Berkeley, Berkeley, CA; 3) Department of Evolutionary Biology, Uppsala University, Uppsala, Sweden; 4) Science for Life Laboratory, Uppsala University, Uppsala, Sweden.

Genetic differentiation among human populations is greatly influenced by geography due to the accumulation of local allele frequency differences. However, little is known about the possibly different increment of genetic differentiation along the different directions (north-south, east-west, ...). We analyzed genome-wide polymorphism data from African (n=29), Asiatic (n=26), Native American (n=9) and European (n=38) populations, and we found that the major orientations of genetic differentiation are north-south in Europe and Africa, east-west in Asia, but no preferential orientation was found in the Americas. A practical consequence of the anisotropic pattern of genetic differentiation is that the localization of an individual's geographic origin based on SNP data should be more precise along the orientation of maximum differentiation. We compared the localization of geographic origin obtained with principal component regression with a baseline method and confirmed that the largest improvement was obtained along the orientation of maximum differentiation. Our findings have implications for interpreting the making of human genetic variation in terms of isolation by distance and spatial range expansion processes.

3357T

On the use of deterministic functions for deriving approximate coalescent distributions. *E. Jewett, N. Rosenberg.* Biology, Stanford University, Stanford, CA.

Under the coalescent model, the number $n(t)$ of lineages as a function of time is nearly deterministic and is well approximated by its expected value $E[n(t)]$. In turn, this expectation is well approximated by deterministic functions that are simple and easy to compute (Slatkin and Rannala 1997; Maruvka et al., 2011). Such deterministic functions have been applied to the estimation of allele age (Slatkin and Rannala 1997; Slatkin 2000) and the estimation of effective population size (Maruvka et al., 2011). Here, we apply these deterministic functions to a variety of additional problems in estimating population divergence times and population trees, and in modeling faster-than-exponential growth. We find that deterministic approximations to $E[n(t)]$ can be used to derive accurate, computationally fast, and numerically stable approximations to coalescent probability distributions and expectations. These approximations can be very good, even when the number of lineages sampled is small. Such approximations provide alternatives to exact formulas that are computationally intractable or numerically unstable when the number of sampled lineages is moderate or large, making them ideal for applications involving large genotyping or sequencing data sets.

3358F

Exploring Population Admixture Dynamics Using Empirical and Simulated Genome-Wide Distribution of Ancestral Chromosomal Segments. *W. Jin[1], S. Wang[2], L. Jin[1,3], S. Xu[1].* 1) CAS-MPG Partner Institute for Computational Biology, Chinese Academy of Sciences, China; 2) Department of Organismic and Evolutionary Biology, Harvard University, MA; 3) School of Life Sciences, Fudan University, China.

The history and processes of admixture determine the haplotype structure and linkage disequilibrium patterns of the admixed population, which is of significant importance to medical and evolutionary studies such as identification of disease-associated genetic variants and natural selection signatures. However, most previous studies relied on simplified models that do not consider the inherent complexity of admixture process or admixture dynamics. Here we proposed two approaches to explore population admixture dynamics, and we demonstrated the approach based on distribution of chromosomal segments of distinct ancestry (CSDA) was more powerful than that based on the distribution of individual ancestry proportions by analyzing genome-wide empirical and simulated data. Analysis of 1,890 African-Americans showed that a continuous gene flow (CGF) model, in which African-Americans continuously received gene flow from European populations over about 14 generations, best explained the admixture dynamics of African-Americans among several putative models. Interestingly, we observed some African-Americans who had much more European ancestry than those under simulation, indicating substructures of local ancestries in African-American that could be resulted from individuals from some particular lineages having continuously inter-married with people of European ancestry. On the contrary, the admixture dynamics of Mexicans was more likely to be explained by a gradual admixture (GA) model in which Mexicans continuously received gene flow from both European and Amerindian populations for about 24 generations. The higher proportion of long CSDA observed in Mexican-Americans relative to Mexican-Mestizos suggested recent gene flow from European-Americans to Mexican-Americans. The genetic components of Sub-Saharan Africans in Middle Eastern populations such as Mozabite, Bedouin and Palestinian could be explained by one early admixture followed by some recent gene flow from Africans. In summary, this study not only provides new approaches to explore the admixture dynamics of admixed populations, but also advances our understanding of historical admixture processes of African-Americans, Mexicans and Middle Eastern populations.

3359W

Population History and Y-Chromsome Diversity of the Poqomchi' Maya of Eastern Guatemala. *S. Johnson[1], A. Justice[2], M. Crawford[1].* 1) University of Kansas, Lawrence, KS; 2) Department of Epidemiology, University of North Carolina at Chapel Hill.

The Poqomchi' Maya is a group of Native Americans that were peacefully converted to Christianity in the 16th century by the Spanish. This study aims to characterize the source and level of paternal admixture within the Poqomchi' Maya of Eastern Guatemala, while making inferences into their historical relationships to surrounding Central Americans. DNA was extracted from 25 Poqomchi' males and analyzed for Y-Chromosome variation and diversity compared to other Native American populations of Central America. Using SNPs that were characterized with HyBeacons® PCR probes, the presence of Native American haplogroup Q-M242 was determined to have a frequency of 90%, (the highest among all comparative populations), the majority of which, (84%)belonged to the Native American-specific lineage Q1a3a-M3. STRs (DYS-19, 385 a/b, 389 I/II, 390, 391, 392, 393, 438, and 439) were characterized on an Applied Biosystems's ABI Prism 3730 for haplogroup assignment. The Poqomchi' exhibited 25 unique haplotypes within the sample. Interestingly, neither of the Native American haplogroups C-M130 and C3b-P39, was detected. Haplotype diversity, locus diversity, and MPD were computed to characterize the intra-population variation. The Poqomchi' exhibit higher than expected haplotype diversity (0.990) compared to surrounding populations, but average gene diversity (0.569) and MPD (3.98). This difference in diversity measures is likely the result of a high occurrence of low frequency haplotypes. Mismatch analysis for STR data indicate a unimodal distribution for the frequency of pairwise differences, indicating population expansion. Slatkins Rst, R-matrix analysis, and SAMOVA were used to highlight the relationship among populations in Meso- and Central America. Results indicate that little haplotype variation is explained by geography. Despite a long history of colonization by Europeans, population decline, and forced migration, the Poqomchi' have maintained a relative isolation from non-Native populations and therefore exhibit a majority of Native American Y-lineages with little non-native admixture.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3360T

Local Ancestry Inference in the Thousand Genomes Project Admixed Populations. *E.E. Kenny[1], B. Maples[1], C. Churchhouse[2], A. Gupta Hinch[2], A. Williams[3], Y. Baran[4], S. Gravel[1], A. Moreno Estrada[1], C. Gignoux[5], F. Zakharia[1], A.W. Bigham[6], M.D. Shriver[7], A. Ruiz-Linares[8], G. Bedoya[9], T.K. Oleksyk[10], J. Dutil[12], J.C. Martinez-Cruzado[10], E. Gonzalez Burchard[5], S. Myers[2], E. Halperin[11], J. Marchini[2], C.D. Bustamante[1], The 1000 Genomes Consortium.* 1) Department of Genetics, Stanford University, Stanford, CA; 2) Department of Statistics, 1 South Parks Road, Oxford OX1 3TG, UK; 3) Department of Genetics, Harvard Medical School, New Research Building, 77 Ave. Louis Pasteur, Boston, Massachusetts 02115, USA; 4) School of Computer Science, Tel-Aviv University, Tel-Aviv, Israel; 5) Institute for Human Genetics, University of California San Francisco, California, USA; 6) University of Michigan, Ann Arbor, Michigan, USA; 7) Penn State University, University Park, Pennsylvania, USA; 8) Department of Genetics, Evolution and Environment. University College London, UK; 9) University of Antioquia, Medellin, Colombia; 10) Department of Biology, University of Puerto Rico at Mayagüez, Puerto Rico 00680; 11) International Computer Science Institute, Berkeley, California, USA; 12) Department of Biochemistry, Ponce School of Medicine, Ponce, Puerto Rico.

Local Ancestry Inference (LAI) from genome-scale data is an essential analysis step in biomedical genomics of mixed ancestry populations for improved power in association studies, identification of disease genes through admixture mapping and as a biomarker for disease outcome. The task of LAI is to determine the population-specific ancestry of the chromosomes of an individual, from several hypothesized source populations (i.e., a Hispanic/Latino individual may have a combination of European, African and Indigenous American ancestry). Here we present the results of the Thousand Genomes Project (TGP) Local Ancestry Inference Working Group to compare four commonly used LAI methods (LAMP-LD, HAPMIX, RFMIX and MULTIMIX) and produced high-quality LAI calls for Phase 1 TGP data from African-American (ASW), Colombian (CLM), Puerto Rican (PUR) and Mexican (MXL) populations. We first tested the accuracy of the LAI algorithms using `known-truth' simulated ASW and MXL individuals (n=10 each) and using reference panels of European, Indigenous American and African populations (n=60 each). Accuracy for each method is high when ancestral populations are genetically distinct (98.9–99.5% accuracy for ASW) but less accurate when source populations are genetically closer (93.9–95.3% accuracy for MXL). Using diploid ancestry calls at each site, we can define high confidence regions with >= 3 methods agree and improve our per method accuracy (95.5–96.4% and 99.1–99.6% for MXL and ASW, respectively) that capture 97.4% and 99.3% of the MXL and ASW sites. The consensus call at these high confidence regions gives us marginally better accuracy than any one method (96.8% and 99.7%). Applying this strategy to the TGP admixed populations, we obtain consensus calls for 99.5%, 96.5%, 96.4% and 95.2% of the ASW, PUR, CLM and MXL samples, respectively. We note the average proportion ancestries are 0.25:0.75 (European:African) for ASW, 0.47:0.04:0.49, 0.75:0.12:0.13 and 0.67:0.08:0.25 (European:African:Indigenous American) for MXL, PUR and CLM, respectively. Finally, the proportion of the genome that could not be assigned ancestry due to discordant calls between methods was 0.013, 0.049, 0.033 and 0.037 for ASW, MXL, PUR and CLM, respectively, and was strongly correlated to the proportion of short ancestry tracts (<5cM) per individual. Consensus LAI for all TGP admixed individuals are available from the TGP Phase 1 release ftp repository.

3361F

A Likelihood Ratio Test for Sex-Bias and its Application to Whole-Genome Sequencing Data of a Set of Globally-Distributed Populations. *S. Musharoff[1], J.M. Kidd[1,2], B.M. Henn[1], S. Gravel[1], B. Maples[1], K. Eilertson[3], M.C. Yee[1], H.M. Cann[4], G. Euskirchen[1], M. Snyder[1], C.D. Bustamante[1], S. Ramachandran[5].* 1) Genetics, Stanford Univ Sch Medicine, Stanford, CA; 2) Human Genetics, University of Michigan, Ann Arbor MI; 3) Bioinformatics Core, J. David Gladstone Institutes, San Francisco CA; 4) Ecology and Evolutionary Biology, Brown University, Providence RI; 5) Fondation Jean Dausset - Centre d'Étude du Polymorphisme Humain, Paris France.

The human X chromosome presents a unique opportunity to study the effect of population genetic forces. Because the X chromosome is carried in two copies in females and one copy in males, its effective population size is determined by the ratio of breeding males to females, and, as a result, the X chromosome is sensitive to sex-biased demographic processes (i.e. those involving unequal numbers of breeding males and females). Previous estimates of the level of sex-bias in human populations from Y chromosome and mitochondrial data, or from X chromosome and autosomal data, found evidence for both male bias and female bias, which may have occurred at different times during human history. Here we present a likelihood ratio test for sex-bias based on the inference of demographic parameters from genomic sequence data. Simulation studies enable us to assess the sensitivity of this test, to determine the timescales on which sex biases can be detected, and to assess the effect of sequencing coverage. We previously sequenced to moderate coverage the whole genomes of 53 individuals from 7 diverse human populations (KhoeSan, Mbuti-Pygmy, Mozabite, Pathan, Cambodian, Yakut, and Maya) drawn from the HGDP-CEPH Diversity Panel. To aid in variant calling, we have sequenced the exomes of the same individuals to high coverage. To apply our test for sex-bias to these populations, we first correct each population's site-frequency spectrum (SFS) to account for genomic variation missed due to sequencing technology using an empirical correction based on the exome sequences. We infer demographic histories from the SFS using the program dadi and we see striking differences between the observed patterns of X chromosome variation and those that are expected under the null hypothesis of no sex-bias. We perform our likelihood ratio test for sex-bias on each population separately, as well as jointly. To assess the power of this method, we use the forward simulator SFScode for a wide range of demographic scenarios including that of sex-biased admixture between modern humans and archaic humans. This characterization of the effect of demography on the X chromosome enables us to study the effects of selection, the mode of which may differ from that on the autosomes due to the X chromosome's hemizygosity in males. The test for sex-bias presented here is not limited to human or to whole-genome data and will enable demographic insights beyond those obtainable from the autosomes alone.

3362W

Population Structure and Genetic Diversity in Argentinean populations. *M. Muzzio[1,2,3], J.M.B. Motti[2,3], E. Kenny[1], L.S. Jurado-Medina[2], M.C. Yee[1], J. Beltramo[2], R. Santos[2,3], T. Cooke[1], V. Ramallo[4], M. Schawb[2], O. Cornejo[1], G. Bailliet[2], C.M. Bravi[2,3], C.D. Bustamante[1].* 1) Stanford University School of Medicine, Stanford, CA; 2) IMBICE CCT-La Plata CONICET-CICPBA, Argentina; 3) Facultad de ciencias Naturales y Museo, Universidad Nacional de La Plata, Argentina; 4) Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil.

Argentina was one of the last places peopled when humans entered the Americas. Its earliest archeological sites are dated from ~13K years b.p., and throughout its history the region has exhibited broad cultural diversity, involving both hunter-gatherer and agro-pastoral societies, different linguistic families, the Southern expansion of the Inca Empire, the arrival of the Spanish conquistadors and slave trade from Africa. More recently, Argentina has had the highest proportion of influx of transoceanic migration between ~1850 and 1940 of any Latin American country. Furthermore, the geographical distribution of immigrants was oriented towards the Pampa region of central-eastern Argentina and the Northeast, so different patterns of mixed ancestry throughout the country are expected. In this present work we aim to describe the history and structure of the Argentinean population. Thus far, we have collected DNA samples from 2904 participants from 15 dispersed regions in Argentina, from the Andean populations in the Northwest to the river lands of the Northeast. Our preliminary data includes genome-scale Illumina Exome Array 250K genotypes, plus next-generation sequencing data from uniformly spaced genomic regions targeted with a genotype-by-sequencing strategy (which covers 1.5% of the genome), for 375 participants from 9 of the 15 populations. We will identify proportions and patterns of European, African and Indigenous American ancestry in the Argentinean cohort via local ancestry inference. The first series of analyses will focus on genomic tracts of Indigenous American ancestry for fine-scale examination of sub-continental structure by comparing to a panel of 493 samples from 52 Indigenous American populations from throughout the Americas. Furthermore, this work will serve as a test of different low-cost strategies (exome genotype array vs. genotype-by-sequencing) for accurate inference of population structure in extant populations.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3363T

Rediscovering the people who greeted Columbus through mtDNA sequence analysis. *M.A. Nieves-Colon[1], E.P. Tascón-Peñaranda[2], M. Diaz-Matallana[3], I. Carrero-Gonzalez[4], F. Curbelo-Canabal[2], J. Galanther[5], M.L. Perez-Quiñones[2], H.J. Diaz-Zabala[6], C. Eng[5], O.R. Wever[7], F. Mercedes de la Cruz[8], A. Alvarez-Serrano[9], E.G. Burchard[5], J.C. Martínez-Cruzado[2].* 1) School of Human Evolution & Social Change, Arizona State University, Tempe, AZ; 2) Department of Biology, University of Puerto Rico at Mayagüez, Mayagüez, PR; 3) Instituto de Genética Humana, Pontificia Universidad Javeriana, Bogotá, Colombia; 4) Baylor College of Medicine, Houston, Texas; 5) Institute of Human Genetics, University of California, San Francisco, California; 6) Ponce School of Medicine, Ponce, Puerto Rico; 7) Department of Internal Medicine, Dr. Horacio Oduber Hospital, Oranjestad, Aruba; 8) Facultad de Ciencia y Tecnología, Universidad Central del Este, San Pedro de Macorís, Dominican Republic; 9) Museo Arqueológico Regional Altos de Chavón, La Romana, Dominican Republic.

In an effort to understand the Pre-Columbian migrations that brought the people found by Columbus to the Greater Antilles, we sequenced the mtDNA control region of 504 admixed people from the Caribbean with Native American maternal ancestry. These included 140 samples from Puerto Rico (PR), 138 samples from Honduras, and 120 samples from the Dominican Republic (DR). In addition, we obtained 49 samples from Colombia, 46 from Aruba, and a smaller number of samples from other populations in the region. Haplogroups A2 and C1 predominated in both PR and DR with joint frequencies of 88.0% and 83.1%, respectively. At the same time, median joining network analyses showed few lineages that were shared between the islands, probably as a result of limited gene flow combined with genetic drift. With the exception of two cosmopolitan control region lineages, the lineages shared between the islands always showed diversity in PR but not in DR, suggesting migrations occurred mostly in the east-to-west direction. Non-cosmopolitan lineages on both islands showed affinity with either Central or South America, with the exception of one Dominican lineage that appeared in both Central America and Aruba. In PR, where lineages were well represented, those with Central American affinity held higher diversity, suggesting that the islands were peopled first from Central America and later from South America. However, it was South American and not Central American lineages that presented signs of population expansion, consistent with the archaeological evidence that suggested the arrival of agro-ceramic cultures to the Greater Antilles from South America approximately 2,500 YBP. Consequently, most of the autosomal Native American ancestry in PR and possibly DR may be of South American origin. Complete mtDNA sequences of both a Honduran lineage related to DR, and a Venezuelan lineage related to PR, showed differences with their island counterparts. Further examination and additional samples will be necessary to refine the number, geographic origin, and the migration times for the peopling of the Greater Antilles.

3364F

Continent-wide decoupling of Y-chromosomal genetic variation from language and geography in native South Americans. *M. Nothnagel[1], L. Roewer[2], L. Gusmao[3,5], V. Gomes[3], M. Gonzalez[3], D. Corach[4], A. Sala[4], E. Alechine[4], T. Palha[5], N. Santos[5], A. Ribeiro-dos-Santos[5], M. Geppert[2], S. Willuweit[2], M. Nagy[2], S. Zweynert[6], M. Baeta[7], C. Nunez[7], B. Martinez-Jarreta[7], F. Gonzalez-Andrade[8], E. Fagundes de Carvalho[9], D. Aparecida da Silva[9], J. Jose Builes[10,11], D. Turbon[12], A. Maria Lopez Parra[13], E. Eduardo Arroyo-Pardo[13], U. Toscanini[14], L. Borjas[15], C. Barletta[16], S. Santos[5], M. Krawczak[1].* 1) Institute of Medical Informatics and Statistics, Christian-Albrechts University, Kiel, Germany; 2) Institute of Legal Medicine and Forensic Sciences, Charité - Universitätsmedizin Berlin, Berlin, Germany; 3) Institute of Molecular Pathology and Immunology, University of Porto, Porto, Portugal; 4) Laboratório de Genética Humana e Médica, Universidade Federal do Pará, Belém, Pará, Brazil; 5) Facultad de Farmacia y Bioquimica, Servicio de Huellas Digitales Geneticas, Universidad de Buenos Aires, Buenos Aires, Argentina; 6) Department of Psychiatry and Psychotherapy, Charité - Universitätsmedizin Berlin, Berlin, Germany; 7) Department of Forensic Medicine, University of Zaragoza, Zaragoza, Spain; 8) Science and Technology Department, Ministry of Public Health, Quito, Ecuador; 9) Laboratorio de Diagnósticos por DNA, Instituto de Biologia, Universidade do Estado do Rio de Janeiro, Brazil; 10) GENES Ltda., Lab. Genética Forense y Huellas Digitales del DNA, Medellín, Colombia; 11) Instituto de Biología, Universidad de Antioquia, Medellín, Colombia; 12) Unitat d'Antropologia, Dept. Biologia Animal, Facultat de Biologia, Universitat de Barcelona, Barcelona, Spain; 13) Laboratorio de Genetica Forense, Dpto. de Toxicología y Legislación Sanitaria, Fac. de Medicina, Universidad Complutense de Madrid, Madrid, Spain; 14) PRICAI-Fundación Favaloro, Buenos Aires, Argentina; 15) Laboratorio de Genetica Molecular, Unidad de Genetica Medica, Facultad de Medicina, Universidad del Zulia, Maracaibo, Venezuela; 16) Laboratorio de Genética Humana, Facultad de Ciencias Biológicas, UNMSM-Universidad, Nacional Mayor de San Marcos, Lima, Peru.

The way of initial habitation creates a primordial spatial pattern of human genetic variation that is subsequently modified by demographic processes. Under the assumption that most such changes follow trajectories set by climate as well as geographic and cultural conditions, some correlation between the genetic structure on the one hand, and the linguistic and geographical structure on the other, is to be expected in extant human populations even a long time after an initial colonization event and has been documented in numerous studies on European and Asian populations. In contrast, there is a notable absence of such descriptions for South America. Here, we examined Y-chromosomal genetic diversity and its relation to geographic location and linguistic classification on a continental scale in the so far largest study on South American natives, involving up to 17 microsatellite markers genotyped in a total sample of 1011 individuals representing 50 tribal populations from 81 settlements as well as single-nucleotide polymorphisms defining the founding native American phylogenetic lineages Q and C and its sub lineages. We observed a large decoupling of Y-chromosomal genetic diversity from geographical habitats and in parts from language groups on a continental scale, which is consistent with a rapid peopling of the continent and subsequent long periods of isolation of relatively small-sized tribal groups. Our results highlight the fact that a pronounced correlation between genetic and geographic/cultural structure can be expected only under very specific conditions, most of which are likely not to have been met by native South Americans.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3365W

Using Haplotypes to Infer Ancestral Origins For Recently Admixed Individuals. *K. Noto, J.K. Byrnes, N. Myres, N. Angerhofer, S.R. Woodward, C.A. Ball, K. Chahine.* AncestryDNA, Provo, UT, USA.

Identification of individual ancestral origins from genetic data is useful in many fields including admixture analysis and medical genetics. Several approaches to this task have been developed, and most rely on allele frequency differences observed at single nucleotide polymorphisms (SNPs) to assign individuals to one or more ancestral source populations. In most cases, these methods assume marker independence, treating each SNP as an independent observation. With the advent of genotyping arrays in which millions of SNPs are typed, neighboring SNPs are frequently close enough to be in linkage disequilibrium (LD). In this case the alleles observed at neighboring SNPs are strongly correlated due to shared genetic history. Using this type of data requires LD thinning to remove linked pairs of SNPs and satisfy the independence assumption. Unfortunately LD thinning also removes significant amounts of information in the data, reducing assignment accuracy. This is particularly problematic in high resolution analyses, such as identifying countries of origin within Europe. Here we explore the possibility of using phased haplotypes in place of SNPs for ancestry inference. This work is motivated by our initial observation that each SNP contributes more than twice as much ancestry information when it is part of a haplotype. What is more, some previously developed ancestry inference methods such as FRAPPE accept multi-allelic data, readily lending themselves to use with haplotypes. To examine the effect of using haplotypes on ancestry inference, we apply an algorithm based on FRAPPE to estimate ancestry proportions from SNPs alone, and then from phased haplotypes built from the same SNP set. We then compare these estimates to known sample origins. This process is repeated for two European data sets with single-origin individuals; POPRES and an unpublished set of 1000 European samples with detailed pedigrees. In both cases, all four grandparents for each sample are confirmed to be from the same country. Next, using single-origin individuals as ancestors, we use a forward simulation to construct individuals with recent admixture between various European populations and compare estimated ancestry from haplotypes or SNPs to the known mixture proportions. Finally, we assemble sets of ancestry informative markers (AIMs) from SNPs and haplotypes and assess the performance of these selected sets.

3366T

Haplotype sharing distributions for fine-scale demographic reconstruction. *P. Palamara[1], T. Lencz[2], A. Darvasi[3], I. Pe'er[1,4].* 1) Department of Computer Science, Columbia University, New York, NY; 2) The Zucker Hillside Hospital Division of the North Shore-Long Island Jewish Health System, Glen Oaks, NY, USA; 3) Division of Psychiatry Research, The Zucker Hillside Hospital, The Feinstein Institute for Medical Research, North Shore - Long Island Jewish Health System, New York, NY; 4) Center for Computational Biology and Bioinformatics, Columbia University, New York, NY.

Identical-by-descent (IBD) chromosomal segments shared by pairs of individuals can now be reliably detected in large genomic datasets of purportedly unrelated samples, and summary statistics of IBD sharing were recently shown to convey information about population-level features such as demography, natural selection and heritability of common traits. However, while key theoretical quantities of IBD sharing have been extensively investigated in the context of pedigree structures, limited literature is available to describe IBD haplotypes in pairs of unrelated individuals from arbitrary populations. We present analytical results for the relationship between haplotype sharing and demographic history. We express the distribution of IBD sharing across pairs of individuals for segments of arbitrary length as a function of the population's demography. We derive a maximum-likelihood procedure for the inference of demographic parameters, and evaluate it using several datasets of 500 synthetic individuals. Analyzing a single chromosome, we reconstruct the population size within 3% of the true size in 95% of Wright-Fisher populations. For exponentially expanding or contracting populations, the reconstructed size is within 10% of the true size for the past 100 generations. We analyze genome-wide SNP data for 500 Ashkenazi Jewish individuals. The most likely scenario we reconstructed reveals two periods of expansion, separated by a strong bottleneck. A first period of mild growth ($r \approx .02$) leads to a population of ~60,000 individuals in the 12th century A.D. Following a severe bottleneck, the reconstructed scenario includes ~300 individuals rapidly ($r \approx .28$) expanding to ~4,000,000 in the past 32 generations. This demographic profile is consistent with historical accounts for the Ashkenazi population. We further analyze 56 Kenyan Masai individuals from the HapMap 3 dataset. High levels of cryptic relatedness are pervasive in this sample. While this may suggest a strong ($r \approx -.14$) recent bottleneck in this population, we show that a model of multiple small-sized demes interacting through high migration rates, consistent with the social structure of the Masai, results in a compatible pattern of haplotype sharing. This mismatch between single-population models and historically structured demographies may explain the observed abundance of runs of homozygosity within several outbred populations. An implementation of the described methods, DoRIS, is publicly available.

3367F

Inference of population splits and mixtures from genome-wide allele frequency data. *J.K. Pickrell[1], J.K. Pritchard[2,3].* 1) Department of Genetics, Harvard Medical School, Boston, MA; 2) Department of Human Genetics, University of Chicago, Chicago, IL; 3) Howard Hughes Medical Institute, University of Chicago, Chicago, IL.

Many aspects of the historical relationships between populations in a species are reflected in genetic data. Inferring these relationships from genetic data, however, remains a challenging task. In this paper, we present a statistical model for inferring the patterns of population splits and mixtures in multiple populations. In this model, the sampled populations in a species are related to their common ancestor through a graph of ancestral populations. Using genome-wide allele frequency data and a Gaussian approximation to genetic drift, we infer the structure of this graph. We applied this method to a set of 55 human populations and a set of 82 dog breeds and wild canids. In both species, we show that a simple bifurcating tree does not fully describe the data; in contrast, we infer many migration events. While some of the migration events that we find have been detected previously, many have not. For example, in the human data we infer that Cambodians trace approximately 16% of their ancestry to a population ancestral to other extant East Asian populations. In the dog data, we infer that both the boxer and basenji trace a considerable fraction of their ancestry (9% and 25%, respectively) to wolves subsequent to domestication, and that East Asian toy breeds (the Shih Tzu and the Pekingese) result from admixture between modern toy breeds and "ancient" Asian breeds. Software implementing the model described here, called TreeMix, is available at http://treemix.googlecode.com.

3368W

Inferring Y Chromosome Phylogeny by Sequencing Diverse Populations. *G.D. Poznik[1,2], P.A. Underhill[3], B.M. Henn[3], M.C. Yee[3], E. Sliwerska[4], G.M. Euskirchen[3], L. Quintana-Murci[5], E. Patin[5], M. Snyder[3], J.M. Kidd[4], C.D. Bustamante[3].* 1) Program in Biomedical Informatics, Stanford University School of Medicine, Stanford, CA; 2) Department of Statistics, Stanford University, Stanford, CA; 3) Department of Genetics, Stanford University School of Medicine, Stanford, CA; 4) Department of Human Genetics, University of Michigan, Ann Arbour, MI; 5) Human Evolutionary Genetics Unit, Centre National de la Recherche Scientifique URA3012, Institut Pasteur, Paris, France.

The male-specific region of the Y chromosome (MSY) harbors the longest stretch of non-recombining DNA in the human genome and is therefore a unique tool that enables the tracking of migrations and inference of demographic history. We have sequenced 69 male samples from nine globally diverse populations, including three African hunter-gatherer groups. Due to inefficient selection, a relatively high mutation rate, and a small effective population size, the Y chromosome is particularly subject to drift. It has accumulated large expanses of highly repetitive sequence, which pose considerable challenge within a short read sequencing paradigm. To overcome this hurdle, we have built an informatics pipeline to reliably call Y chromosome alleles from moderate coverage short read shotgun sequence data. First, we defined a callability mask, learned from the mapping quality and depth of coverage patterns in the data, and then we tuned base-pair level quality control thresholds. Based on 13,000 provisional SNP calls, we inferred a tree of the 69 sequenced Y chromosomes. Using this tree, we then called individual genotypes for each SNP with a custom-built, phylogeny-aware, EM algorithm. With these high quality calls in hand, samples were assigned haplogroup labels using standard YCC nomenclature; 29 distinct named haplogroups were represented. We find that the maximum likelihood tree we construct recapitulates the extant Y chromosome phylogeny, thus confirming the fruits of decades of work based on ascertained SNPs. Further, we resolve a major long-standing polytomy by identifying a variant for which one haplogroup retains the ancestral allele, whereas its brother clades share the derived allele, thus indicating common ancestry and uniting the latter two branches. This finding has been confirmed by genotyping a larger panel. Finally, we estimate the MSY rate of mutation recurrence and the time to the most recent common ancestor of the sampled chromosomes.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3369T

Towards understanding the population substructure and ancestry in Kuwaiti population. *R. Rajagopalan, O. Alsmadi, F. Alkayal, G. Thareja, S. John, P. Hebbar, A. Thangavel.*Dasman Genome Centre Dasman Diabetes Institute PO BOX 1180 Dasman, Kuwait 15462.

Owing to high rate of consanguinity and well known genealogy in most families, the population of Kuwait offers a great opportunity in genetic epidemiology to map recessive components of monogenic and complex disorders as well. While much has been published in rare monogenic disorders from this region, we present here for the first time the genetic makeup of Kuwaiti population in the context of genome-wide runs of homozygosity (ROH) and Identity-By-Descent (IBD) among unrelated individuals. 124 unrelated Kuwaiti individuals were genotyped using Illumina Human OmniExpress platform and more than 700,000 SNPs were used in this analysis. We used fastIBD algorithm in Beagle software [Browning & Browning, 2011] to find shared genomic regions that are IBD in this dataset and found 13,097 segments greater than 2 Mb and average length of a segment was 4.4 Mb. There were 716 long IBD segments (>10 Mb) between 360 unique pairs of individuals and the average IBD length in that subset was 23.4 Mb. 119 out of 124 individuals had at least one IBD segment that is greater than 10 Mb. We then used PLINK [Purcell S et. al. 2007] to find 3074 ROH segments that are greater than 2 Mb with mean ROH size of 3 Mb. Total ROH length ranged from 11 Mb to 533 Mb with a mean of 75 Mb and number of segments ranged from 8 to 79 with a mean of 25. To find common ROHs, we calculated the frequency of each SNP falling into an ROH segment using a sliding window approach and found 23 segments in chromosomes 1, 10 and 17 with greater than 30% of individuals having an ROH, of which 10 with greater than 40%. We also inferred regions that are homozygous-by-descent (HBD) using Beagle to compare with ROH statistics from PLINK. Though, the number of ROH segments seems to be overestimated (mostly due to error in genotype calls), the total amount of HBD is nearly the same (r2 = 0.99). Results from IBD and ROH analyses suggest the existence of both recent and ancient shared ancestry among Kuwaiti population placing them between outbred and population isolates. This could well be used in advantage to better understand the population substructure of Kuwait and in general the Arabs. With high prevalence of complex diseases like Obesity and Diabetes in the region, we hope this will also serve as a tool to find recessive components of complex diseases and may help explain some of the missing heritability.

3370F

The evolutionary history of mutations associated with leukemia. *S. Ramachandran[1], M. Kantesaria[1], A. Walia[1], J. Yang[2].* 1) Ecology & Evolutionary Biol, Brown Univ, Providence, RI; 2) Pharmaceutical Sciences, St. Jude Children's Research Hospital, Memphis, TN.

Studies of worldwide human variation have repeatedly observed genome-wide signatures predicted by simulations of a serial founder effect, in which populations expand outward from Africa through a process where each population in the expansion is formed from a subset of the previous population in the expansion. Two of the genome-wide signatures are a decrease in heterozygosity at multiallelic markers with increasing distance from Africa and increasing linkage disequilibrium with increasing distance from Africa (Jakobsson et al. 2008, DeGiorgio et al. 2009). Genome-wide association studies (GWAS) of acute lymphoblastic leukemia have observed that both incidence and response to therapy exhibit racial disparities. A recent study confirmed that response to therapy co-segregates with Native American ancestry, even after adjusting for known prognostic factors (Yang et al. 2011). We examine haplotype variation at this and other loci in the genome using genome-wide data. We find that the evolutionary history of mutations associated with leukemia is consistent with a history of hitchhiking along with beneficial alleles in linkage disequilibrium with the risk alleles found through GWAS.

3371W

Identification of regions under selection in the Gullah African American population. *P.S. Ramos[1], S. Sajuthi[2], Y. Huang[3], J. Divers[2], K.M. Kaufman[4], J.B. Harley[4], R.P. Kimberly[5], D.L. Kamen[1], C.D. Langefeld[2], M.M. Sale[3], W.T. Garvey[6], G.S. Gilkeson[1].* 1) Department of Medicine, Medical University of South Carolina, Charleston, SC; 2) Department of Biostatistical Sciences and Center for Public Health Genomics, Wake Forest School of Medicine, Winston-Salem, NC; 3) Department of Medicine and Center for Public Health Genomics, University of Virginia, Charlottesville, VA; 4) Division of Rheumatology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH; 5) Department of Medicine, University of Alabama at Birmingham, Birmingham, AL; 6) Department of Nutrition Sciences and Birmingham VA Medical Center, University of Alabama, Birmingham, AL.

PURPOSE: The Gullah are a unique population of African ancestry in the U.S. In addition to their genetic and environmental homogeneity and low European admixture (less than 3.5%), a shorter genetic distance between the Gullah and Sierra Leonean tribes has also been reported, suggesting that population genetic signals, such as regions under recent selection, may be more easily detected in the Gullah than in other African American (AA) populations. Given the increased prevalence of many common diseases in AA and the increasing evidence of selection at loci associated with human diseases, identification of alleles under selection may provide insight into disease susceptibility. Since population-specific selection may cause allele frequency differences, the goal of this study was to identify regions with minor allele frequency (MAF) differences between Gullah and Sierra Leoneans. METHODS: We had available 271 Gullah and 400 Sierra Leonean samples genotyped on the Illumina Human1M and Affymetrix Genome-Wide Human SNP Array 6.0, respectively. After stringent QC was applied to each population, 185,569 SNPs common to both arrays with MAF>5% were used to compute the significance of the MAF differences between the two populations. In order to exclude MAF differences that are spurious or assay artifacts, only regions where at least two SNPs in LD showed significant Bonferroni-adjusted MAF differences were considered. RESULTS: The largest region showing significant MAF differences between the Gullah and Sierra Leoneans was the MHC at 6p21.32, where multiple SNPs show P-values between E-06 and E-13, and has previously been shown to be under selection in other populations. The region showing the most significant MAF differences was that of the complement-regulatory CUB and Sushi multiple domains 1 precursor gene (CSMD1) at 8p23.2, where several SNPs show P-values between E-06 and E-25. Interestingly, this gene has been associated with, among other traits, emphysema, multiple sclerosis, insulin resistance in AA and coronary heart disease in AA. CONCLUSION: We have identified several regions with significant allele frequency differences between the Gullah and Sierra Leoneans, suggesting that population-specific selective pressures may be operating at these loci. Given the increased prevalence of many complex traits in AA and the homogeneity of the Gullah, identification of these regions in the Gullah has the potential to elucidate disease risks in AA.

3372T

A genomewide map of Neandertal ancestry in modern humans. *S. Sankararaman[1,2], N. Patterson[2], S. Mallick[1,2], S. Pääbo[3], D. Reich[1,2].* 1) Harvard Medical School, Boston, MA; 2) Broad Institute of MIT and Harvard, Cambridge, MA; 3) Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany.

Analysis of the genomes of archaic hominins, such as Neandertals and Denisovans, has revealed that these groups have contributed to the genetic variation of modern human populations. Yet, we know little about how these ancient mixtures have shaped the genetic structure of human populations and even less about their impact on human evolution. To answer these questions, we need a map of archaic ancestry.

Building such a map is technically challenging because of the antiquity of these gene flow events. We have developed a method based on the statistical framework of Conditional Random Fields (CRFs) that integrates single SNPs as well as haplotypes informative of ancient gene flow to give highly accurate predictions.

We applied this method to polymorphism data in European and East Asian individuals from the 1000 genomes project, in conjunction with the draft sequence of the Neandertal genome, to obtain the first genome-wide map of Neandertal ancestry. Analysis of this map leads to several novel findings:

1. We identify around 35,000 Neandertal-derived alleles in Europeans and 21,000 in East Asians.

2. We also identify over 100 Neandertal-derived alleles that are likely to have been the target of selection since introgression. One of these has a frequency of about 85% in Europe and overlaps CLOCK, a key gene in Circadian function in mammals. This gene has been found in other selection scans in Eurasian populations, but has never before been linked to Neandertal gene flow.

3. Several of the Neandertal-derived alleles identified are found among the >6,000 SNPs associated with common diseases listed in the NHGRI catalog. This list of Neandertal-derived variants include a risk variant associated with obesity and a protective variant against breast cancer.

4. Using the uncorrelated ancestries in Europe and East Asia, we can reconstruct about 600 Mb of the genome of the introgressing Neandertal.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3373F

Genetic diversity and migration history of indigenous populations from the Central Valley of Mexico. *T. Schurr[1], M. Vilar[1], R. Gomez[2], H. Zillges[1], A. Sanders[1], A. Owings[1], J. Gaieski[1], The Genographic Consortium.* 1) Dept Anthropology, Univ Pennsylvania, Philadelphia, PA; 2) Dept Toxicologia, CINVESTAV Mexico City, Mexico.

Home to the Aztec Empire, the Central Valley of Mexico was the most densely populated area of North America upon the arrival of Europeans in the early 16th century. Today, more than a million indigenous people from a number of distinct linguistic groups inhabit this region, including the Otomi (Oto-Manguean family) and the Nahua, or former Aztec (Uto-Aztecan family). Although archeologically well-studied, this complex cultural and linguistic region remains genetically understudied.

In order to assess the genetic diversity of indigenous Mexicans and infer prehistoric patterns of interaction among Mesoamerican groups, we genotyped 536 participants from twenty-three communities in four Mexican states: Guanajuato, Hidalgo, Morelos and Queretaro. Genotyping included mtDNA haplogroup and haplotype identification through control region sequencing and coding-region SNP analysis; Y-chromosome haplogroup designation through SNP and STR typing; and admixture assessment through analysis of autosomal AIMs and STRs. The resulting genotypes were compared to those of dozens of other Native American populations.

MtDNA analyses of the Otomi revealed a very high frequency (>98%) and great diversity of the four major Native American haplogroups (A2, B2, C1, D1), and also the rare D4h3. The Nahua belonged mostly to haplgroups A (75%) and B (15%). Y-chromosome analysis also showed high frequency and haplotype diversity for indigenous haplogroup Q (>70%) among the Otomi, while the Nahua were nearly 100% for that haplogroup. In addition, autosomal analysis revealed a high percentage (>75%) of indigenous AIMs in both groups. The observed genetic patterns suggest that the two populations remained genetically differentiated, even after living in proximity for several millennia. The high frequency and diversity of haplogroup A2, as compared to that in other indigenous American populations, also suggests the Central Valley may have been the center of growth and expansion for that pan-American haplogroup, one possibly associated with the domestication and spread of Mesoamerican crops such as maize. The high frequency of both haploid and bi-parental indigenous markers further suggests that the Central Valley experienced limited African and European genetic introgression in the 500 years since Spanish colonization.

3374W

Unexpected relationships and inbreeding in HapMap phase III populations. *E. Stevens[1], J. Baugher[2], M. Shirley[2], L. Frelin[3], J. Pevsner[3,4].* 1) Human Genetics, Johns Hopkins Univ Sch Med, Baltimore, MD; 2) Biochemistry, Cellular, and Molecular Biology, Johns Hopkins Univ Sch Med, Baltimore, MD; 3) Neurology, Hugo W. Moser Research Institute at Kennedy Krieger, Baltimore, MD; 4) Psychiatry and Behavioral Sciences, Johns Hopkins Univ Sch Med, Baltimore, MD.

Correct annotation of the genetic relationships between samples is essential for population genomic studies that could be biased by errors or omissions. To this end, we used identity-by-state (IBS) and identity-by-descent (IBD) methods to assess genetic relatedness of individuals within HapMap Phase III. We analyzed data from 1,397 individuals from 11 ethnic populations. Our results support previous studies (Pemberton et al., 2010; Kyriazopoulou-Panagiotopouloue et al., 2011) assessing unknown relatedness present within this population. Additionally, we present evidence for 1,657 novel pairwise relationships across 9 populations. Surprisingly, significant IBD1 values were estimated between known parents. Furthermore, significant IBD2 values were estimated in 32 previously annotated parent-child relationships. Consistent with a hypothesis of inbreeding, regions of homozygosity (ROH) were identified in which a subset overlapped those reported in previous studies (Gibson et al. 2010; Johnson et al. 2011) in the offspring of related parents. In total, we inferred 25 inbred individuals with ROH that overlapped areas of relatedness between the parents and/or IBD2 sharing at a different genomic locus between a child and a parent. Finally, 8 previously annotated parent-child relationships had unexpected IBD0 values (indicative of a deletion or genotype errors), and 10 previously annotated second-degree relationships along with 38 other novel pairwise relationships had unexpected IBD2 (indicating two separate paths of recent ancestry). These newly described types of relatedness may impact the outcome of previous studies and should inform the design of future studies relying on the HapMap Phase III resource.

3375T

Class I MHC region genetic markers have different dinamics than neutral genomic regions in Brazilian populations. *R.C.P. Toledo[1], A.C. Arcanjo[1], C.T. Mendes-Júnior[2], E.C. Castelli[3], J.A. Peña[4], R.V. Andrade[5], M.N. Klautau-Guimarães[1], S.F. Oliveira[1].* 1) Programa de Pós-Graduação em Biologia Animal, Instituto de Ciências Biológicas, Universidade de Brasília, Brasília, Distrito Federal, Brazil; 2) Departamento de Química, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, São Paulo, Brazil; 3) Departamento de Biologia Geral, Instituto de Ciências Biológicas, Universidade Federal de Goiás, Goiânia, Goiás, Brazil; 4) Departamento de Genetica y Antropología Física, Facultad de Ciencia y Tecnología, Universidad del País Vasco, Bilbao, Spain; 5) Pró-Reitoria de Pós-Graduação e Pesquisa, Universidade Católica de Brasília, Brasília, Distrito Federal, Brazil.

Brasília and Kalunga are Brazilian Midwest populations with different settlement history as well as genetic constitution. Kalunga is a rural afro-descendent population, established in the 19th century, formed by runaway afro-descendent slaves, whereas Brasília is an urban population, established in 1960 with the transfer of the capital of Brazil from the Southeast region. These populations were previously analyzed with neutral genetic markers such as AIMs, autosomic STRs, DNAmt and chromosome-Y markers which have shown, solely and altogether, that those are two distinct populations, with very different genetic structure. It has been suggested that the MHC region of the human genome might be under selection, and therefore do not behave as neutral regions. In this work, a set of loci in the Class I MHC region was analyzed to verify whether both those populations behave equally regarding markers under selection. The set of markers included five Alu insertions (MICB, TF, HJ, HG and HF) comprising the whole Class I MHC, and a 360bp sequence in the 3'UTR region of the HLA-G gene. Statistical analysis included linkage disequilibrium, haplotypic inference, population differentiation, Hardy-Weinberg Equilibrium and neutrality tests. Eight different SNPs were observed in the 3'UTR region: indel 14pb, +3003C/T, +3010G/C, +3027 A/C, +3035 C/T, +3142 G/C, +3187 G/A and +3196 G/C. All 13 loci in both populations were in HWE and Kalunga showed heterozygote excess for the AluHG locus. Brasília showed 11 3'UTR haplotypes and 21 Alu insertion haplotypes, while Kalunga presented 9 3'UTR haplotypes and 18 Alu haplotypes. The Ewens-Watterson neutrality test showed significant values only for the 3'UTR region (Brasília: F=-1.4140, $p<0.05$; Kalunga: F=-1.6354, $p<0.05$). Attempting to confirm these results, Tajima's D and Fu & Li's F tests were estimated for the 3'UTR region, but only Brasília showed significant values (D=2.1419, $p<0.05$; F=1.8458, $p<0.05$). Population differentiation test (FST) showed similarity between Brasília and Kalunga for the whole (FST = 0.01526) and separate sets (3'UTR: 0.01320; Alu: 0.01916) of markers. Therefore, there is indication of ongoing selection in the HLA-G 3'UTR, but not in the Class I MHC Alu insertions. Also, both populations are similar regarding the Class I MHC region, contrasting with results previously described using genetic markers located within genomic neutral regions. Acknowledgment: CAPES, CNPq and FAPDF.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3376F

Coalescence-based Estimation of Population History in the Presence of Admixture from Genetic Variation Data. *M. Tsai[1], G. Blelloch[2], R. Ravi[3], R. Schwartz[4].* 1) Computational Biology, Carnegie Mellon University, Pittsburgh, PA; 2) Department of Computer Science, Carnegie Mellon University, Pittsburgh, PA; 3) Tepper School of Business, Carnegie Mellon University, Pittsburgh, PA; 4) Department of Biological Science, Carnegie Mellon University, Pittsburgh, PA 15238.

Learning how modern human populations have arisen, dispersed, and intermixed since we emerged as a species is an important but challenging problem. Learning precisely how the human population is structured and how different subgroups are related to one another is not only a fundamental issue in understanding human origins, but also has practical relevance to improving models of genome evolution and better controlling for population substructure in association studies. Despite considerable attention to the general problem of identifying population substructure in large-scale variation data, the field lacks automated methods for reconstructing the relationships among population subgroups and the process by which such substructure has emerged over time (the population history). One especially challenging problem in such inference is inferring correct orders and timing of events in the presence of admixture. We describe here a novel two-step approach for joint inference of a quantitative population history in the presence of admixture from large variation datasets. The method first identifies a set of phylogenetic splits that are likely have occurred during the evolutionary history and assign each observed variation site to one of the splits. The resulting split set and the number of variation sites assigned to each split are then used to infer a model of population-level evolution describing times of the divergence and admixture events as well as admixture proportions using a coalescence-based Markov chain. To evaluate the algorithm, we generated a series of simulated data sets consisting of combinations of parental and admixed populations. On simple models of two parental and one admixed population, our algorithm reliably identifies the true population model and provides reasonable estimates of event times and admixture parameters. Performance degrades on more complicated admixture scenarios, but still shows at least 70% accuracy for four-population, two-admixture scenarios at identifying the correct model as the maximum likelihood scenario. The results suggest that automated reconstruction of complex population histories in the presence of admixture from genome-scale data is feasible, although further algorithmic improvements may be needed to infer more complicated scenarios.

3377W

The GenoChip: a new tool for genetic anthropology. *S. Wells[1], E. Greenspan[2], S. Staats[2], T. Krahn[2], C. Tyler-Smith[3], Y. Xue[3], S. Tofanelli[4], P. Francalacci[5], F. Cucca[6], L. Pagani[7], L. Jin[8], H. Li[8], T.G. Schurr[9], J.B. Gaieski[9], C. Melendez[9], M.G. Vilar[9], A.C. Owings[9], R. Gomez[10], R. Fujita[11], F. Santos[12], D. Comas[13], O. Balanovsky[14], E. Balanovska[14], P. Zalloua[15], H. Soodyall[16], R. Pitchappan[17], G. Arun Kumar[17], M.F. Hammer[18], B. Greenspan[2], E. Elhaik[19].* 1) Mission Programs, National Geographic Society, Washington, DC; 2) Family Tree DNA, Houston, TX; 3) Wellcome Trust Sanger Institute, Hinxton, UK; 4) University of Pisa, Italy; 5) University of Sassari, Italy; 6) National Research Council, Monserrato, Italy; 7) University of Cambridge, UK; 8) Fudan University, Shanghai, China; 9) University of Pennsylvania, Philadelphia, PA; 10) CINVESTAV, Mexico City, Mexico; 11) University of San Martin de Porres, Lima, Peru; 12) Federal University of Minas Gerais (UFMG), Belo Horizonte, Brazil; 13) Pompeu Fabra University, Barcelona, Spain; 14) Russian Academy of Medical Sciences, Moscow, Russia; 15) Lebanese American University, Byblos, Lebanon; 16) University of the Witwatersrand, Johannesburg, South Africa; 17) Chettinad Academy of Research and Education, Chennai, India; 18) University of Arizona, Tucson, AZ; 19) Johns Hopkins University, Baltimore, MD.

Background: The Genographic Project is an international effort aimed at charting human history using genetic data. The project is non-profit and non-medical, and through the sale of its public participation kits it supports cultural preservation efforts in indigenous and traditional communities. To extend our knowledge of the human journey, interbreeding with ancient hominins, and modern human demographic history, we designed a genotyping chip optimized for genetic anthropology research. **Methods:** Our goal was to design, produce, and validate a SNP array dedicated to genetic anthropology. The GenoChip is an Illumina HD iSelect genotyping bead array with over 130,000 highly informative autosomal and X-chromosomal SNPs ascertained from over 450 worldwide populations, ~13,000 Y-chromosomal SNPs, and ~3,000 mtDNA SNPs. To determine the extent of gene flow from archaic hominins to modern humans, we included over 25,000 SNPs from candidate regions of interbreeding between extinct hominins (Neanderthal and Denisovan) and modern humans. To avoid any inadvertent medical testing we filtered out all SNPs that have known or suspected health or functional associations. We validated the chip by genotyping over 1,000 samples from 1000 Genomes, Family Tree DNA, and Genographic Project populations. **Results:** The concordance between the GenoChip and the 1000 Genomes data was over 99.5%. The GenoChip has a SNP density of approximately (1/100,000) bases over 92% of the human genome and is highly compatible with Illumina and Affymetrix commercial platforms. The ~10,000 novel Y SNPs included on the chip have greatly refined our understanding of the Y-chromosome phylogenetic tree. By including Y and mtDNA SNPs on an unprecedented scale, the GenoChip is able to delineate extremely detailed human migratory paths. The autosomal and X-chromosomal markers included on the GenoChip have revealed novel patterns of ancestry that shed a detailed new light on human history. Interbreeding analysis with extinct hominids confirmed some previous reports and allowed us to describe the modern geographical distribution of these markers in detail. **Conclusions:** The GenoChip is the first genotyping chip completely dedicated to genetic anthropology with no known medically relevant markers. We anticipate that the large-scale application of the GenoChip using the Genographic Project's diverse sample collection will provide new insights into genetic anthropology and human history.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3378T

Genetic, Genomic and Phenotypic Characterization of Qinghai Mongolians. *J. Xing[1, 2], U. Tanna[3], T.S. Simonson[2, 4], W.S. Watkins[2], D.J. Witherspoon[2], W. Wu[2], G. Qing[3], C.D. Huff[2, 5], L.B. Jorde[2], R.L. Ge[3].* 1) Department of Genetics, Rutgers, the State University of New Jersey, Piscataway, NJ; 2) Department of Human Genetics, University of Utah, Salt Lake City, UT; 3) Research Center for High Altitude Medicine, Qinghai University Medical School, Xining, Qinghai, People's Republic of China; 4) Department of Genetics, University of California in San Diego, San Diego, CA; 5) Department of Epidemiology, the University of Texas M.D. Anderson Cancer Center, Houston, TX.

Qinghai Mongolians (QH Mongolians) migrated from the Mongolian steppes to the Qinghai-Tibetan Plateau approximately 500 years ago. Living at about 3000 meters, QH Mongolians have been challenged by extreme environmental conditions similar to neighboring highland Tibetans, including metabolic demands imposed by diet, cold, and to a lesser degree, hypobaric hypoxia. Identification of adaptive genetic factors in this population could provide insight into coordinated physiological responses to these extreme conditions. In this study, we examined genetic and phenotypic variation in 42 QH Mongolian individuals. Using high-density SNP array data, we demonstrate that QH Mongolians represent a distinct population from the neighboring Tibetan, Buryat, and Han Chinese populations (Fst = 0.008 in all cases). We then performed a genome-wide selection scan in QH Mongolians and compared the selection candidates with other Mongolian and Tibetan populations. Our analyses identified a number of selection candidates that are shared between Mongolians or between Mongolians and Tibetans, including genes involved in body composition and hypoxia-related selection candidate genes previously identified in Tibetans. To understand the genomic variation and identify potential functional variants in QH Mongolians, we performed whole-genome sequencing on a QH Mongolian individual (Tianjiao1). Mitochondrial DNA (MtDNA) and Y-chromosome lineage assignment put Tianjiao1 in the H18 and Q1a1 haplogroups, respectively. Overall, we found more than 2% of the variants in the Tianjiao1 genome are not present in other populations, including more than 1,000 coding variants. Lastly, to determine the phenotypic variation in QH Mongolians, we collected a number of phenotypic traits, including levels of hemoglobin (Hb) and several metabolites. QH Mongolians have a significant lower Hb level compared to Tibetans, and the putative selected Tibetan haplotypes showed no significant association with Hb level for QH Mongolians. Overall, our study showed that QH Mongolians share genetic ancestry with Tibetan populations as well as other Mongolian populations. They present a good candidate population for studying adaptation to extreme conditions.

3379F

Paleolithic human migrations in East Eurasia by sequencing Y chromosomes. *S. Yan[1,2], C.C. Wang[1], L. Jin[1,2].* 1) MOE Key Lab of Contemporary Anthorpology, Fudan University, Shanghai, China; 2) Computational Genetics, CAS-MPG Partner Institute for Computational Biology, Shanghai, China.

Paleolithic human migrations in East Eurasia remains largely unknown due to the lack of sufficient markers derived from the mutations that occurred during that time frame. To tackle this problem, using the sequence capturing, barcoding technology and next-generation sequencing, we identified more than 4,000 new SNPs encompassing most single copy non-recombining region of human Y chromosome. New clades for haplogroups O, C, N, D, and Q could be geographically located. Especially, a few star-like expansions were unveiled, showing strong population growth. The phylogeny of Haplogroup N was radically rearranged, and all the N individuals could now be categorized into either a northern clade N1 or southern clade N2, revealing a Paleolithic migratory routes of the ancestors of Uralic speaking populations. Haplogroup C, especially the East Eurasia-dominant clade C3, could also be separated into at least two ancient clades, suggesting Paleolithic migrations in East Asia. Three major clades under O, M117+, M134xM117, and 002611+, each could be now further classified into several subclades. With these new findings, we proposed the modified the routes and dates for human populations' migration, especially those in Paleolithic time. A few Y-chromosomal expansions could now be linked to certain prehistoric cultures or ancestors of language families.

3380W

Reconstructing the history of Indigenous American populations from the Thousand Genomes Project admixed populations of the Americas. *F. Zakharia[1], M. Muzzio[1, 2, 3], K. Sandoval[1], A. Moreno-Estrada[1], S. Gravel[1], J. Byrnes[4], C.R. Guignoux[7], J.L. Rodriguez-Flores[6], B. Maples[1], W. Guiblet[5], J. Dutil[8], E.E. Kenny[1], A.W. Bigham[10], M.D. Shriver[11], . The 1000 Genomes Consortium[12], A. Ruiz-Linares[9], T.K. Oleksyk[5], J.C. Martinez-Cruzado[5], E. Gonzalez Burchard[7], C.D. Bustamante[1].* 1) Department of Genetics, School of Medicine, Stanford University, Stanford, CA, USA; 2) Facultad de Ciencias Naturales, Universidad Nacional de La Plata, 1900, Argentina; 3) IMBICE CCT-La Plata, CONICET, 1900, Argentina; 4) Ancestry.com; 5) Department of Biology, University of Puerto Rico at Mayagüez, Puerto Rico; 6) Department of Genetic Medicine, Weill Cornell Medical College, New York, NY, USA; 7) Institute for Human Genetics, University of California San Francisco, CA, USA; 8) Ponce School of Medicine, Puerto Rico; 9) Department of Genetics, Evolution and Environment. University College London, UK; 10) University of Michigan, Ann Arbor, MI, USA; 11) Penn State University, University Park, PA, USA; 12) The 1000 Genomes Project.

The Americas exhibit a rich cultural and biological diversity, including many countries with different languages, and have been deeply affected by historical events such as the arrival of the Europeans in the late 15th century. Yet little is known about the original peopling of the Americas, or about the interactions that may have taken place between the early settlers. The Thousand Genomes Project (TGP) has generated low coverage whole genome sequence, high coverage exome sequence and high quality genotype calls for three admixed populations from the Americas: Puerto Rico (PUR), Colombia (CLM), and Mexico (MXL). This dataset not only provides novel insights into these admixed populations, but also offers a unique opportunity to explore the genetics of their respective Indigenous American (IA) ancestors. To this end, we performed local ancestry deconvolution to identify tracts of IA ancestry in each of the admixed individuals in the sample. We found that the PUR, CLM and MXL populations contain an average 12%, 24%, and 47%; IA ancestry, respectively, and that 99.9%; of genomic sites are covered by at least one IA chromosome, with an average coverage of 17X, 35X, and 54X IA chromosomes for the PUR, CLM, and MXL respectively. We developed ancestry-specific (AS) extensions of both PCA and ADMIXTURE (PCA-AS and ADMIXTURE-AS) to analyze population structure considering only the IA components of ancestry in the admixed TGP genomes. We trained these methods using a large dataset comprising 493 participants from 52 IA populations across the Americas, genotyped at high density. We found that IA components in admixed individuals cluster with different linguistic groups, with the MXL and CLM appearing to be closest to Northern Amerind and Chibchan-Paezan speaking groups respectively, and the PUR near Equatorial-Tucanoan speaking populations endemic to the Amazon. Finally, we used the joint site frequency spectra of the admixed populations to make novel inferences about the demographic history of their IA ancestors. Preliminary results estimate an initial split of the MXL ancestors from the others around 12.5 kya, followed by the divergence of the PUR and CLM ancestors around 8 kya. We also found evidence of strong bottlenecking events in the PUR and CLM lineages, consistent with the low levels of heterozygosity observed in these populations.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3381T

The genetic aspect of the Slavic settlement in the East Alps region. *A. ZUPAN[1], K. VRABEC[1], T. DOV{{Cak}} DRNOV{{Sak}}EK[2], D. GLAVA{{Cak}}[1].* 1) Department for Molecular Genetics, Institute for Pathology, Faculty of Medicine, University of Ljubljana, Ljubljana, Slovenia; 2) Blood Transfusion Centre of Slovenia, Ljubljana, Slovenia.

The territory of East Alps region experienced turbulent historical events which are reflected in different archeological cultures found in this area. The archeology findings and historical writings from the sixth century A.D. revealed a sharp social transition from Roman-type social structure to unimpressive pan-Slavic social structure. The main reason for the social transition is contributed to the degradation of the Roman Empire and to the Slavic settlements. However, the historical process of the Slavic settlement in the territory of East Alps and the impact it had on the Romanized indigenous population was not convincingly resolved by the archaeological, ethnological, topological and linguistic studies. In order to provide a genetic aspect on the Slavic settlement of the East Alps region and to reconstruct past demographic events, we analyzed 42 Y-chromosome biallelic markers and 17 Y-chromosome short tandem repeats (Y-STR) in 399 Slovenians from five geographical regions. The biallelic genetic markers revealed 29 haplogroups, with the most common being R1a1a* (M198), R1b1b1a1 (xL11), I2a1 (M423) and I1* (M253). To obtain phylogenetic relationships between different populations, the biallelic and Y-STR data were compared with the data obtained from eleven Slavic and three non-Slavic populations across central Europe and Balkan. Two different Slavic clusters emerged from principal component analysis (PCA), based on frequency of bialellic markers and from multi dimensional scaling analysis (MDS), based on Y-STRs. One cluster encompassed West-Slavic groups, together with the population of Slovenia and the second encompassed the Balkan states. The genetic affiliation of Slovenian population with West-Slavic groups was additionally confirmed by the analysis of molecular variance (AMOVA) which showed the highest genetic variation among groups, when the Slovenian population was grouped together with other West-Slavic populations. The genetic data suggests a dominant North Slavic migration route into the area of East Alps region, through a process of demic diffusion, which allowed cohabitation and subsequent assimilation of indigenous Romanized population. The population genetic results from this study, when combined with archeological findings and historical writings, provide a new insight into the process of Slavic migrations into the territory of East Alps region and beyond.

3382F

Principal component analysis produces an inaccurate representation of population structure under certain conditions. *J. Bentham[1], D.L. Morris[1], D.S. Cunninghame Graham[1], C.L. Pinder[1], V. Anand[1], J.E. Wither[2], J.D. Rioux[3], M.E. Alarcón-Riquelme[4,5], T.J. Vyse[1].* 1) Department of Medical and Molecular Genetics, Division of Genetics and Molecular Medicine, King's College London, UK; 2) Toronto Western Hospital, Main Pavilion, 399 Bathurst St, Toronto, Ontario, M5T 2S8, Canada; 3) Montreal Heart Institute, Université de Montréal, 5000 rue Belanger, Montreal, Quebec, H1T 1C8, Canada; 4) GENYO, Centro de Genómica e Investigación Oncológica Pfizer, Universidad de Granada, Junta de Andalucía, Avenida de la Ilustración 114, 18007 Granada, Spain; 5) Arthritis and Clinical Immunology Program, Oklahoma Medical Research Foundation, Oklahoma City, 73104, USA.

The successes of genome-wide association studies have produced an explosion in the number of loci associated with genetically inherited diseases. Whilst many results have been replicated robustly, other loci identified appear to be false positives. Although false positive associations may be caused by various sources of systematic error, our results relate to one well-known cause of inflation of test statistics, population structure; we show that when using principal component analysis (PCA), individuals with very different ancestries may appear together, having values of zero for each principal component.

Population structure has long been recognised as a potential cause of false positives; study designs mean that subjects may be heterogeneous in ancestry, and where there are both different allele frequencies and disease prevalence between groups of subjects, spurious associations with disease may be identified. One method used to counteract this problem is to carry out PCA and use the first few principal components as covariates in logistic regression; this method relies on the principal components accurately reflecting population structure. We suggest that a proportion of false positive associations are caused by a deterministic problem encountered when applying PCA to genotype data.

We show that for the high dimensional datasets produced by genome-wide studies, PCA can fail to identify gross population outliers. Under certain conditions, the eigenvectors of subjects from different continents are superimposed on the origin of the principal component space; this is an effect of the curse of dimensionality, a set of phenomena associated with high dimensional spaces. The complete loss of information describing the ancestry of these subjects causes unacceptable inflation of the Type I error rate; counterintuitively, this effect is more likely to occur as the number of SNPs used to calculate principal components increases.We provide two examples of the effect: the first uses a dataset taken from a large genome-wide association study of systemic lupus erythematosus; the second uses simulated data. We explain how the outliers may be identified and removed; in our lupus GWAS the inflation factor was reduced from 1.9 to 1.1 by removing 200 samples found at the origin in our initial PCA.

3383W

The Kaiser Permanente/UCSF Genetic Epidemiology Research Study on Adult Health and Aging: Local ancestry estimations in the admixed individuals of the Genetic Epidemiology Research on Aging (GERA) cohort. *S.I. Candille[1], N.A. Johnson[1], Y. Banda[2], T. Hoffman[2], M. Kvale[2], C. Sabatti[3], S. Hesselson[2], P.Y. Kwok[2], C. Schaefer[4], N. Risch[2,4], H. Tang[1].* 1) Department of Genetics, Stanford University School of Medicine, Stanford, CA; 2) Institute for Human Genetics, University of California, San Francisco, CA; 3) Department of Health Research and Policy, Stanford University School of Medicine, Stanford, CA; 4) Kaiser Permanente Research Program on Genes, Environment and Health, Kaiser Permanente Northern California Division of Research, Oakland, CA.

Individuals with recent ancestors from different continents have admixed genomes composed of discernible blocks of sequence inherited from genetically distinct populations. Knowledge of these blocks, their length and position within the genome enables population genetic inferences of the timing and characteristics of admixture events. Admixture mapping uses the information contained in admixed genomes to identify trait loci, for which risk variants occur at different frequencies between ancestral populations. The Genetic Epidemiology Research on Adult Health and Aging (GERA) cohort, part of the Kaiser Permanente Research Program on Genes, Environment and Health (RPGEH), is a cohort of over 100,000 individuals recruited in Northern California. While the majority of subjects are of European ancestry, the cohort also comprises a large number of individuals presenting genetic admixture between continents including ~3000 African Americans, ~8000 Latinos, and ~1560 Filipinos. To develop a resource for population genetics and genotype/phenotype correlation studies in the GERA cohort admixed groups, we used the software SABER+ to estimate local genetic ancestry at a set of genome-wide polymorphisms for each individual. These local ancestry estimates can be used to perform admixture mapping of quantitative traits and we present our results on the mapping of several important traits displaying ethnic variation in these groups. The ethnically diverse GERA cohort represents a unique opportunity to extend the benefits of genetic research until now largely focused on populations of European descent to several admixed groups.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3384T

Multi-way admixture deconvolution using phased or unphased ancestral panels. *C. Churchhouse, J. Marchini.* Statistics, Univ Oxford, Oxford, United Kingdom.

Large collaborative projects have provided a resource of high quality genetic data on a variety of populations across the world, from a wider range of geographic locations than previously sampled including populations that are known to be admixed. An increasing number of disease association studies are being conducted in admixed individuals, such as African Americans and Latinos, and the variation in ancestry within these groups is being utilised to carry out admixture mapping. In such practices, it is important to be able to accurately decompose the population of origin along the genome of an individual with a mixed ancestral background. We have a developed a method named MULTIMIX that can perform ancestry inference in admixed individuals. MULTIMIX is novel in possessing all of the following features (a) the method can handle any number of ancestral populations, making it especially useful for analysing individuals with ancestry from three or more populations, (b) it is applicable to a dense set of genome-wide linked SNPs, (c) it will make use of phased ancestral haplotypes if they are available, but is also applicable to unphased panels of ancestral genotypes, (d) the admixed samples may be analysed as either phased or unphased, (e) the statistical model underlying our method that describes the distribution of the haplotype/genotype given the ancestry of the admixed sample makes MULTIMIX very fast to implement and capable of handling whole-genome analysis even in the case of more than two ancestral populations. We illustrate our approach using simulated data and show that we obtain up to 99.79% correct ancestry calls in European and African admixed samples and up to 99.42% in European, African and Chinese samples. We also demonstrate our method applied to real data from the HGDP and 1000 Genomes project and show that it calls 94.3% of the genome of the Mexican individuals of HapMap3 with a confidence of at least 95%.

3385F

A panel of ancestry informative markers for the South African Coloured population. *M. Daya[1], L. Van der Merwe[2], M. Möller[1], E. Hoal[1], C.R. Gignoux[3], J.M. Galanter[3], B.M Henn[3].* 1) Human Gen & Molecular Biol, Stellenbosch Univ, Cape Town, Tygerberg, South Africa; 2) Biostatistics Unit, Medical Research Council, Tygerberg, South Africa; 3) University of California San Francisco, San Francisco, California, United States of America.

Admixture is a well known confounder in genetic association studies. If genome-wide data is not available, as would be the case for candidate gene studies, ancestry informative markers (AIMs) are required in order to adjust for admixture. Previous work has shown that the South African Coloured population has 5 ancestral populations: African San, African non-San, European, South Asian and East Asian. Consequently, none of the existing published AIM sets can be applied to the population. Using genome-wide data to find SNPs with large allele frequency differences between ancestral populations, as quantified by Rosenberg et. al's *In* statistic, we developed a panel of AIMs for the South African Coloured population by experimenting with various selection strategies. Subsets of different sizes were evaluated by measuring the correlation between ancestry proportions estimated by each AIM subset with ancestry proportions estimated using genome-wide data. Since the accuracy of subsets converge with increasing size, we used linear mixed-effects modelling to select the optimal subset of markers.

3386W

A scalable pipeline for local ancestry inference using thousands of reference individuals. *C.B. Do, E. Durand, J.M. Macpherson, B. Naughton, J.L. Mountain.* 23andMe, Inc, Mountain View, CA.

Ancestry deconvolution, the task of identifying the ancestral origin of chromosomal segments in admixed individuals, is straightforward when the ancestral populations considered are sufficiently distinct. To date, however, no approaches have been shown to be effective at distinguishing between closely related populations (e.g., within Europe). Moreover, due to their computational complexity, most existing methods for ancestry deconvolution are unsuitable for application in large-scale settings, where the reference panels used contain thousands of individuals.

We describe Ancestry Painting 2.0, a modular three-stage pipeline for efficiently and accurately identifying the ancestral origin of chromosomal segments in admixed individuals. In the first stage, an out-of-sample extension of the BEAGLE phasing algorithm is used to generate a preliminary phasing for an unphased, genotyped individual. In the second stage, a support vector machine (SVM) using a specialized string kernel assigns tentative ancestry labels to short local phased genomic regions. In the third stage, an autoregressive pair hidden Markov model simultaneously corrects phasing errors and produces reconciled local ancestry estimates and confidence scores based on the SVM labels.

We compiled a reference panel of over 7,500 individuals of homogeneous ancestry, derived from a combination of several publicly available datasets and over 5,000 individuals reporting four grandparents with the same country-of-origin from the customer database of the personal genetics company, 23andMe, Inc, and excluding outliers identified through principal components analysis (PCA). In cross-validation experiments, Ancestry Painting 2.0 achieves high sensitivity and specificity (in most cases >90%) for labeling chromosomal segments across over 20 different populations worldwide. We also demonstrate the robustness of the algorithm via simulations of individuals of known local admixture, and compare Ancestry Painting 2.0 with existing state-of-the-art tools for multi-population local and global ancestry inference, including LAMP, ALLOY, PCA-ADMIX, and ADMIXTURE.

3387T

Global genome-wide variations comparisons show a generally homogeneous Levant that has been recently structured by culture. *M. Haber[1,2], D. Comas[1], P.A. Zalloua[2,3].* 1) Institut de Biologia Evolutiva (CSIC-UPF), Departament de Ciències de la Salut i de la Vida, Universitat Pompeu Fabra, 08003 Barcelona, Spain; 2) The Lebanese American University, Chouran, Beirut, Lebanon; 3) Harvard School of Public Health, Boston, MA 02215, USA.

Background: The Levant is a region in the Middle East with an impressive record of continuous human existence and major cultural developments that has been well documented since the Paleolithic. Genetic and archeological studies present solid evidence in placing the Middle East and the Arabian Peninsula as the first stepping grounds outside Africa. There is however little understanding of how the Middle East, and more precisely the Levant, was populated since the first Out of Africa Expansion, and how the Levantines genetically relate to each other and to their neighbors. **Methods:** We analyze newly generated genome-wide data from 1,341 subjects from Lebanon with complete demographic information in the context of already published data of 994 individuals from 48 global populations. We implement a series of cluster and structure-like analyses on unlinked SNPs (more than 200,000 SNPs), we also construct a co-ancestry matrix using information from haplotypes phased from more than 500,000 linked SNPs. We use this information to generate a novel population relationship tree and to identify genetic components specific to the Levantines. **Results:** We show that fine stratifications within the Levant are highly influenced by the religious affiliation of the populations within the region. We find a Levantine substructure that is distinctly split into two ancestral branches; one sharing more genetic characteristics with modern day Europeans and Central Asians while the other sharing more genetic characteristics with other Middle Easterners and Africans. Finally, we identify a Levantine component that has diverged from other Middle Easterners ~23,700–15,500 years ago during the last glacial period, and diverged from Europeans ~15,900–9,100 years ago at the start of the Neolithic. **Conclusions:** We show for the first time that the Levantines are today genetically closer to Europeans than to other Middle Easterners. We found that the Levantines and Europeans diverged at the start of the Neolithic age, giving support to a replacement model of Neolithic expansion rather than upper Paleolithic continuity of Europeans. Finally, we show that although population movements and expansions during the Epipaleolithic marked the emergence of a Levantine component and made the Levantines genetically similar, recent cultural developments, such as the inception and spread of religion, have had an impact on fine-scale population stratifications in the Levant.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3388F

Human and *Helicobacter pylori* Ancestries in Colombian Populations with Contrasting Risks of Gastric Cancer. *N. Kodaman[1], A. Pazos[2, 8], M.B. Piazuelo[2], B.G. Schneider[2], C.L. Shaffer[3], J. Romero-Gallo[2], T. de Sablet[2,7], L.E. Bravo[4], K. Wilson[2,5,7], T. Cover[3,6,7], P. Correa[2], S. Williams[1].* 1) Center for Human Genetics, Vanderbilt University, Nashville, TN; 2) Division of Gastroenterology, Department of Medicine, Vanderbilt University School of Medicine, Nashville, TN; 3) Department of Pathology, Microbiology and Immunology, Vanderbilt University School of Medicine, Nashville, TN; 4) Department of Pathology, Universidad del Valle School of Medicine, Cali, Colombia; 5) Department of Cancer Biology, Vanderbilt University School of Medicine, Nashville, TN; 6) Division of Infectious Diseases, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN; 7) Veterans Affairs Tennessee Valley Healthcare System, Nashville, TN; 8) Universidad de Nariño, Pasto, Colombia.

Helicobacter pylori, a chronic gastric pathogen that infects more than half of all humans, is the principal cause of gastric cancer, the second leading cause of cancer mortality worldwide. *H. pylori* infects over 90% of some Colombian populations, with individuals living in the Andean mountain region 25 times more likely to develop gastric cancer than their coastal counterparts, despite similar rates of infection. *H. pylori* strains undergo frequent recombination and thus often contain the genetic signatures of multiple ancestral populations. Previous studies have used the Bayesian clustering algorithm STRUCTURE to uncover six such populations, believed to have evolved in geographical isolation. Using multi-locus sequencing data from those studies and from *H. pylori* isolates obtained from gastric biopsies of 188 Colombian subjects, we found that Colombian strains derive mainly from three previously characterized ancestral *H. pylori*populations: ancestral Africa1 (AA1), ancestral Europe1 (AE1), and ancestral Europe2 (AE2), with most isolates containing genomic regions from all three. Distribution of strain ancestry varied with geography, with AA1 the major ancestral cluster in coastal region samples (mean=50.3%, N=78), and AE1 the major cluster in mountain region samples (mean=51.2%, N=110). We also used STRUCTURE to analyze 8181 loci from an Immunochip study of the biopsied individuals to assess human ancestry. Distributions of human ancestry also varied with geography, with mean proportions of 10.4% European, 54.4% African, and 35.2% Amerindian ancestry in the coastal region, and 13.4% European and 85.8% Amerindian ancestry in the mountain region. Proportion of Amerindian ancestry correlated strongly with the proportion of AA1, AE1, and AE2 ancestry in biopsy isolates ($r = -0.64$, $r=0.58$, $r=0.42$, respectively, with $p<0.001$ for all three). This is the first study to quantitatively assess human and *H. pylori* ancestry in a large number of Colombians. Our findings provide insight into the gastric cancer risk discrepancy between geographically stratified Colombian populations.

3389W

Deep ancestry within a Neandertal-like region on human chromosome 12. *F.L. Mendez[1], J.C. Watkins[2], M.F. Hammer[1].* 1) Arizona Research Laboratories, University of Arizona, Tucson, AZ; 2) Department of Mathematics, University of Arizona, Tucson, AZ.

The importance of the archaic ancestry of modern humans is becoming a topic of increasing interest. The identification of introgressive sequences for which a Neandertal ancestry can be assigned with confidence is essential to correctly characterize the process of introgression, understand the phenotypic consequences of introgression and study the population structure of archaic and modern humans. With a new method of analysis that provides high confidence estimates for sequences with high error rates we identify a region in chromosome 12 that has introgressed from Neandertals into modern humans, we characterize its geographic distribution and observe levels of intra-allelic variation. Furthermore, we analyze how levels of divergence change along the sequence. Consistent with previous expectations we find that the introgressive allele is mostly restricted to Eurasians, exhibits low levels of diversity and has introgressed recently into modern humans. In addition, we observe that segments of very deep divergence suggest a complex demographic scenario for the ancestry of Neandertals.

3390T

Sub-continental ancestry of admixed Caribbean populations and implications for medical genomics in the Americas. *A. Moreno Estrada[1], S. Gravel[1], M.L. Cuccaro[2], F. Zakharia[1], P. Ortiz-Tello[1], R.J. Martinez[2], J.L. McCauley[2], E.E. Kenny[1], D.J. Hedges[2], R.W. Morris[2], J.K. Byrnes[1], S. Acevedo[3], P.J. Norman[4], Z. Layrisse[5], P. Parham[4], C.D. Bustamante[1], E.R. Martin[2].* 1) Genetics, Stanford University, Stanford, CA; 2) Center for Genetic Epidemiology and Statistical Genetics, University of Miami Miller School of Medicine, Miami, USA; 3) Department of Biology, University of Puerto Rico, Mayaguez, PR; 4) Department of Structural Biology, Stanford University School of Medicine, Stanford, USA; 5) Center of Experimental Medicine "Miguel Layrisse", Venezuelan Research Institute (IVIC), Caracas, Venezuela.

Particular dynamics of admixture in the Americas have imprinted a unique genomic pattern in present day Hispanic/Latino populations. A more detailed understanding of how 500 years of admixture have shaped the genomes of Latino populations today is critical to empower medical genetic studies in admixed populations and provide important knowledge about their demographic history. While Mexican populations have been genomically characterized to some extent, genetic studies of populations from the Caribbean and South America have been largely underrepresented. To that end, we have densely genotyped 339 samples from parent-offspring triads of Caribbean-descent sampled in South Florida, US, including Puerto Ricans, Cubans, Dominicans, Haitians, Hondurans and Colombians. We combined these SNP data with other publically available genomic resources, including HapMap and 1000 Genomes data, as well as Native American SNP data from putatively ancestral populations. Global ancestry proportions, estimated using the ADMIXTURE clustering algorithm, revealed substantial variation in admixture proportions among and within Caribbean populations. To assign local, ancestry-specific haplotypes across the genome, we implemented a novel PCA-based admixture deconvolution approach (PCADMIX) for three ancestral populations. Given the optimal accuracy obtained from trio phasing, we were able to determine exact posterior probabilities along each chromosome and measure the length of tracts attributable to distinct ancestries. We observed patterns consistent with continuous influx of African ancestry across the islands and a pulse migration across populations in contact with the continental landmass. By contrast, we find less evidence for continuous flow of Native American ancestry (characterized by shorter ancestry tracts), compatible with a single migration event between 15–20 generations ago and no further contribution of Native Americans into the Caribbean. We have also used local ancestry estimates to scan the genome for ancestry-specific enriched regions potentially indicative of natural selection. Insights gained from the application of new statistical methods for admixture deconvolution will improve the design and genetic analysis of medical genomic studies in Hispanic/Latinos and other complex admixed groups, broadening the ensemble of populations included in current catalogs of human variation and contributing to reduce health disparities across the globe.

3391W

Population structure and ancestry inference under sample selection bias. *S.S. Shringarpure, E.P. Xing.* School of Computer Science, Carnegie Mellon University, Pittsburgh, PA.

Genotyping studies for detecting population structure are often constrained by economic considerations, consent etc. As a result, the genotyped sample may not be statistically representative of the underlying populations. A similar situation is encountered when samples from multiple studies are combined in an attempt to increase statistical power or for better resolution in detecting population structure. This problem is known as sample selection bias and it affects analyses by leading to results which are not consistent with the larger underlying populations. We examined the effects of different levels of sample selection bias on the performance of individual ancestry inference in simulated genotype data. Our experiments show that sample selection bias adversely affects the results of population structure analysis and ancestry inference. Our experiments on data from the Human Genome Diversity Project also show that sample selection bias can result in population structure results which can lead to different interpretations of human evolutionary history. We propose a correction for sample selection bias that is easy to implement and is effective in practice.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3392T

Genome-wide effects of sex-specific incompatibility on neutral introgression. *M. Uyenoyama.* Dept Biol, Duke Univ, Durham, NC.

A widely-used paradigm holds that while selection affects variation in a region-specific manner, all regions of the genome experience the same demographic history. This view justifies identification of outlier segments as candidates for regions that are associated, through tight linkage or epistasis, with targets of selection. On the contrary, we show that strong, sex-specific selection induced upon the introduction by migration of variants adapted to other environments generates region-specific barriers to neutral introgression. For example, an autosomal incompatibility factor that induces more severe effects in male than in female carriers can induce a higher barrier against introgression of Y chromosomes than unlinked autosomal regions, and even lower barriers to X-linked regions. Further, Y-linked factors inhibit introgression by autosomes but pose no barrier to mitochondria. These examples illustrate that sex-specific incompatibility factors can induce region-specific backward migration rates throughout the genome, even to the extent of generating higher FST in autosomal than X-linked regions, a pattern that might be interpreted as a signature of selection or sex-biased dispersal. We present analytical expressions for backward migration rates induced by sex-specific incompatibility. We then use these results to develop a coalescent-based algorithm for sampling from the exact distributions of topologies and internode lengths of the gene genealogy of a sample of neutral sites. We present preliminary results from application of our methods to genomic sequences sampled from Drosophila pseudoobscura and D. persimilis, a model system for introgression between recently diverged species for which locations of hybrid incompatibility factors have been determined by direct experiment.

3393W

Admixture Analysis of Spontaneous Hepatitis C Virus Clearance among Individuals of African-Descent. *G.L. Wojcik, P. Duggal on behalf of the International Consortium of HCV Spontaneous Resolution.* Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD.

Background: Hepatitis C virus (HCV) is an important global public health problem with 130–170 million people worldwide chronically infected and 350,000 deaths each year. Associations between ethnicity and spontaneous viral clearance have been reported with African-Americans having a 25% lower rate of spontaneous and treatment-related viral clearance than European-Americans. Host genetics plays a large role in spontaneous clearance with significant associations in MHC and IL28B. Methods: From the International Consortium of HCV Spontaneous Resolution, we identified 451 individuals of African-descent, as determined by Principal Components Analysis (PCA). Local ancestry was estimated using LAMP (Local Ancestry in Admixed Populations), with two ancestral populations being assumed. The association between local ancestry and HCV spontaneous clearance was determined using logistic regression across the genome. Suggestive regions were then adjusted for SNP effects and a possible interaction with IL28B. Results: On average, individuals were genetically 82% of African-descent. This average is consistent with prior estimates in African-Americans. Approximately 26,000 local ancestry blocks were estimated across the genome, with no difference in average global ancestry between groups. A region on chromosome 15 had a suggestive association with spontaneous clearance in the gene SCAPER. Local ancestry of African descent conferred 1.9 times higher odds of clearance (p=5.9×10-4). When the region was adjusted for SNP effects, the top SNP was rs4886800 with an odds ratio of 1.8 (p= 0.005), and African ancestry was associated with 2.4 times the odds of clearance (p=4.4×10-5). Both signals were independent of IL28B (p>0.05). No other region approached statistical significance (p>0.0001). Conclusions: SCAPER is an S-phase cyclin A-associated protein in the endoplasmic reticulum. It is involved in recruiting cyclin A, which is essential for the progression of a cell through the S phase. Abnormal expression in the cytoplasm has been found to be associated with halted cell cycle progression and chronic HCV infection. This study demonstrates the need to estimate local ancestry and its effect on outcomes that may differ between ancestral backgrounds. The independent association of SCAPER may explain some of the ancestral differences in spontaneous resolution of HCV between African- and European-Americans and warrants further investigation.

3394T

A Genome-Wide Perspective of Population Structure and Individual Admixture in Xinjiang: Implications for Demographic History and Complex Trait Gene Mapping. *S. Xu[1], S. Li[2], D. Lu[1], P. Qin[1], H. Lou[1], W. Jin[1], X. Pan[2], L. Jin[2].* 1) CAS-MPG Partner Institute for Computational Biology, Shanghai, Shanghai, China; 2) Fudan University, Shanghai, China.

Many human populations settled at Xinjiang, China demonstrate an array of mixed anthropological features of East Eurasian (EEA) and West Eurasian (WEA), indicating a possible scenario of biological admixture between already differentiated EEA and WEA populations. However, their complex biological origin and genomic make-up, as well as their genetic interaction with surrounding populations are not well understood. In an attempt to decipher their genetic structure and population history, we conducted, to our knowledge, the first genome-wide study of four populations residing in Xinjiang. Analysis was performed using over 900,000 genome-wide single nucleotide polymorphisms (SNPs) in 200 unrelated individuals and made comparisons with 3,165 individuals representing 128 world-wide populations. Based on both genome-wide random markers and ancestry informative markers (AIMs), we demonstrated that the Kazakhs derived their ancestries from both EEA and WEA. Only modest variation in admixture proportions was observed among individuals within each population. We concluded that the present-day Xinjiang populations can be considered as admixed populations by genomic make-up, and their admixture features hold great promise to applications of admixture mapping and advance genetic-based medical research. We finally constructed a genome-wide admixture map with ancestry informative SNPs for gene discovery using admixture mapping in Xinjiang populations.

3395W

Genomics variants in the endogamous Mashadi Jewish population identified by pooled whole exome sequencing. *H. Ostrer[1], W.J. Huh[2], G. Akler[3], L. U[2], M. Popovic[4], D. Kural[4], J. Sheffield[4], E. Burns[1], G. Atzmon[1], C. Oddoux[2].* 1) Albert Einstein College Med, Bronx, NY; 2) Montefiore Medical Center, Bronx, NY; 3) Mount Sinai School of Medicine, New York, NY; 4) Seven Bridges Genomics, Cambridge, MA.

The Mashadi Jewish population is a highly endogamous group of ~20,000 members who ancestors once resided in Mashad, Iran. To understand the genetic variation within this group, whole exome sequencing was performed on four pools of 100 unrelated individuals, following capture using Nimblegen SeqCap EZ Human Exome Library v2.0 or Illumina SureSelect (2 pools each) to a coverage of 35–55-fold. Variant calls were made using GATK and found to be quantitative over the range of alleles (r ~0.90). The summary statistics demonstrated 762123 variants with 29036 mapped by ANNOVAR. Of these 538 variants had allele frequencies ≤.10 in the Mashadi Jewish population and were shown by PolyPhen2, SIFT, Mutation Taster, GERP, PhyloP to be potentially deleterious. Seventy-six demonstrated 4-fold enrichment compared to the 5400 samples in the NHLBI Exome Sequencing Project database and 6 occurred in genes known to be associated with Mendelian phenotypes (ALS2, CLCNKB, GP6, KRT14, PRR4, NR3C2). These findings demonstrate that pooled sequencing is a quantitative and cost-efficient means for generating genomic variants for previously uncharacterized populations and potentially can identify previously unknown risks for Mendelian disorders within the population.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3396T

An evaluation of genetic characteristics of two population isolates from Greece: the HELIC-Pomak and MANOLIS studies. *G. Dedoussis[1], I. Tachmazidou[2], AE. Farmaki[1], L. Southam[2], K. Palin[2], A. Kolb-Kokocinski[2], W. Rayner[2], E. Daoutidou[1], I. Ntalla[1], K. Panoutsopoulou[2], E. Tsafantakis[3], M. Karaleftheri[4], E. Zeggini[2].* 1) Dietetics-Nutrition, Harokopio Univ, Athens, Greece; 2) Wellcome Trust Sanger Institute The Morgan Building Wellcome Trust Genome Campus Hinxton Cambridge CB10 1HH; 3) Health Center, Anogeia, 74051 Crete, Greece; 4) Health Center, Exinos, 67300 Xanthi, Greece.

The study of low-frequency and rare variants can be empowered by focusing on isolated populations, in which rare variants may have increased in frequency and linkage disequilibrium tends to be extended. Sequencing is efficient in isolates, because variants are shared in extended haplotype contexts, supporting accurate imputation. Here we assess sample sets collected from two Greek populations: the Pomak villages are a set of religiously-isolated mountainous villages in the North (population size 11,000) with high incidence of metabolic-related phenotypes (eg. Hypertension >23%; hyperlipidemia >16%) of cardiovascular disease; Anogia is a mountainous village on Crete, with high levels of longevity (population size 4,000). 747 and 1118 individuals respectively were typed on the Illumina OmniExpress platform. We calculated genome-wide IBS statistics to assess the degree of relatedness and compared it with the general Greek population (707 samples with OmniExpress data, TEENAGE study). We additionally calculated the proportion of individuals with at least one “surrogate parent” as a means for accurate long-range haplotype phasing and imputation, as proposed by Kong et al, Nature Genetics 2008. We find 1–1.4% of individual pairs with pi-hat>0.05, and ~0.4% with pi-hat>0.1 in the isolates compared to 0% in the general Greek population. We also find that ~80%-82% of subjects have at least one surrogate parent in the isolates, compared to ~1% in the outbred Greek population. We have established the HELIC-Pomak and MANOLIS cohorts as genetic isolates and we are currently whole-genome sequencing 250 individuals to enable imputation and subsequent association testing. This approach has the potential to identify novel robust associations with disease-related complex traits.

3397W

Genomic variants in the Dominican population identified by whole exome genotyping using Affymetrix Axiom Arrays. *C. Oddoux[1], C. Cambell[2], L. U[1], H. Ostrer[2].* 1) Montefiore Medical Center, Bronx, NY; 2) Albert Einstein College of Medicine, Bronx, NY.

The Dominican population is an admixed population with native, African, and European origins. To understand the genetic variation within this group, whole exome genotyping was performed on 192 unrelated individuals, using the Affymetrix Axiom array. The summary statistics demonstrated 319,284 variants with 260,478 mapped by ANNOVAR. Of these 10,995 variants had allele frequencies ≤10 in the Dominican population and were shown by PolyPhen2, SIFT, MutationTaster, LRT, or PhyloP to be potentially deleterious the majority showing four-fold or greater enrichment as compared to the 5400 samples in the NHLBI Exome Sequencing Project database and variants were observed in 1110 genes known to be associated with Mendelian phenotypes. These findings demonstrate that whole exome genotyping using Axiom arrays is a cost-efficient means for generating genomic variant frequencies for previously uncharacterized populations and potentially can identify previously unknown risks for Mendelian disorders within the population.

3398T

A Population-Based Study and Subsequent Disclosure of Autosomal Recessive Disease-Causing Mutations in a Founder Population. *J.X. Chong[1], R.L. Anderson[1], R. Ouwenga[1], D.J. Waggoner[1,2], C. Ober[1,3].* 1) Human Genetics, University of Chicago, Chicago, IL; 2) Pediatrics, University of Chicago, Chicago, IL; 3) Obstetrics and Gynecology, University of Chicago, Chicago, IL.

The decreasing costs of exome and whole genome sequencing have resulted in a renaissance for identifying Mendelian disease mutations, and for the first time it is possible to survey the distribution and characteristics of these mutations in large population samples. We conducted carrier screening for all autosomal recessive mutations known to be present in 1,644 members of a founder population, the Hutterites of South Dakota, and revealed surprisingly high carrier frequencies for many of these mutations. By utilizing the rich demographic, genetic, and phenotypic data available on these subjects and simulations in the exact pedigree that these individuals belong to, we show that the majority of mutations were likely introduced into the population by a single Hutterite founder and then drifted to the high frequencies observed. In addition, we find that carriers of the same founder mutation often share extremely small segments (<0.5Mb) of the mutation-bearing haplotype, which has implications for homozygosity mapping approaches used in founder populations. We further show that while there is an increased incidence of autosomal recessive diseases overall, the mean carrier burden is likely to be lower in the Hutterites than in the general population. Finally, based on simulations, we predict the presence of 30 or more undiscovered recessive mutations among these subjects, which would at least double the number of AR diseases that have been reported in the Hutterites to date. Given the potential clinical and reproductive relevance of the genotypes discovered during this project, we initiated a phased approach to disclosure of results, which included assessment of attitudes towards disclosure of carrier information, an educational program, and a consent process that would allow us to disclose genetic results to participants in our studies who desired this information. We also met directly with individuals and families to share information related to medically actionable diseases that were present in themselves or their family. Overall, ~80% of individuals who participated in the educational program signed consent forms for the release of their carrier status for 14 diseases.

3399W

Deep-coverage whole genome sequencing of 100 Qatari reveals elevated prevalence of population-specific variants linked to disease. *J.L. Rodriguez-Flores[1], K.A. Fakhro[2], N.R. Hackett[1], J. Salit[1], J. Fuller[1], J.A. Malek[2], R. Badii[3], A. Al-Marri[3], L. Chouchane[2], R.G. Crystal[1], J.G. Mezey[1,4].* 1) Dept of Genetic Medicine, Weill Cornell Medical College, New York, NY; 2) Dept of Genetic Medicine, Weill Cornell Medical College, Doha, Qatar; 3) Hamad Medical Corporation, Doha, Qatar; 4) Dept of Biological Statistics and Computational Biology, Cornell University, Ithaca, NY.

The Arabian Peninsula is the migration crossroads between Africa and Eurasia, and hence is an important part of the story of human genetic variation. The Qatari population is a broad sample of the genetic variation in the region, including individuals of Bedouin (Q1), Persian (Q2), and African (Q3) ancestry. The Q1 population is of particular interest because of a high rate of consanguinity within the population, with long runs of homozygosity in the genome. In order to study the effects of extreme consanguinity on the Qatari genome while sampling the complete spectrum of population genetic variation, we selected for whole genome sequencing 100 Qatari with varying degrees of Bedouin ancestry. Over half the individuals selected for sequencing are of primarily Bedouin ancestry (>70%), with the rest having varying proportions of Bedouin admixture, and eight being highly admixed (<50% Q1, Q2 or Q3). Each genome was sequenced to >30x depth using Illumina paired-end 100bp reads, these reads were mapped to the GRCh37 human reference using BWA and variants were called using a pipeline based on GATK. We identified an overall higher proportion of homozygosity in the Q1, with a median homozygosity in coding regions higher than Q2, Q3 and all 1000 Genomes Project populations. Using a manually curated version of the OMIM database, we identified multiple heterozygous carriers for recessive variants linked to severe childhood-onset Mendelian disorders with prevalence up to 2%. Four of these variants were observed only in one sub-population (Q1, Q2 or Q3), and hence the sub-population allele frequency was higher than for the general population. Examples include an arterial tortuosity variant at 4% prevalence in Q1, and a retinitis pigmentosa variant at 5% prevalence in Q2. Of these two, only the retinitis pigmentosa variant was previously observed in the 1000 Genomes project (1% prevalence in Europe). Analysis of identity-by-descent showed shared ancestry of the local haplotypes flanking these variants, indicating a common founder origin. We hypothesize that these observations are a combined result of a founder effect in the population and a high rate of consanguinity, resulting in an unusually high prevalence of a severely deleterious mutation within a subpopulation that frequently intermarries. This whole genome dataset can serve as a foundation for future studies of genomics in Qatar and the Arabian Peninsula.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3400T

Single gene disorders in Western India: Cohort of high risk group. *C. Ankleshwaria, J. Sheth, H. Patel, J. Lekshami, F. Sheth.* Biochemical and Molecular Genetics, Frige's Institute of Human Genetics, Ahmedabad, Gujarat, India.

Present study comprises of 755 cases referred for various single gene disorders from 2007 to 2012. This includes 313 (41.45%) for β-Thalassemia, 139 (18.41%) for Cystic fibrosis (CF), 116 (15.36%) for Spinal Muscular Atrophy (SMA) and 74 (9.80%) for Duchenne Muscular Dystrophy (DMD), 32(4.23%) for Spinocerebellar Ataxia (SCA), 24(3.17%) of Achondroplasia, 22 (2.91%) of Fragile- X, 20(2.64%) of Albinism and 15(1.98%) of Huntington The most common mutation observed in β-Thalassemia was IVS 1, 5 (G-C) in 108(34.50%) followed by FS 41/42 in 33(10.54%), 619 bp deletion in 32(10.22%), FS 8/9 in 25(7.98%) and IVS 1, 1 (G-T) in 27(8.62%) cases. For CF, δf508 was observed in 24(17.26%), R117H in 3(2.15%), heterozygous forδf508 and S549N in 1(0.71%). For SMA, exon 7 and exon 8 of SMN1 and SMN2 was deleted in 57(49.13%), only exon 7 in 3(2.58%) cases of SMN1, only exon 8 in 3(2.58%) cases of SMN1, only exon 7 of SMN2 gene in 3(2.58%) cases and exon 7and 8 of SMN2 gene in 1(0.86%) case. For DMD, single exon deletion was observed in 10(13.51%) cases, 2 exon deletion in 6(8.10%), three exon deletion in 14(18.91), four exon deletion in 6(8.10%) and remaining 38(51.35%) showed multiple exon deletions. Of 32 cases suspected with SCA, 1(3.12%) was found with SCA 1, 8(25%) with SCA 2, 3(9.37%) with SCA 3. For Achondroplasia, 3(12.5%) cases were found to have G1138A and 2(8.33%) were found to have G1138C in heterozygous status. Cases suspected with Fragile-X, only 3(13.63%) were confirmed to have more than 100 repeats of CGG. For Albinism, R278X was the most common mutation with heterozygous status in 5(33.33%) and homozygous in 5(33.33%) cases. For Huntington, 6(40%) of 15 were found to have increased CAG repeats confirming the disease status. Overall study demonstrates that β-Thalassemia, CF, SMA and DMD are the most common single gene disorder and national screening planning may be needed to attenuate the burden of these diseases in addition to carrier screening and prenatal diagnosis.

3401W

Study on the distribution of rare genetic variations in the French Canadian population. *S. Girard[1,2], L.-P. Lemieux Perreault[1], I. Mongrain[1], G. Lettre[3], J. Rioux[3], M. Phillips[3], S. de Denus[3], G.A. Rouleau[4], J.-C. Tardif[3], M.-P. Dube[1,3].* 1) Université de Montréal, Montreal, Quebec, CanadaCentre de Pharmacogénomique Beaulieu-Saucier Université de Montréal 5000, rue Belanger Est Montreal, Quebec H1T 1C8; 2) CHUM Research Center, Y-3624, Pav JA-de-Sève 2099 rue Alexandre-de-Sève, Montreal, Qc; 3) Institut de Cardiologie de Montréal 5000, rue Bélanger Montréal (Québec) H1T 1C8; 4) Centre Hospitalier Universitaire Sainte-Justine Research Center, Montréal, Québec, Canada.

Several studies have recently demonstrated that unlike common genetic variations, rare genetic variants have a unique distribution profile specific to each population. Until recently, it was not possible to conduct population studies on the distribution of rare variants due to limitations linked to technologies and the number of identified rare variants. Using a specifically designed genotyping platform to identify variations in exons (Illumina ExomeChip), we examined the distribution of rare variations in a cohort of over 10,000 individuals (the Montreal Heart Institute Hospital Cohort). Among this cohort, we compared the prevalence of rare variations in 7713 individuals of French Canadian origin to those of European populations (1000 genome project and NHLBI GO Exome Sequencing Project). We subdivided the French Canadian population samples according to the birthplace of the four grandparents of the individuals. Using the geographical coordinates of the birthplace, we kept only individuals whose grandparents were born less than 25km apart. Using a MDS plot, we were able to identify a divergent subpopulation that characterizes the Charlevoix-Saguenay region, a known geographic isolate. The distribution of selected rare variants from this region was found to be distinct from the rest of the French Canadian population. This allows us to better characterize the founder effect specific to the French-Canadian population in addition to improving our understanding of the evolution of allele frequencies of rare variations.

3402T

Identification of a novel Native American Y chromosome founding lineage in North-west South America. *L. Roewer[1], M. Nothnagel[2], L. Gusmão[3,4], V. Gomes[3], M. Gonzaléz[3], D. Corach[5], A. Sala[5], E. Alechine[5], T. Palha[4], N. Santos[4], A. Ribeiro-dos-Santos[4], M. Geppert[1], S. Willuweit[1], M. Nagy[1], S. Zweynert[6], M. Baeta[7], C. Núñez[7], B. Martínez-Jarreta[7], F. González-Andrade[8], E. Fagundes de Carvalho[9], D. Aparecida da Silva[9], J.J. Builes[10,11], D. Turbón[12], A.M. Lopez Parra[13], E. Arroyo-Pardo[13], U. Toscanini[14], L. Borjas[15], C. Barletta[16], S. Santos[4], M. Krawczak[2].* 1) Department of Forensic Genetics, Institute of Legal Medicine and Forensic Sciences, Charité - Universitätsmedizin Berlin, Germany; 2) Institute of Medical Informatics and Statistics, Christian-Albrechts University Kiel, Germany; 3) Institute of Molecular Pathology and Immunology, University of Porto, Portugal; 4) Universidade Federal do Pará, Laboratório de Genética Humana e Médica, Belém, Brazil; 5) Universidad de Buenos Aires, Facultad de Farmacia y Bioquimica, Servicio de Huellas Digitales Geneticas, Buenos Aires, Argentina; 6) Department of Psychiatry and Psychotherapy, Charité - Universitätsmedizin Berlin, Germany; 7) Department of Forensic Medicine, University of Zaragoza, Spain; 8) Science and Technology Department, Ministry of Public Health, Quito, Ecuador; 9) Laboratorio de Diagnósticos por DNA, Instituto de Biologia, Universidade do Estado do Rio de Janeiro, Brazil; 10) GENES Ltda., Lab. Genetica Forense y Huellas Digitales del DNA, Medellin, Colombia; 11) Instituto de Biología, Universidad de Antioquia, Medellín, Colombia; 12) Unitat d'Antropologia, Dept. Biologia Animal, Facultat de Biologia, Universitat de Barcelona, Spain; 13) Laboratorio de Genetica Forense, Dpto. de Toxicología y Legislación Sanitaria, Fac. de Medicina, Universidad Complutense de Madrid, Spain; 14) PRICAI-Fundación Favaloro, Buenos Aires, Argentina; 15) Laboratorio de Genetica Molecular, Unidad de Genetica Medica, Facultad de Medicina, Universidad del Zulia, Maracaibo, Venezuela; 16) Laboratorio de Genética Humana, Facultad de Ciencias Biológicas, UNMSM-Universidad, Nacional Mayor de San Marcos, Lima, Peru.

We initiated the largest population genetic study of Y-chromosomal genetic variation in South American natives to date. A total of 1011 individuals, representing 50 tribal populations from 81 settlements, were genotyped for up to 17 short tandem repeats and up to 16 single-nucleotide polymorphisms (Y-SNPs), defining native American phylogenetic lineages Q and C and their sub lineages. We report upon the detection of a novel Native American founding lineage C-M217 (C3*) in different tribal and linguistic groups from Ecuador (Kichwa, Waorani). This finding is intriguing in view of the high prevalence of the same haplogroup in Central, East and North East Asia, and its concurrent absence from North America. Possible scenarios include (i) later migratory waves that quickly passed the existing populations in North America, and (ii) long-distance contact with East Asia. The introduction of C3* chromosomes into South America from Eastern Asia by sea, either along the American West coast or across the Pacific could revitalize the hypothesis for the origins of New World ceramics found in coastal Ecuador from the Japanese Jōmon culture. The striking differences observed between the C3* Y-STR haplotypes of Ecuadorians and Asians on the one hand and between the tribes within Ecuador on the other would be explicable in terms of a long divergence time after the arrival and the process of tribalization.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3403W

Single Nucleotide Variation Analysis of Protein Domain Signatures of Human Genes. *J. Freudenberg.* Ctr Human Gen, Feinstein Inst Med Res, Manhasset, NY.

We recently proposed a simple method to estimate relative density of nonsynonymous SNVs (rdnsv) in a set of genes as compared with its expectation (Freudenberg et al., PLoS ONE 7(6): e38087, 2012). In short, rdnsv is estimated by correcting the ratio between nonsynonymous SNVs (nsSNVs) per nonsynonymous site and synonymous SNVs (sSNVs) per synonymous site for the bias of synonymous sites towards transitions in the genetic code and different mutation rates for transitions and transversions. We then fit a linear regression model to rdnsv over different SNV allele frequency bins. The y-intercept of this model (rdnsv0) predicts rdnsv for an allele frequency close to 0, which can be interpreted as the proportion of nonsynonymous sites where mutations may segregate with a frequency notably greater 0. The slope of the model further determines rdnsv for the allele frequency close to 1 (rdnsv1), which measures the proportion of nonsynonymous sites where nsSNVs may reach fixation. We now apply our method to evaluate the levels of nsSNVs among the 30 most prevalent Pfam domain signatures in the human genome, using different published exome sequencing datasets. The greatest rdnsv estimates are displayed by PF00067:Cyt_P450, PF00005:ABC_transpt-like, and PF00001:GPCR_Rhodpsn. Interestingly, Cyt_P450 and ABC_transpt-like stand out by high rdnsv0 and small rdnsv1 estimates, indicating many sites where nonsynonymous SNVs may segregate at low frequency, but few sites where they may reach high frequency or fixation. In contrast, the sequence annotated by PF00001:GPCR_Rhodpsn displays both high rdnsv0 and rdnsv1 estimates. This domain signature is typically found in olfactory receptors, where nonsynonymous variants easily reach high frequency due to relaxed constraint. On the other end of the spectrum, we find the smallest rdnsv0 and rdnsv1 estimates for the domain signatures PF07714:Tyr_pkinase, PF00069:Se/Thr_pkinase, and PF00076:RRM_RNP1. These protein domains play important signaling roles in many cell types, which may explain their high levels of constraint and reduced tolerance even for rare variants. Importantly, our results are consistent across parameters and SNV datasets. In particular, similar estimates of rdnsv are obtained by the approximative strategy that multiplies the observed ratio between nsSNVS and sSNVs with the factor f'=0.4 and other strategies that consider the exact numbers of synonymous, nonsynonymous, transition, and transversion sites.

3404T

Investigating the evolution of deleterious mutations in the major histocompatibility complex as a potential explanation for frequently observed HLA-disease associations. *T.L. Lenz[1], D.M. Jordan[1], S.R. Sunyaev[1,2].* 1) Division of Genetics, Brigham and Women's Hospital, Harvard Medical School, Boston, MA; 2) Broad Institute of MIT and Harvard, Cambridge, MA.

Understanding the evolution of disease-causing mutations in the genome is a key objective in human genetics and evolutionary medicine. The accumulating evidence for a growing number of disease variants from genome-wide association studies raises the question as to how these causal variants evolve and are maintained in human populations despite their usually deleterious effects. Both neutral as wells as selective processes have been proposed to counteract purging of disease variants. Here we investigate these processes for one of the most gene-dense regions in the human genome, the major histocompatibility complex (MHC). While this genomic region on the short arm of chromosome 6 appears to harbor an unexpected large fraction of disease variants, it exhibits also an exceptional level of heterozygosity, which is thought to be maintained by pathogen-mediated balancing selection on the antigen-presenting HLA genes in this region. It has therefore been suggested that high levels of haplotypic heterozygosity in the MHC region provide a `shelter' for deleterious mutations by limiting expression of their homozygous phenotype and thus inhibiting purging mechanisms. This scenario would explain the otherwise unexpected high frequencies of deleterious mutations in this region and thus the frequent association of the MHC region with genetic disorders and diseases. Here we use frequency information of sequence variants from 5,379 human exomes, publicly available through the exome variant server of the NHLBI GO Exome Sequencing Project, to investigate the frequency of deleterious sequence variants in the MHC region at an unprecedented depth. This dataset comprises 17,538 protein-coding genes with nearly 900,000 amino acid-altering point mutations (SNPs). From this dataset we identify computationally predicted deleterious mutations inside and outside the MHC region and compare the observed frequency distribution of such mutations in the MHC region against simulated gene sets from the rest of the genome. We then employ forward simulations to identify the most likely evolutionary scenarios that have led to the observed distribution of damaging sequence variants in the MHC region. Our results thus shed light on the processes governing the evolution and maintenance of deleterious sequence variants in the human genome.

3405W

Quantifying the Degree of Purifying Selection in Genes Associated with Nephrotic Syndrome. *M.G. Sampson[1,4], M. Kretzler[2,4], H.M. Kang[3,4].* 1) Dept. of Pediatrics & Communicable Disease-Nephrology; 2) Dept. of Internal Medicine-Nephrology; 3) Dept. of Biostatistics; 4) University of Michigan Ann Arbor, MI.

Background: Studying the genetic basis of nephrotic syndrome (NS), a disease resulting from protein loss through the urine due to abnormalities of the glomerular filtration barrier, has resulted in the discovery of both rare and more common variants in genes associated with increased risk of disease. However, the contribution of genetic variation to the pathogenesis of NS cases remains unknown in the majority of cases. Given the observation that deleterious single nucleotide variants (SNV) in disease genes tend to undergo purifying natural selection, defining the degree of this selection in known and candidate NS genes may help to identify disease-associated loci in subjects with this disease. Leveraging genome-wide SNV data from the 1000 Genomes Project (1000G) coupled with functional prediction software, we developed and implemented a number of novel and known methods to quantify the degree of purifying selection on genes associated with NS and compared them genome-wide.**Methods:** Using 1000G data, we measured the frequency of non-synonymous SNV present in coding regions across the genome and used functional prediction software to stratify them as deleterious or benign. We then created two metrics of selection: 1) number of SNV/kilobase ("Suppressed Mutation") & 2) percentage of SNV that are private ("Enriched Rare Variants"). Using a Wilcoxon rank sum test, we contrasted the difference of frequency spectrum between benign and silent mutations for 17,506 genes across the genome and established the quantile in which the 12 NS-associated genes were located.**Results:** The NS genes demonstrating most enrichment for rare variants were *APOL1* (.008 quantile of all genes tested), *LAMB2* (.016 quantile), and *TRPC6* (.036 quantile). The NS genes with the strongest suppression of deleterious SNV were *INF2* (.008 quantile), *MYH9* (.011 quantile), and *ACTN4* (.045 quantile). *INF2* had a 66x lower rate of deleterious vs benign SNV.**Conclusions:** Multiple NS genes exhibit evidence of strong negative selection, suggesting that deleterious SNV in these genes are likely functional and could have large effects. This can be further investigated by sequencing these genes in NS-cohorts large enough to pick up rare, deleterious SNV. Additionally, annotating NS-candidate genes with these metrics of purifying selection may help prioritize genes for further study.

3406T

An Assessment of Population Stratification in Rare Variant Association tests Using an Analytic Model of Joint Site Frequency Spectra. *M. Reppell[1], M. Zawistowski[1], D. Wegmann[2], R. Weyant[1], P.L. St Jean[3], M.G. Ehm[3], M.R. Nelson[3], J. Novembre[4], S. Zöllner[1,5].* 1) Department of Biostatistics, University of Michigan, Ann Arbor, MI, USA; 2) School of Life Sciences, École polytechnique fédérale de Lausanne, Lausanne, Switzerland; 3) Quantitative Sciences, GlaxoSmithKline, RTP, NC, USA; 4) Department of Ecology and Evolutionary Biology, University of California-Los Angeles, Los Angeles, CA, USA; 5) Department of Psychiatry, University of Michigan, Ann Arbor, MI, USA.

Understanding the role rare variation plays in disease etiology is a fundamental question in human genetics today. A critical step in determining valid genotype-phenotype relationships is properly accounting for population structure. The impact of population stratification and necessary correction methods for gene-based association tests are not well understood. To address this we devised a novel analytical model for the Joint Site Frequency Spectrum (JSFS) of two populations, extending the classic hierarchical beta model of population allele frequencies. This model captures patterns of rare variant diversity observed in real sequence data without requiring explicit demographic assumptions, as would be necessary in a coalescent-based approach. We calibrated our model using data from 202 drug target genes sequenced in 14,001 individuals, with a particular focus on accurately modeling singletons, and simulated datasets with different degrees of population stratification by sampling case and control at different proportions from two populations. Gene-based association tests that combine information from multiple rare alleles into a single statistic can be broadly divided into those that accumulate rare variant counts within individual samples (for example, CMC and CMAT), and those that combine statistical evidence across individual markers (for example, SKAT). Interestingly, the P-value inflation induced by population stratification differs substantially between the two classes of rare variant tests: tests that combine evidence across markers typically have a higher rate of false positives than tests that combine variants across genes. This behavior is well predicted by summary statistics that quantify the burden and sharing of rare variants across populations. We further demonstrate that genomic control as currently defined is unable to properly correct for test statistic inflation, but find that an alternative control that uses observed test statistic variance accounts well for over- or under dispersion in gene-based test statistics. A better understanding of how population stratification affects rare variant association tests, and how to properly correct it, is critical for determining the role rare variants play in disease development.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3407W

Estimating and Interpreting F_{ST}: the Impact of Rare Variants. *G. Bhatia[1,2], N. Patterson[2], S. Sankararaman[2,5], A.L. Price[2,3,4].* 1) Harvard-Massachusetts Institute of Technology (MIT) Division of Health, Science and Technology, Cambridge, MA; 2) Broad Institute of Harvard and MIT, Cambridge, MA; 3) Department of Epidemiology, Harvard School of Public Health, Boston, MA; 4) Department of Biostatistics, Harvard School of Public Health, Boston, MA; 5) Department of Genetics, Harvard Medical School, Boston, MA.

F_{ST} is a widely used tool for studying population structure, but many different definitions, estimation methods and interpretations exist in the literature. Thus, wide variation in published estimates of F_{ST} is important to understand. For example, the F_{ST} between European (CEU) and East Asian (CHB) populations is 0.111 when estimated from HapMap3 data, but only 0.052 when estimated from 1000 Genomes data (1kG). While, changes in F_{ST} from sequencing data might be expected from including rare variants we show that this is largely through bias introduced by the estimation method and not population genetic factors. We describe a method that is shown to avoid these biases. We consider three specific aspects of estimation: (1) defining F_{ST} for a single SNP, (2) combining estimates of F_{ST} across multiple SNPs, and (3) selecting the set of SNPs used in the computation. Correcting for differences in each of these aspects of estimation yields estimates of F_{ST} that are much more concordant between genotype and sequence data. For example, our estimate of F_{ST} between CEU and CHB from 1kG is 0.106, only slightly lower than the HapMap3 estimate. This decrease is due to ascertainment bias of SNPs included in the HapMap3 project, not to properties of rare variants. In general, F_{ST}at rare variants in a population will be sensitive to demographic events affecting that population. When comparing CEU to CHB, for example, we show that rare variants in CEU and CHB have higher F_{ST} than common variants. This is consistent with the influence of strong bottlenecks on F_{ST} at rare variants. We note that ascertainment in an out-group—for example, Yoruba (YRI)—will remove this frequency dependence of F_{ST}. Finally, we show that single-SNP estimates of F_{ST} based on a common definition (Weir and Cockerham 1984) can become inflated in a setting of very different sample sizes. This inflation can result in false-positive signals of natural selection. Indeed, we show that in a recent study of selection that compared 1,890 African-American and 113 YRI samples, (Jin et al. 2011), F_{ST} estimates at 9 of the 10 reported novel loci are inflated by the disparity in sample size, and, after correction, only 7 of these 10 loci remain nominally significant. This suggests that caution is warranted when using this definition to rank single-SNP estimates of F_{ST}. Our results indicate that a careful protocol is needed for producing F_{ST} estimates. We provide such a protocol.

3408T

Are rare and common variant stratification patterns the same? *E. Genin[1,2], M-C. Babron[1,2], M. de Tayrac[1,2], D.N. Rutledge[3], E. Zeggini[4].* 1) U946 Genetic variability and human diseases, INSERM, Paris, France; 2) Institut Universitaire d'Hématologie, Univ Paris Diderot, Paris, France; 3) AgroParisTech, Paris, France; 4) Wellcome Trust Sanger Institute, Hinxton, Cambridge, United Kingdom.

If the population stratification of common variants has been extensively studied on empirical data, this is not the case for rare variants. Rare variants are more likely than common variants to be under moderate levels of negative selection and/or to have arisen recently and be population-specific. It is thus plausible that rare variants could exhibit stratification patterns different from common variants. With the development of high-throughput sequencing, the role of rare variants in diseases can now be tested in association studies. To avoid false positive associations, it is important to understand how rare variants are stratified and how to adjust for their stratification in association tests. Using the WTCCC1 Affymetrix 500K genome scan data, we show that the stratification patterns of rare and common variants are indeed different. The top principal component extracted from the rare variant category shows poor correlations with any principal component or combination of principal components from the low frequency or common variant categories. These results could suggest that a suitable solution to avoid false positive association due to population stratification would involve adjusting for the respective PCs when testing for variants in different allele frequency categories. However, we found this was not the case both on type 2 diabetes case-control data and on simulated data. Indeed, adjusting rare variant association tests on PCs derived from rare variants does no better to correct for population stratification than adjusting on PCs derived from more common variants. Mixed models perform slightly better for low frequency variants than PC based adjustments but less well for the rarest variants. These results call for the need of new methodological developments specifically devoted to address rare variant stratification issues in association tests.

3409W

Variant frequency and the scale of population structure. *P. Raska, O. De la Cruz Cabrera.* Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, OH.

Proper characterization of population structure is important for understanding current and historical demographics of human populations and for adjusting for this structure when conducting tests of genetic association. Principal component analysis (PCA) has proven effective for this purpose. Adding kernel methodology to PCA brings with it two added benefits, the possibility of dealing with non-linear structure and the convenient notion of kernel bandwidth. Here we make the connection between kernel bandwidth and variant frequency as both being delimiters of the scale of population structure that can be characterized. We investigate the effect of variant frequency on inferred population structure through simulation. The underlying hypothesis is that rare variants, which tend to be shared only between individuals with more recent common ancestry, will tend to reveal structure at a smaller scale, while common variants, which can be shared by unrelated individuals, reveal differences at a larger scale. Conversely, common variants will not vary sufficiently to inform on smaller scale variation and rare variants will be less informative than common variants for larger scale variation. The latter becomes obvious with the realization that it would not be possible to infer the relative distances between subpopulations on the basis of completely private variants. We further hypothesize that as the geographic region under consideration is reduced, rare variants will become more common reaching an optimum frequency at which the variant will be most informative for the regional population structure. Finally,we predict that the population structure that is present for common and rare variants will differ not only in scale but also within the finer scale of rare variants, where selective forces may have shaped population structure independently of the neutral drift pattern represented by common variants. What population structure is best characterized by variants of different frequency is a question of paramount importance with the advent of next generation sequencing. In rare variant association analyses, now becoming common with the availability of sequencing data, the type of confounding structure present in the data may not be well represented nor controlled for by common variants. This study wishes to address this question and in so doing, give guidelines for population stratification control in the new era of genetic association studies.

3410T

Male-specific common CNVs and complex gender differences in CNV detected among 808 olfactory receptor loci, in 150 phenotypically normal individuals from the 1000 Genome Project. *F. Shadravan.* None, San Francisco, CA.

Gender differences due to both autosomal and sex chromosomes play a pivotal role in the human genetic identity and are also manifested in many genetic disorders. In this study their effects on copy number variation (CNV), the gain or loss of genomic materials, which is also known to cause some diseases were explored. As the olfactory receptor (OR) repertoire comprises the largest human gene family, it was selected for this study. The CNV genotypes came from open access data, studying 808 OR loci in 150 phenotypically normal individuals from the 1000 Genome Project. Out of 808 loci studied, 188 pseudogenes and 134 genes showed a total of 4,595 CNVs, and their analysis revealed the following two novel findings: 1- Detection of 22 gene duplications among nine OR loci in subtelomeric region of chromosome X in 13 out of 57 males, while absent in 93 females studied. One of these loci was located between two genes previously identified as part of a “gay genes” region. 2- Detection of significant gender differences in OR pseudogenes, and not OR genes, in which the male genome showed more CNVs, among 494 non-common variants detected in 260 OR loci; whereas the comparison of 4,101 common CNV variants detected among 62 OR loci did not show any significant differences among the genders, either between OR pseudogenes or genes (the standard Bonferroni correction was used for multiple testing). Using a CNV public database for sick children (ISCA), the application of these findings for improving clinical molecular diagnostics was discussed through showing an example of gender differences in CNV among kids with autism. Additional clinical relevance is discussed, as the most CNV-enriched polymorphic OR cluster in the human genome, located at chr 15q11.2, was found near the PWS/AS bi-directionally imprinted region associated with two well-known mental retardation syndromes. In light of finding recent gender differences among OR pseudogenes, an indirect role for OR processed pseudogenes in the embryonic development of primate brain is proposed. It is based on their parent-of-origin regulatory suppression of their intact counterpart genes, thus switching off temporally a more primitive mental cognitive system based heavily on olfaction and pheromones, as seen in most mammals to allow the development of a newer cognitive system based on a larger brain in primates. Syndromic mental retardation consequently might arise due to developmental delays in this switching process.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3411W

Transcriptome sequencing across seven diverse populations characterizing novel transcript isoforms and allele-specific expression. *H.A. Costa[1], A.R. Martin[1], J.M. Kidd[2], B.M. Henn[1], M.C. Yee[1], F. Grubert[1], S.B. Montgomery[1], H.M. Cann[3], M.P. Snyder[1], C.D. Bustamante[1].* 1) Stanford University School of Medicine, Department of Genetics, Stanford, CA, 94305; 2) University of Michigan School of Medicine, Department of Human Genetics, Ann Arbor, MI, 48109; 3) Foundation Jean Dausset, Centre d'Etude du Polymorphisme Humain, Paris, 75010, France.

RNA sequencing studies enable the characterization and quantification of the full complement of coding elements within the human genome. These transcribed elements play an important role in human evolution and phenotypic diversity. Present studies have shed insight into novel transcript structures and transcript variability in individual populations. However, in order to obtain a more complete understanding of how gene expression variation influences complex human phenotypes and to better understand overall human diversity, further studies investigating under-sampled and isolated populations are necessary. Thus, we have sequenced the genomes and mRNA fractions of 45 lymphoblastoid cell lines from seven populations within the Human Genome Diversity Project. These populations include: San Bushmen of southern Africa, Mbuti Pygmies of central Africa, Mozabites of north Africa, Pathans of central Asia, Cambodians of east Asia, Yakut of Siberia, and Mayans of Mexico. This cohort of populations mirroring major human migratory routes will enable us to leverage the robust and quantitative nature of RNAseq to identify novel transcript isoforms as well as rare transcripts and sites of allele-specific expression. Preliminary results reveal ~100 novel transcript isoforms significantly expressed across several individuals, with several showing population unique expression. Additionally, functional annotation of a subset of these isoforms shows enrichment in metabolism and immune response genes. Further analysis will be required to correlate these variants with functional significance. This dataset will allow for the characterization of human transcriptome variation across diverse populations.

3412T

African Genome Variation Project. *I. Tachmazidou[1], F. Ayele[2], C. Pomilla[1,3], J. Asimit[1], D. Gurdasani[1,3], R. Nsubuga[4], D. Shriner[2] on behalf of the APCDR investigators.* 1) Wellcome Trust Sanger Insitute, Hinxton, United Kingdom; 2) Center for Research on Genomics and Global Health, Bethesda, US; 3) Department of Public Health, University of Cambridge, Cambridge, United Kingdom; 4) MRC/UVRI Uganda Research Unit on AIDS, Uganda.

Genome-wide association studies in populations from sub-Saharan Africa (SSA) are eagerly anticipated, but there is a paucity of genetic data to inform powerful study design. Pronounced genetic diversity across ethnic groups within SSA, in conjunction with low levels of LD and differences in haplotype structure, give rise to statistical genetics challenges when designing and conducting genomic epidemiology studies. The African Genome Variation Project is a collaboration across the African Partnership for Chronic Disease Research (www.apcdr.org), which aims to facilitate large-scale genetic association studies in diseases of relevance to SSA populations by providing insight into the genetic variation landscape of different ethnic groups. To achieve this, we have genotyped 100 unrelated individuals from each of 12 ethnolinguistic groups on the 2.5 million SNP Illumina platform. This project has been designed to provide a currently unavailable genome-wide view of common variation and a framework for statistically imputing genetic variants for genome-wide association studies in anticipation of future larger-scale whole-genome sequencing endeavours in relevant populations. We are actively examining: 1) the allele frequency spectrum of variants on the chip; 2) patterns of LD and haplotype structure; 3) the utility of the 2.5M Illumina platform for relevant populations; 4) imputation-based approaches aiming to increase genetic association study power; 5) analytical challenges and the need for new statistical genetics methods to address them; 6) the data as a valuable resource for the scientific community; 7) collaboration and synergies among contributing parties; and 8) the set-up of a research framework and resource for additional analyses, for example studies of natural selection, mapping genetic variation to linguistic and demographic differences, and evaluating powerful approaches for trans-ethnic meta-analysis and fine-mapping. We find that between 1.32 and 1.39 million SNPs have MAF>0.05 and that between 200 and 270 thousand SNPs are monomorphic in the individuals examined. We also find that there are high levels of redundancy on the chip, as calculated based on pairwise correlation between variants in each ethnolinguistic group. We will be submitting all genotype data to the European Genotype Archive (EGA). These data will be available to bona fide researchers under a managed access procedure, in keeping with ethical approval for the samples utilised.

3413W

Inference of genetic ancestry using genotype data from a single individual. *C. Chen[1], S. Lee[2], S. Pollack[1,2,3], D. Hunter[1,4], P. Kraft[1,2,4], J. Hirschhorn[3,5,6,7,8], A. Price[1,2,3].* 1) Department of Epidemiology, Harvard School of Public Health, Boston, MA; 2) Department of Biostatistics, Harvard School of Public Health, Boston, MA; 3) Broad Institute of Harvard and MIT, Cambridge, MA; 4) Channing Laboratory, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA; 5) Division of Genetics, Children's Hospital, Boston, MA; 6) Division of Endocrinology, Children's Hospital, Boston, MA; 7) Program in Genomics, Children's Hospital, Boston, MA; 8) Department of Genetics, Harvard Medical School, MA.

Existing methods to infer genetic ancestry generally require raw genotypes from a large set of individuals. Applications in personal genomics and genetic association studies motivate inference of genetic ancestry using raw genotypes from just a single individual, given the logistical challenges surrounding the sharing of raw genotypes, and because the samples in a particular study may have insufficient diversity or may have family structure which can impede ancestry inference. Here, we present methods and software to infer genome-wide ancestry using genotypes from a single individual by predicting principal components (PCs) with SNP weights pre-computed from external reference panels of genotypes. Notably, the reference panels of genotypes do not need to be shared for predicting ancestry. The predicted PCs can be transformed into ancestry from continental populations such as West Africans, East Asians and Europeans, or more closely related subpopulations such as Northwest Europeans (NW), Southeast Europeans (SE) and Ashkenazi Jews (AJ). We used 394 samples from HapMap 3 continental populations and 6,546 European Americans from Framingham Heart Study SHARe data as our external reference panels of genotypes. We applied our method to two independent data sets: 49 African Americans from HapMap 3 (ASW) and 1,636 European Americans from a study of bipolar disease (BD). We used results from PCA incorporating genotypes from external reference panels as our gold standard. For ASW samples, the predicted first PC obtained using pre-computed SNP weights attained an R^2 of 1.00 (95% CI: 1.00 to 1.00) while first PC obtained using genotypes from ASW samples alone attained an R^2 of 0.32 (95% CI: 0.16 to 0.44), which reflects the confounding effects of family structure in ASW samples on ancestry inference. For BD samples, the predicted ancestry attained 98% accuracy when assigning individuals to categories of NW, SE, and AJ ancestry. However, the ancestry obtained using genotypes from BD samples alone attained only 51% accuracy, which reflects the lack of diversity in the BD samples. In a study of simulated phenotypes with population stratification (max. $\chi^2 = 71.5$, corrected $P = 8.5\times10^{-12}$), application of our method was sufficient to correct for population stratification (max. $\chi^2 = 21.3$, corrected $P > 0.05$). In summary, we present the first publicly available software that use genotypes from a single individual to provide accurate estimates of genetic ancestry.

3414W

A Unified Model of Meiosis Combining Recombination, Non-Disjunction, Interference and Infertility. *H.R. Johnston IV, D.J. Cutler.* Department of Human Genetics, Emory University School of Medicine, Atlanta, GA.

Human male and female recombination rates and patterns differ greatly across the broad scale of human chromosomes. Rates of infertility and non-disjunction differ widely between males and females. No simple cause is known for these observations. To this end, we have created a unified model of meiosis that combines recombination, non-disjunction, interference and fertility. The model correctly predicts the rate of fertility, trisomy 21 occurrences and the number and, most interestingly, the different patterns of recombination between the sexes. The model we create is based on the observation that chiasmata are the mechanism that enables the normal segregation of chromosomes during meiosis. Non-disjunction is the result of a failed segregation event. In our model, non-disjunction occurs both when no chiasmata are present between pairs of non-sister chromatids as well as when multiple chiasmata are present close together between pairs of non-sister chromatids. Other elements of our model include having no chiasmata occur between sister chromatids as well as concluding male meiosis immediately while arresting female meiosis between birth and the mother's age at conception. This period of arrest requires that females begin with far more chiasmata than males. It also allows for physical interference to initiate from anywhere on a chromosome arm. In males, this initiation event is always telomeric. These elements combine to generate the unique patterns of recombination in each gender that have, heretofore, not been explained. They also generate the unique patterns of non-disjunction and infertility, helping to explain why these phenomena are seen far more often in eggs relative to sperm. Overall, this model argues that gross differences between male and female patterns of non-disjunction, infertility, and recombination are substantially the result of the period of meiotic arrest during oogenesis.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3415T

Landscape of DNA methylation in meningiomas. *F. Gao[1], L. Shi[1], J. Russin[2], L. Zeng[1], X. Chang[1], T. Chen[2], S. Giannotta[2], D. Weisenberger[3], G. Zada[2], W. Mack[1,2], K. Wang[1].* 1) Zilkha Neurogenetic Institute, USC, Los Angeles, CA; 2) Department of Neurosurgery, USC, Los Angeles, CA; 3) Epigenome Center, USC, Los Angeles, CA.

Meningiomas are central nervous system tumors that originate from the meningeal coverings of the brain and spinal cord. The vast majority of meningiomas are pathologically benign or atypical, but 3–5% of them become malignant. Several studies have investigated the genetics of meningiomas, but the key molecular pathways involved in malignant transformation remains enigmatic. Additionally, the extent of epigenetic alterations in malignant meningioma remains unknown, despite previous studies on benign/atypical meningiomas. Here we explored the landscape of DNA methylation in 19 meningiomas (11 benign, 4 atypical and 4 malignant). Compared to benign tumors, atypical and malignant tumors show increasing trends of global hypomethylation. Clustering analysis using DNA methylation data readily separates malignant tumors from atypical and benign tumors (but not between benign and atypical tumors), implicating that the DNA methylation pattern may serve as a diagnostic biomarker for malignancy. Biological pathway analysis on genes differentially methylated in malignant and benign tumors suggested that the glioma-related pathway is significantly enriched (p=0.002). Genes hypomethylated at promoters show a general increase of expression level in malignant meningiomas, whereas genes with hypermethylated CpG islands (such as *HOXA6* and *HOXA9*) coincide with the binding sites of polycomb repressive complex (PRC) in early developmental stage. Most of the genes with hypermethylated CpG islands at promoter regions were suppressed in both malignant and benign meningiomas, suggesting the switching of epigenetic machinery on gene silencing from PRC binding to DNA hypermethylation. One exception is *MAL2*, which is highly expressed in normal brain tissue and benign meningiomas, but silenced in malignant tumors, suggesting that DNA hypermethylation of *MAL2* represents de novo gene silencing induced by DNA methylation in malignant meningiomas. In summary, our results suggested that malignant meningiomas have distinct DNA methylation patterns than benign and atypical meningiomas, and that some of the differentially methylated genes may serve as diagnostic biomarkers or candidate causal genes for malignant transformation.

3416T

A Panel of Epigenetic Biomarkers of NSCLC Identified from a Comprehensive Microarray-based DNA Methylation Analysis. *S.Cheng Guo[1], X.Tian Wang[1], J.Bin Xu[2], Felix.Li Jin[1], J.Cun Wang[1].* 1) human genetics, Ministry of Education Key Laboratory of Contemporary Anthropology School of Life, shanghai, China; 2) Department of Cardiac and Thoracic Surgery, Changhai Hospital, Shanghai, China.

Lung cancer, a complex disease involved both genetic and epigenetic changes, is the leading cause of cancer deaths worldwide. Aberrant DNA methylation status is an important event in lung cancer initiation.So DNA methylation has emerged as a highly promising biomarker for early cancer screening or diagnosis. The analysis of DNA methylation-based biomarkers is rapidly advancing, and a large number of potential biomarkers have been identified. However, the sensitivity and specificity of a single site still can't meet the need of the clinical diagnosis or screening, and thus exploring the best combination or biomarker panels to benefit clinical detection precision is pressing. In this study, DNA methylation profiling data of 400 non-small cell lung cancer (NSCLC) associated samples, which included 294 NSCLC, 99 cancer adjacents, and 7 normal lung tissues with detailed clinical information, were obtained. All the datasets were background-corrected and normalized separately, then merged and conducted quantile normalization collectively. Batch effect analysis tools, ComBat, was used to eliminate the batch effect existing in combined datasets. Four different methods based on PCA, Decision tree, LDA, and SVM were used to make clinical classification. Four genes (NTSR-1, GALR1, AGTR1, SLC5A8) was uncovered highly methylated in NSCLC with methylation rates of 61%, 47%, 31%, and 27%, respectively, while no methylation signal was found in normal lung tissues. Hypermethylation of NTSR1 was significantly associated with cancer metastasis (P = 0.0012) and subtype (P = 0.000892). Hypermethylation of GALR1 was also significantly associated with subtype of lung cancer (P = 1.55e-09). Multiple classification methods showed combined DNA methylation biomarkers had great power to distinguish NSCLC from adjacent cancer tissues with average sensitivity and specificity of 92.9% and 96.4% at least after 10 fold cross validation. Another dataset from our lab confirmed the methylation of galanin receptor subtype 2 (GALR1), with a similar methlyation rate of 66%, whose hypermethyaltion in promoter region has been identified in several other cancers and caused its gene inactivation. we have developed a ComBat-Meta pipeline to analyze the array-based methylation data and found four hypermethylated genes in caucasian NSCLC patients with high sensitivity and specificity. The methylation status of two of them was further confirmed in Chinese NSCLC patients.

3417T

Methylome Correlation Structure in Human Aging Brain. *G.P. Srivastava[1,2,3], M.L. Eaton[3,4], B.T. Keenan[1,2,3], L.B. Chibnik[1,2,3], J.A Schneider[5], M. Kellis[3,4], D.A. Bennett[5], P.L. De Jager[1,2,3].* 1) Neurology, Harvard Medical School, Boston, MA; 2) Program in Translational NeuroPsychiatric Genomics, Institute for the Neurosciences, Departments of Neurology and Psychiatry, Brigham and Women's Hospital, Boston, MA; 3) Program in Medical and Population Genetics, Broad Institute, Cambridge, MA 02142; 4) Computer Science and Artificial Intelligence Laboratory (CSAIL), Massachusetts Institute of Technology, Cambridge, MA; 5) Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago IL.

DNA methylation is one of the most studied epigenetic marks. The human genome has a non-uniform distribution of approximately 22 million CpGs mostly concentrated in apparently nonfunctional repetitive elements and a tiny fraction of functionally important genomic regions such as CpG islands. Though methylation of CpG sites at single nucleotide resolution and in CpG dense regions like CpG islands is well studied, inter-CpG relation in terms of their methylation and the correlation structure amongst methylation of different CpG sites across the genome is not well understood. We used genome-wide DNA methylation profiles from 748 frozen human post-mortem brains using Illumina 450K HD infinium technology to explore the correlation structures in the brain's methylome: one-half of samples were used in the discovery screen and the second half of subjects were used for replication. We find that the methylation of neighboring CpGs within a certain physical distance is strongly correlated, and we define such areas of high correlation in “cis” as a methylation block (mBlock). We propose a flexible and efficient definition for an mBlock and have developed an algorithm to explore all such mBlocks along with their boundaries and other characteristics. Limiting the analysis to regions deeply explored with our beadset (density >= 5 CG/kb, 146,041/460,067(~1\4) beadset content) and applying our definition of >75% of probe pairs with r >=0.5, we find 20,355 CG sites distributed in 2,537 mBlocks; 13.9% of surveyed CGs fall in an mBlock. The replication samples validate this correlation structure: r^2=0.996 when comparing the extent of correlation within each mBlock defined in the discovery sample set to that of the replication sample set. The mean mBlock size is approximately 2 kb ranging from 52 bp to ~263 kbps, but large outliers are noted. mBlocks are found throughout each chromosome. Strong correlation of mBlocks in "trans” is also noted and is being characterized. Overall, we provide a methodology for empirical data reduction in DNA methylation data that (1) provides an approach to estimating the number of independent loci sampled by the Illumina beadset in brain to facilitate downstream analysis, (2) enables the design of future, more efficient and informative content for beadsets or other CG screening strategies, and (3) offers one insight into the structure of the methylome that could highlight functional units found in the tissue of cell type being sampled.

3418T

Quantifying and Normalizing Methylation Levels in Illumina Arrays. *H. Xu[1], D. Ryu[1], S. Su[2], X. Wang[2], R. Podolsky[3], V. George[1].* 1) Dept Biostatistics, Georgia Health Sciences University, Augusta, GA; 2) Georgia Prevention Center, Georgia Health Sciences University, Augusta, GA; 3) Ctr Biotechnology & Genomic Medicine, Georgia Health Sciences University, Augusta, GA.

Illumina has developed a 450K BeadArray for examining methylation levels genomewide and enabling methylome-wide association studies. Most analyses of Illumina methylation arrays depend on β, the proportion of total signal that is methylated. Little attention has been paid to finding a summary measure that has good statistical properties for testing for differential methylation. Further, little attention has been paid to the potential need to normalize these data. We proposed the N-value, a type of weighted logistic transformation of that accounts for signal variability among beads.Analyses based on the N-value appear to be more robust based on an obesity data set as well as based on simulations.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3419T

Bone health at age 49–51 years is associated with *IGF2* DNA methylation levels in the Newcastle Thousand Families Study cohort. *C. Potter[1], J.C. McConnell[1], L. Barrett[1], R.M. Francis[2], M.S. Pearce[3], C.L. Relton[1].* 1) Institute of Genetic Medicine, Newcastle University, Newcastle upon Tyne, United Kingdom; 2) Regional Medical Physics Department, Freeman Hospital, Newcastle upon Tyne, United Kingdom; 3) Institute of Health & Society, Newcastle University, Sir James Spence Institute, Newcastle upon Tyne, United Kingdom.

Purpose: Inter-individual variation in DNA methylation is believed to play a role in the development and progression of common complex diseases. The hypothesis that both global and gene-specific DNA methylation is associated with bone health at age 49–51 years was investigated in this study. **Methods:** Markers of global DNA methylation (LINE-1 and LUMA) and *IGF2*promoter methylation were quantified by Pyrosequencing in peripheral blood-derived DNA samples from 215, 127 and 203 individuals, respectively, from the Newcastle Thousand Families Study (NTFS). Associations between DNA methylation and contemporary DXA-based measures of bone health, namely femoral neck shaft angle (NSA), bone mineral density (BMD) and bone area of the hip and lumbar spine, were examined. Data were analysed using linear regression, adjusted for significant covariates (including sex, height and/or weight). Presented regression coefficients indicate the unit increase in the outcome variable (NSA:°c, BMD: g/cm2, area: cm2) per unit increase in methylation (% or ratio). No corrections were made for multiple testing. **Results:** *IGF2* methylation demonstrated a significant association with hip BMD (coefficient (95% CI): 0.003 (0.000, 0.006), p=0.045), the significance of which was attenuated following adjustment for weight (0.002 (0.000, 0.005), p=0.075). Upon stratification for sex, a significant association was demonstrated in females (0.004 (0.001, 0.007), p=0.012) but not males (-0.001 (-0.006, 0.005), p=0.762). There was no evidence of a statistical interaction between *IGF2* methylation and sex upon hip BMD. *IGF2* methylation also demonstrated association with spine BMD (0.003 (0.000, 0.007), p=0.042), which was again attenuated following adjustment for weight and sex (0.003 (0.000, 0.006), p=0.066) and upon stratification for sex (females: 0.004 (-0.002, 0.009), p=0.175, males: 0.003 (-0.001, 0.007), p=0.195). Neither LINE-1 nor LUMA global DNA methylation was associated with hip or spine BMD. Bone area (hip, spine and total) and NSA were not associated with any measure of DNA methylation. **Conclusions:** *IGF2* DNA methylation demonstrated associations with hip and spine BMD in the NTFS population, particularly in females. However, these associations did not appear to be wholly independent of known covariates. Hence, further work is required in larger datasets to elucidate a possible causal relationship.

3420T

NSD1 mutations in Sotos syndrome alter the DNA methylation landscape of genes involved in somatic growth and neuronal transmission. *R. Weksberg[1, 2], H. Jin[2], Y. Chen[2], D. Grafodatskaya[2], M. Inbar-Feigenberg[1], BHY. Chung[3], C. Cytrynbaum[1], S. Choufani[2].* 1) Div Clin & Metabolic Gen, Hosp Sick Children, Toronto, ON, Canada; 2) Program in Genetics and Genome Biology, Hosp Sick Children, Toronto ON, Canada; 3) University of Hong Kong, Hong kong.

Sotos syndrome (OMIM 117550) is a rare genetic overgrowth disorder associated with malformations and neurodevelopmental problems including intellectual and behavioural issues. Individuals with Sotos syndrome are haploinsufficient for NSD1 (nuclear receptor-binding SET domain protein 1), a histone lysine methyltransferase important for multiple aspects of normal development. Recently, NSD1 was shown to bind near various promoter elements, regulating multiple genes via interactions with H3K36me and RNA polymerase II. These data suggest a potential role for NSD1 in regulating transcription initiation via epigenetic mechanisms such as DNA methylation (DNAm). We utilized a novel approach to identify the functional impact of NSD1 mutations on the methylome of patients with Sotos syndrome. Bisulfite modified blood DNA from 6 NSD1 mutation- positive cases was investigated by genome-wide profiling of DNAm using the Illumina Human Infinium 27K microarray covering the promoter of 14500 genes. Differential DNAm analysis using Mann-Whitney and FDR identified a total of 1285 CpG sites with loss of DNAm and 54 CpGs with gain of DNAm that were significantly different (>10%) compared to controls (n=36). To investigate whether these DNAm differences are NSD1 specific, we examined the DNAm profiles of 2 other overgrowth syndromes Simpson-Golabi-Behmel syndrome (n=5) with mutations in GPC3 and PTEN Hamartoma Tumor syndrome (n=6)with mutations in PTEN using data from the Illumina 27k array. Differential DNAm analysis did not identify any significant DNAm at any of the 27578 CpG sites in these two disorders compared to controls. These data suggest that DNAm changes at specific gene promoters occur in concert with NSD1 haploinsufficiency. Using ingenuity pathway software, we discovered that the significant candidate CpG sites associated with NSD1 mutations are enriched within two main functional networks: mass growth and neurotransmission. These networks are highly relevant to the clinical manifestations observed in patients with Sotos syndrome and highlight the relevance of such DNAm changes. This is the first report of a functional effect of the NSD1 haploinsufficiency on the DNA methylation landscape. Further work is currently underway to identify common sequence motifs in the promoters of genes with altered DNAm signatures in NSD1 haploinsufficiency that will be critical for the characterization of genome-wide binding targets for NSD1.

3421T

Aberrant expression of DNA methylation, cellular proliferation and DNA repair genes in gastric carcinogenesis? *D.Q. Calcagno[1], F. Wisnieski[1], T.B. Brilhante[1], C.O. Gigek[1], E.S. Chen[1], M.F. Leal[1], S. Demachki[2], P.P. Assumpcao[3], R. Artigiani[4], L.G. Lorenco[5], R.R. Burbano[6], M.A.C. Smith[1].* 1) Disciplina de Genética, Departamento de Morfologia e Genética, UNIFESP, São Paulo, Brazil; 2) Servico de Anatomia patologica, Hospital Universitário João de Barros Barreto, UFPA, Belém, Brazil; 3) Servico de Cirurgia, Hospital Universitário João de Barros Barreto, UFPA, Belém, Brazil; 4) Departamento de Patologia, UNIFESP, São Paulo, Brazil; 5) Disciplina de Gastroenterologia Cirúrgica, Departamento de Cirurgia, UNIFESP, São Paulo, Brazil; 6) Laboratório de Citogenética Humana, Instituto de Ciências Biológicas, UFPA, Belém, Brazil.

Aberrant DNA methylation contributes significantly to the development and progression of gastric cancer, one of the leading causes of cancer death worldwide. A greater understanding of methylation status into these processes are necessary to help in gastric cancer patient treatment. The methylation status is controlled and mediated mainly by DNA methyltransferases (DNMT1, DNMT3A and DNMT3B) and methyl-CpG binding domain proteins (MeCP2, MBD1, MBD2, MBD3 and MBD4). We compared quantitative reverse transcriptase PCR real time in eighteen gastric tumors and their paired non-neoplastic gastric tissue samples or gastric tissue from health patients to examine the expression of DNMT1, DNMT3A, DNMT3B, MeCP2, MBD1, MBD2, MBD3 and MBD4. mRNA. We also investigated PCNA, Ki67 and MGMT gene mRNA expression, proliferative biomarkers and repair gene, respectively. In present study, DNMT1A and PCNA were overexpressed in gastric tumors than gastric tissue from health patients (U= 81, p=0.0196; U=5, p=0.0307, respectively). Additionally, we also observed that DNMT3A and Ki67 were overexpressed in paired non-neoplastic gastric tissue than gastric tissue from health patients (U=6, p=0.0443; U=11, p= 0.0029; respectively). Moreover, DNMT3A, MBD1, MBD2, MBD3, MBD4 and MGMT mRNA expression were reduced in gastric tumors than their paired non-neoplastic gastric tissue (W= 112, p=0.016; W= 108, p=0.027; W= 96, p=0.048; W= 109, p=0.027; W= 107, p=0.0210; W=79, p=0.0143, respectively). These results suggested that early onset gastric cancer involves an increase of mRNA expression of DNMT1A, DNMT3A, PCNA and Ki67. In the future, the aberrant mRNA expression of DNMT1A, DNMT3A, MBD1, MBD2, MBD3, MBD4, PCNA, Ki67, and MGHT may become valuable diagnostic and prognostic markers as well as potential therapeutic targets for gastric cancer.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3422T

Novel MicroRNAs and Epigenetic Biomarkers for Gastrointestinal Stromal Tumors (GIST). *D. Glavac[1], M. Ravnik-Glava{{cak}}[1,2].* 1) Department of Molecular Genetics, Faculty of Medicine, University of Ljubljana, Vrazov trg 2, 1000, Ljubljana, Slovenia; 2) Institute of Biochemistry, Faculty of Medicine, University of Ljubljana, Vrazov trg 2, 1000, Ljubljana, Slovenia.

Gastrointestinal stromal tumors (GISTs) are the most common mesenchymal neoplasms in the gastrointestinal tract. A small minority of GISTs are associated with hereditary syndromes. Large-scale screening of multiple types of molecular aberrations (e.g., mutations, copy number variations, DNA methylations, gene expressions, microRNA expressions) becomes increasingly important in the prognosis and study of cancer, including GISTs. GIST is primarily defined by activating mutations in the KIT or PDGFRA receptor tyrosine kinases. The mutations cause functional changes in KIT and PDGFRA proteins, usually leading to ligand-independent dimerization and constitutional activation. Whereas oncogenic KIT or PDGFRA mutations seem vital to promoting neoplastic transformation, additional genetic and epigenetic alterations are probably required to explain the wide spectrum of clinical behaviour in GIST. In our retrospective study of 126 patients with GIST we investigated KIT and PDGFRA mutational status and further evaluated novel potential biomarkers found by whole genome methylation studies and microRNA microarray analysis. We identified aberrant methylation status using eleven methylation-sensitive CpG islands in GIST with and without mutations. CIMP (CpG islands methylator phenotype, which is defined as methylation involving more than three gene promoters) was present in almoust all GIST without mutation, although there was no statistical difference between GIST with or without gene mutations. On the other hand, CIMP was found in 45 % of GIST with c-kit or PDGFRA gene mutations. In terms of risk categories, CIMP was present in 50% of low-risk GIST and in 60 % of high-risk GIST. Aberrant methylation of multiple gene promoters, as well as the c-kit and PDGFRA genes, therefore plays an important role in the tumorigenesis of GIST. Thirty-five miRNAs were found to be differentially expressed in terms of localization and mutation status. Differential expression was further analysed and confirmed for six miRNAs by qRT-PCR in 39 additional GISTs. Differentially expressed miRNAs were functionally mapped to KIT/PDGFRA signalling and G1/S-phase transition of the cell cycle, revealing several predicted miRNA/mRNA interactions for nine gene targets from KIT/PDGFRA signalling. MicroRNAs and their regulation of gene expression in a tissue-specific manner have also potential application for use as therapeutic targets or disease biomarkers.

3423T

The epigenetic alteration of PPP1R3C in the colorectal cancer. *S.K. Lee[1], J.W. Moon[1], J. Kim[2], H.S. Kim[1], S.H. Park[1].* 1) Dept Anatomy, KOREA Univ, Seoul, South Korea; 2) Dept of General Surgery, Korea Univ, Medical Center, Seoul, South Korea.

Epigenetic DNA methylation is an alternative mechanism to genetic events lead to the inactivation of gene expression. The promoter hypermethylation of gene has been frequently observed during the pathogenesis of colorectal cancer (CRC). We analyzed the methylation profile of 27,578 CpG sites spanning more than 14,000 genes in 20 CRC and 20 adjacent non-tumor containing tissues with methylation bead chip array technology. We identified 125 CpG sites located in promoter CpG islands that were significantly hypermethylated in CRC compared with adjacent normal tissues ($p<0.0001$). Protein phosphatase 1 regulatory subunit 3C (PPP1R3C) is one of hypermethylated genes at the CpG island in CRC tissues. PPP1R3C contributes to glycogen accumulation and is reported the reduced mRNA expression in 35–59% of the melanoma cell lines compared to melanocytes correlated with a high proportion of promoter methylation. We verified the down expressed PPP1R3C in CRC tissues using reverse transcription polymerase chain reaction. Moreover, we figured out the hypermethylated status of PPP1R3C in 3 CRC cell lines compared with normal colon cell line through methylation bead chip array at the 10 promoter CpG sites. The down-regulated expression level of PPP1R3C was also observed both mRNA and protein levels in 2 CRC cell lines, DLD1 and LoVo and the suppressed expression was recovered by treatment of 5-Aza-2'-deoxycytidine, a DNA methyl-transferase inhibitor. These results show that promoter methylation of PPP1R3C gene is frequently occurred and gene silencing may play a remarkable role in colon carcinogenesis.

3424T

EGCG applications on HT-29 and MCF-7 cell lines and evaluation of tumor suppressor gene methylation. *O. Ozer[1], Y.K. Terzi[1], O. Darcansoy Iseri[2], Z. Yilmaz Celik[1], F.I. Sahin[1].* 1) Department of Medical Genetics, Baskent University Faculty of Medicine, Ankara, Turkey; 2) Baskent University Institute of Transplantation and Gene Sciences, Ankara, Turkey.

Cathechins are most active constituents of green tea which include epigallocatechin 3-gallate (EGCG), epigallocatechin, epicatechin-3-gallate (ECG), and epicatechin. EGCG is an antitumor molecule and it shows this activity by binding to the active center of methyl transferase enzyme. Previous studies showed that EGCG decreased DNA hypermethylation and is a potential anti-cancer drug for cancer prevention or treatment. Promoter hypermethylation is a basic mechanism of gene silencing especially for tumor suppressor and DNA repair gene inactivation. It has been shown that methylation of promoter regions impressed in different stages of carcinogenesis. In this study, we aimed to analyze the effect of EGCG on methylation status of tumour suppressor genes in cancer cell lines, HT-29 and MCF-7. MTT assay was utilized to determine EGCG concentrations for each cell line. HT-29 and MCF-7 cells were incubated with 10μM, 20μM and 50μM, and 1μM, 5μM and 10μM EGCG for 48 hours, respectively according to the MTT results. Following incubation with or without EGCG, DNA was isolated from control and study cells. Methylation status was determined by a commercially available methylation spesific MLPA kit, ME002 tumour suppressor-2 probe mix (MRC Holland, Amsterdam, the Nederlands) for 25 different tumor suppressor genes, according to the manufacturer's instructions. In initial samples, promoter regions of CDH13, GATA5 and RARB genes were found to be hypermethylated in the MCF7 cell line. WT1, CHFR, ESR1, and PAX5 genes were hypermethylated in the HT-29 cell line in addition to these three genes. Change in methylation status was not detected in our study samples after EGCG application. Our results suggested that methylation status of tumor suppressor genes did not change with different EGCG doses. As it has been clearly shown that methylation is an important epigenetic factor on behavior of cancer cells and a potential target for epigenetic therapy, further studies would be more informative with different cancer cell lines and large gene methylation profiles to arrive at definite conclusions.

3425T

Methylation, mutational and expression analysis of IL6, IL8 and IL1B on the different tumor tissues. *M. Poulin, A. Meyer, G. Gonzalez, J. Xu, L. Yan.* EpigenDx, Hopkinton, MA.

Interleukins are a group of cytokines that were first seen to be expressed by white blood cell. Inflammatory pathways stimulated by cytokines which are meant to defend the organism against infection and injury, can promote an environment which favors tumor growth and metastasis. Supporting this is the fact that alteration in their expression have been shown to correspond with increased incidence of a variety of cancers. We have isolated the DNA from ten tumor tissues and their adjacent normal tissue from four different tumor types; ovarian, breast, colon and cervical. By utilizing a combination of bisulfite sequencing, quantitative mutation analysis, and expression analysis, we show that there are distinct patterns of methylation levels between IL6, IL18 and IL1B in each of the cancers observed. IL18 showed a dramatic decrease in promoter methylation in all tumor types. In IL6, specific regions within the promoter are differentially methylated inter- as well as intra-genetically. The normal tissue distal promoter shows very high levels of methylation while the proximal promoter methylation levels are very low. When the distal promoter is examined between normal and tumor tissue, there is a distinct hypomethylation observed in all types with the greatest difference being in ovarian and breast tumors. There is very little variation between the normal and tumor tissues in the proximal promoter. IL1B methylation is very variable, where certain cancers and regions within the promoter show hypermethylation while others show hypomethylation. We have also analyzed a number of SNPs within these genes and show a possible correlation with the methylation and /or expression of the genes and the SNP genotypes. These data may provide information that would develop biomarkers for different cancers.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3426T

Methylation Profiling of Tumor Suppressor Genes in Invasive Ductal Breast Carcinoma. *Y.K. Terzi[1], F.I. Sahin[1], O. Ozer[1], Z. Yilmaz Celik[1], B. Demirhan[2], M.C. Yagmurdur[3].* 1) Department of Medical Genetics, Baskent University Faculty of Medicine, Ankara, Turkey; 2) Department of Pathology, Baskent University Faculty of Medicine, Ankara, Turkey; 3) Department of General Surgery, Baskent University Faculty of Medicine, Ankara, Turkey.

Breast cancer is the most common cancer and one of the leading causes of cancer death in females worldwide. Genetic and epigenetic changes play pivotal roles in the development and progression of breast cancer. One important epigenetic mechanism is the methylation of genomic DNA. CpG-islands are located in or near the promoter region or other regulatory regions of approximately 50% of human genes. Many tumors have excessive methylation of CpG islands. Silencing of tumor suppressor genes through promoter hypermethylation is known to be a common event in carcinogenesis, including breast cancer. Its reversible nature makes methylation an attractive biomarker for better diagnosis and prognosis. In this study, we examined the promoter methylation status of 25 tumor suppressor genes in formaline fixed parafin embedded tissues from 18 patients diagnosed with invasive ductal carcinoma. Methylation status was determined by a commercially available methylation specific MLPA kit, ME002 tumour suppressor-2 probe mix (MRC Holland, Amsterdam, the Nederlands) for 25 different tumor suppressor genes, according to the manufacturer's instructions. According to our analysis results, GATA5, CDH13, BRCA2, MSH6, CD44 and CDH13 were found to be hypermethylated in invasive ductal breast carcinoma tissues. Promoter hypermethylation of these genes in breast carcinoma tissues, except GATA5, has been reported previously. We found hypermethylation of GATA5 promoter in 3 of the tissues in the current study. Although altered expression of GATA5 has found to be associated with different tumors, its function in human cancers including breast cancer is not well defined yet. Methylation of tumor suppressor genes is known as an early hallmark of carcinogenesis. Determining the methylation status of genes in different types of breast cancer tumors will help us to unravel disease pathogenesis, and be able to anticipate the prognosis. In addition, identified genes can be used as early biomarkers in preventive medicine.

3427T

Whole genome bisulfite sequencing of acute lymphoblastic leukemia cells. *P. Wahlberg[1], J. Nordlund[1], O. Karlberg[1], C. Bäcklin[2], U. Liljedahl[1], E. Forestier[5,4], M. Gustafsson[2], G. Lönnerholm[3,4], A-C. Syvänen[1].* 1) Molecular Medicine, Department of Medical Sciences, Uppsala University, Sweden; 2) Cancer Pharmacology and Computational Medicine, Department of Medical Sciences, Uppsala University, Sweden; 3) Department of Children's and Women's Health, Uppsala University, Sweden; 4) For the Nordic Society of Pediatric Hematology and Oncology (NOPHO); 5) Department of Medical Biosciences, University of Umeå, Umeå, Sweden.

Acute lymphoblastic leukemia (ALL) is the most common pediatric cancer in the developed countries. With modern treatment protocols children with ALL have a good prognosis, but for unknown reasons, approximately 20% of patients do not respond to treatment. DNA methylation of CpG sites plays an important role in cell differentiation and establishment of cell identity by regulation of gene expression. Aberrant methylation patterns have been observed in a wide variety of cancers. We and others have previously documented large differences in the DNA methylation patterns of ALL epigenomes compared to control immune cells, and common methylation patterns have been identified in ALL subtypes. Because of its importance for cellular function, abnormal methylation patterns could play a major role in ALL pathogenesis. Hence, epigenomic analysis can shed light on specific genes and pathways involved in ALL pathogenesis. Bisulfite sequencing is considered as the golden standard for methylation analysis and offers in combination with high-though sequencing the possibility to determine genome-wide cytosine methylation levels at base-pair resolution. In addition, bisulfite sequencing provides information on haplotype-specific methylation. To determine the patterns of methylation in ALL cells we have generated whole genome bisulfite sequencing data at high coverage to characterize in detail the methylomes of four samples, representing different subtypes of ALL. We will present a detailed view of the methylation patterns in the four bisulfite-sequenced ALL patients together with analysis of RNA-Seq data that we will use to infer how methylation affects transcription. We will also investigate the occurrence of allele-specific methylation and its role in allele-specific gene regulation.

3428T

Integrative genomic and epigenomic profiling of bladder cancer. *A. Woloszynska-Read[1], D. Wang[2], L. Yan[2], Q. Hu[2], J. Wang[2], W. Bshara[3], J. Conroy[4], K. Guru[5], C. Morrison[3], S. Liu[2], D.L. Trump[6], C.S. Johnson[7].* 1) Cancer Prevention and Population Science, Roswell Park Cancer Institute, Buffalo, NY; 2) Biostatistics & Bioinformatics, Roswell Park Cancer Institute, Buffalo, NY; 3) Pathology, Roswell Park Cancer Institute, Buffalo, NY; 4) Center for Genomic Medicine, Roswell Park Cancer Institute, Buffalo, NY; 5) Urology, Roswell Park Cancer Institute, Buffalo, NY; 6) Medicine, Roswell Park Cancer Institute, Buffalo, NY; 7) Pharmacology and Therapeutics, Roswell Park Cancer Institute, Buffalo, NY.

Bladder cancer (BC) is among the most common and prevalent malignancies worldwide, and has remained so for over 30 years. The development of BC is associated with the accumulation of multiple genetic and epigenetic changes. The aim of this study was to characterize BC tissues comprehensively using state-of-the-art DNA methylation and genome sequencing methods, and conduct integrative analyses to identify chromosomal regions affected by genetic and/or epigenetic changes contributing to invasive phenotypes of the disease. We performed a new generation DNA methylation profiling study of 106 bladder tumors and 52 adjacent normal tissues. Our initial findings demonstrate that DNA methylation patterns of more than 450K interrogated CpG loci reliably distinguish tumor and non-malignant bladder tissues, with tumors characterized by global hypermethylation in CpG islands and hypomethylation in CpG shores, shelves, and open seas. Unsupervised clustering of DNA methylation profiles using a recursively portioned mixture model (RPMM) revealed three distinct DNA methylation clusters for BC tissues. The Wilcoxon signed-rank test and the Wilcoxon rank-sum test were used to evaluate differences in DNA methylation β-values for each probe between 52 tumors and 52 matched adjacent-non-malignant tissues, and between the independent 54 tumors and the same 52 non-malignant tissues, respectively. We identified a unique core set of differentially methylated genes enriched for homeobox transcription factors and cell-cell adhesion pathway. Using exome sequencing data from a subset of samples profiled by 450K array, we identified a number of genes harboring non-silent somatic mutations, with MLL2 mutated in more than 40% of tumors. Our study is the first to report MLL2 as a frequent target of somatic mutation in BC. MLL2 is a chromatin modifying protein involved in gene expression and epigenetic control of active chromatin states, and MLL2 mutation may have an effect on DNA methylation patterns in BC. Mutation screening on a set of discovered mutated genes in the remaining samples is currently ongoing to integrate the observed mutations with DNA methylation clusters. We are also examining the clinical and pathologic stage and history of these tumors to discern potential genetic and epigenetic associations. This study could reveal important genomic and epigenomic differences between BC and bladder non-malignant tissues, and define diagnostic, prognostic, and therapeutic targets.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3429T

Differential and sex-specific DNA methylation of the placental *LEP* promoter in early onset preeclampsia. *K. Hogg[1], J.D. Blair[1], M.S. Peñaherrera[1], K. Louie[1], P. von Dadelszen[2], W.P. Robinson[1].* 1) Dept. of Medical Genetics, University of British Columbia, Vancouver, BC, Canada; 2) Dept. Obstetrics and Gynaecology, University of British Columbia, Vancouver, BC, Canada.

Preeclampsia is a common disorder characterized by poor placentation, maternal hypertension and proteinuria and is often associated with fetal intrauterine growth restriction (IUGR). Maternal biomarkers of preeclampsia include elevated plasma leptin and placental *LEP* expression is upregulated. We previously showed that the *LEP* gene promoter is differentially methylated in an imprinted manner in placenta. We hypothesize that altered DNA methylation in the placental *LEP* promoter provides a molecular mechanism for altered expression in these patients. To measure methylation of CpG dinucleotides within the *LEP* promoter, bisulphite pyrosequencing was performed in third trimester placentae from control (n=111), early onset preeclampsia (EOPET; n=19), late onset preeclampsia (LOPET; n=18) and IUGR (n=13) pregnancies. Candidate CpGs were selected based on their proximity to the transcription start site (TSS), TATA box, SP1 and C/EBP transcription factor binding motifs, and a hypoxia response element (HRE). Raw methylation values were adjusted for gestational age and data was additionally stratified by infant gender. *P*-values were adjusted for multiple comparisons. Hypomethylation of the *LEP* promoter was observed in EOPET placentae compared to controls, but not LOPET or IUGR cases. A >10% loss of methylation was observed at CpG sites +9, +14 and +20 bp upstream of the TSS ($P<0.001$). At the TSS (-1; $P<0.01$), and the SP1 (-19) and C/EBP (-51) binding sites, there was a >5% loss of methylation. No methylation change was observed proximal to the HRE (-115,-118). Upon gender-stratification, methylation of the CpG site contained within the C/EBP consensus sequence was significantly hypomethylated in females ($P<0.01$) but not males, compared to controls. Hypomethylation of the placental *LEP* promoter in EOPET, possibly due to loss of methylation of the imprinted allele, may be a mechanism by which *LEP* gene expression is altered in these placentas. Leptin is known to promote proliferation and survival of trophoblast cells, therefore methylation differences may be explained by differing cell populations. The absence of methylation changes in LOPET confirms the differing underlying etiology between early and late onset preeclampsia. Lack of change in IUGR placentas suggests that growth restriction is not a contributory factor. Finally, these results indicate placental sexual dimorphism that may confer altered susceptibility and/or adaption to pathological events.

3430T

A Genome-wide DNA Methylation Association Study of Young-Onset Hypertension in Han Chinese Population of Taiwan. *K.M. Chiang[1,4], H.C. Yang[2], J.W. Chen[3], W.H. Pan[4].* 1) Institute of Biomedical sciences, Academia Sinica, Taipei, Taiwan; 2) Institute of Statistical Science, Academia Sinica, Taipei, Taiwan; 3) National Yang-Ming University School of Medical and Taipei Veterans General Hospital, Taipei, Taiwan; 4) Institute of Biomedical Science, Academia Sinica, Taipei, Taiwan.

Hypertension is a common and complex disorder. Young-onset hypertension (YOH) may be a more feasible target disorder to investigate than the late-onset one due to its stronger genetic component. We performed a two stage genome-wide Methylation association study to map YOH susceptibility methylated loci. In the first stage, we analyzed 45 YOH cases and 45 normotensive controls with age, sex and BMI adjusted in all the analyses. In the second stage, an independent set of 205 YOH cases and 205 normotensive controls were used to validate the findings from the first stage. Illumina HumanMethylation450 BeadChip (> 485,000 methylation sites) was used to generate the β value, which is proportional to the degree of methylated state of any particular loci. SAS and PARTEK Genomic Suit were used for further statistical analyses. We will also integrate the genome-wide CNVs and gene expression data with methylation data to see the functional effect of methylation. Illumina HumanHap 550K Genotyping BeadChip was used for CNV calling and Phalanx Human OneArray v5.1 (30,275 transcripts) was used to measure the whole genome gene expression profiles. In addition to the single locus tests, we will also perform the gene set analyses based on 4277 Gene Ontology categories, 299 BioCarta pathways and 128 KEGG pathways. Several gene sets have been identified in our study.

3431T

Genome wide DNA methylation profiling of monozygotic twins discordant for trisomy 21. *A. Letourneau[1], M.R. Sailani[1], F. Santoni[1], D. Robyr[1], P. Makrythanasis[1], A. Sharp[2], S.E. Antonarakis[1,3].* 1) Genetic Medicine & Development, University of Geneva Medical School, Geneva, Switzerland; 2) Department of Genetics and Genomic Sciences, Mount Sinai School of Medicine, NY 10029 USA; 3) iGE3 institute of Genetics and Genomics of Geneva, Switzerland.

DNA methylation is essential in mammalian development and has an effect in gene expression. We have hypothesized that methylation differences induced by trisomy 21 (T21) contribute to the phenotypic characteristics and heterogeneity in T21. In order to determine the methylation differences in T21 without the interference of the genomic variation, we have used fetal fibroblasts from monozygotic twins discordant for T21. We also used fibroblasts from monozygotic twins concordant for T21, normal monozygotic twins without T21, and unrelated normal and T21 individuals as controls. We applied Reduced Representation Bisulfite Sequencing (RRBS) to generate nucleotide resolution of DNA methylation based on high throughput sequencing (HiSeq 2×100bp) between each pair of twins. An initial analysis of MZ twins discordant for T21 identified 66,491 differentially methylated C nucleotides out of 2,316,600 interrogated nucleotides (DMCs) (FDR<0.01 and at least 25% methylation difference per C) of which 38%, 21%, and 41% reside in gene bodies, promoters, and intergenic regions respectively. The KEGG gene ontology analysis of genes harboring DMCs, showed an enrichment for embryonic morphogenesis (BH_FDR=1.8E-8), neuron differentiation (BH_FDR = 5.6E-7), and cell projection organization (BH_FDR = 2.7E-6); pathway analyses showed an enrichment for genes involved in cancer (BH_FDR=7.6E-5), MAPK signaling (BH_FDR=2.4E-4), and axon guidance (BH_FDR=2.7E-3). The study of methylation differences in monozygotic twins discordant for genomic abnormalities is a promising approach to understand the molecular mechanisms of aneuploidies.

3432T

Genome-wide quantitative DNA methylation analysis of imprinted DMRs in patients with Beckwith-Wiedemann Syndrome by MALDI-TOF MS technology. *T. Maeda, K. Jozaki, H. YAtsuki, K. Higashimoto, H. Soezima.* Biomolecular Sciences, Faculty of Medicine, Saga University, Saga, Saga, Japan.

Beckwith-Wiedemann Syndrome (BWS) is a genomic imprinting disorder characterized by macroglossia, macrosomia, and abdominal wall defect and a predisposition to embryonal tumor in childhood. Dysregulation of imprinted growth regulatory genes due to epimutation, such as loss of methylation at KvDMR (KvDMR-LOM) and gain of methylation at H19-DMR (H19DMR-GOM) within 11p15.5, is the major cause of BWS. It has been reported that BWS with the epimutation also showed hypomethylation at imprinted differentially methylated regions (DMRs) outside of 11p15.5 including ZAC and GNAS loci. However, there has been no report of a genome-wide analysis of imprinted DMRs in BWS patients. In this study, we performed quantitative DNA methylation analysis of 34 imprinted DMRs scattered throughout human genome in 47 BWS patients (37 with KvDMR-LOM, 10 with H19-GOM) by MALDI-TOF MS technology. We also used pyrosequencing to confirm the aberrant methylation detected by MALDI-TOF MS. A strong correlation between the results of MALDI-TOF MS and that of pyrosequencing was found (p = 1.6 × 10-77). By the genome-wide analysis, we identified aberrantly methylated DMRs outside of 11p15.5, which were not reported previously, in both patients with KvDMR-LOM and H19DMR-GOM. In patients with KvDMR-LOM, 14 of 37 patients (37.8%) showed the ectopic aberrant methylation, and about half of patients showed it at plural DMRs. It has been reported that most of all the ectopic aberrant methylation was hypomethylation and occurred at maternally methylated DMRs. However, our result showed hypermethylation at both maternally and paternally methylated DMRs. In patients with H19DMR-GOM, 3 of 10 patients (30.0%) showed the ectopic aberrant methylation. Six showed hypermethylation at either or both of IGF2-DMR0 and IGF2-DMR2, suggesting coordinate biallelic expression of IGF2 by H19-DMR and IGF2-DMR0 and/or IGF2-DMR2. Frequency of the ectopic aberrant methylation among all DMRs analyzed in KvDMR-LOM patients (3.73%) was significantly higher than that in H19DMR-GOM patients (1.03%) (p = 0.0199, chi-square test). Since it is considered that KvDMR-LOM and H19DMR-GOM occur during pre- and post-implantation, respectively, the significant difference suggested that methylation of imprinted DMRs was relatively unstable in pre-implantation stage in comparison with post-implantation stage.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3433T

Epigenomic features and transcription factor binding sites associate with hotspots of genomic instability in human germline and in cancer. *C. Coarfa[1], C.S. Pichot[2], A.R. Jackson[1], A. Tandon[1], S. Raghuraman[1], S. Paithankar[1], A.V. Lee[3], S.E. McGuire[2], A. Milosavljevic[1].* 1) Molecular and Human Genetics, Baylor College of Medicine, Houston, TX; 2) Molecular and Cellular Biology, Baylor College of Medicine, Houston, TX; 3) Pharmacology and Chemical Biology, University of Pittsburgh, Pittsburgh, PA.

Mechanisms of copy-number and other structural variation in human germline and in somatic cancer cells are poorly understood. A large fraction of mutability that selectively affects specific hotspots not associated with NAHR remains to be explained. We recently demonstrated strong association of selective structural mutability with genomic hypomethylation (Li et al, PLoS Genetics 2012 8(5) e1002692). The hypomethylation associates with structurally polymorphic and fast-evolving loci. Rare CNVs occurring in the genomes of individuals diagnosed with schizophrenia, bipolar disorder, and developmental delay and de novo CNVs associated with autism are significantly more concentrated within the hypomethylated loci. We here extend these original findings in the following two directions. (1) To increase resolution and scope, we analyze fifty mate-pair sequencing genome profiles from the pilot 1000 Genome Project and high-resolution breakpoint maps derived by sequencing thirty breast and prostate cell lines and tumor samples, including six mate-pair profiles derived in the course of this project. Analysis of these independent data sets confirms that genomic breakpoints originating in human germline strongly associate with germline-specific genomic hypomethylation in contrast to cancer-associated somatic structural variants which do not. (2) To examine potential role of factors other than methylation, we analyze the variants in the context of the Human Epigenome Atlas, ENCODE, and Cistrome databases, including a total of over 250 relevant epigenomic profiles and transcription factor binding maps. We find that structurally mutable loci in germline associate with ChIP-seq binding sites of Suz12, a key member of the PRC2 Polycomb complex, with H3K27me3 and H3K9me3 histone marks regulated by PRC2 and with ChIP-seq binding sites of other transcription factors regulating the two histone marks. The breakpoints in breast cancer associate with different sets of transcription factor binding sites and epigenomic marks and are particularly concentrated around distal active enhancers. HER2- and HER2- breast cancer cell lines also show distinct distributions of genomic breakpoints. In summary, the observed patterns of association suggest roles for the epigenome and transcription factors in selective structural mutability both in germline and in somatic cells.

3434T

DNA methylation influences lipid and adipokine profiles. *M.A. Carless, H. Kulkarni, D.L. Rainwater, A.G. Comuzzie, M.C. Mahaney, J. Blangero.* Department of Genetics, Texas Biomedical Research Institute, San Antonio, TX, 78227-5301.

Lipid and adipokine imbalances are a critical component of obesity and metabolic syndrome, which are strong risk factors for cardiovascular disease, the leading cause of death worldwide. Although blood lipid levels are heritable, known genetic effects explain only 10–15% of variance within these traits so we set out to identify epigenetic factors that might contribute to variance in lipid and adipokine levels. Using Illumina HumanMethylation450 BeadChips, we generated genome-wide methylation profiles for 1,200 Mexican American individuals from the San Antonio Family Heart Study. In a preliminary subset of 404 individuals, we estimated heritability and performed association analysis of serum lipid and adipokine concentrations using quantitative DNA methylation levels at 482,419 CpG sites. In total, 5,016 CpG sites (~1%) were found to be heritable after Bonferroni correction for multiple tests ($p<1.04\times10^{-7}$). Several associations that approached significance were seen between DNA methylation and serum levels of high density (HDLC) and low density (LDLC) lipoprotein cholesterol, triglycerides, adiponectin and leptin. Notably, triglyceride levels were suggestively associated with methylation in *STARD13* ($p=6.35\times10^{-6}$), a gene previously implicated in lipid variation and metabolic syndrome, as well as with methylation in *GULP1* ($p=1.56\times10^{-5}$), a gene implicated in cellular lipid homeostasis. These genes were also suggestively correlated with HDLC levels, showing an opposite direction of effect. LDLC was suggestively associated with methylation in *PTRF* ($p=4.3\times10^{-6}$), a gene implicated in lipodystrophy. Methylation within *TTC7B* ($p=1.10\times10^{-6}$), a gene implicated in obesity, stroke and cardiovascular disease, and methylation within *FOXA1* ($p=2.30\times10^{-6}$), which may be involved in adipogenesis, were suggestively associated with adiponectin levels. Leptin levels were suggestively associated with methylation in *ASAM* ($p=1.50\times10^{-5}$), a gene involved in adipocyte differentiation and diet-induced obesity. Several other, potentially novel associations were seen, including methylation within genes having transcriptional or translational regulatory functions. We are currently analyzing additional samples and phenotypes to further investigate the role of DNA methylation in lipid and adipokine variation. The identification of genes influencing such variation will contribute to our understanding of the pathophysiology of obesity, metabolic syndrome and cardiovascular disease.

3435T

Identification of epigenetic markers underlying increased risk of type 2 diabetes in South Asians: the EpiMigrant Pilot Study. *A. Drong[1], C. Blancher[1], M.R. Jarvelin[2,3,4,5], C.M. Lindgren[1], P. Elliott[2,3], M.I. McCarthy[1,6], J.S. Kooner[7], J.C. Chambers[2].* 1) Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom; 2) Department of Epidemiology and Biostatistics, Imperial College London, London, United Kingdom; 3) MRC-HPA Centre for Environment and Health, Imperial College London, London, United Kingdom; 4) Institute of Health Sciences, University of Oulu, Oulu, Finland; 5) National Institute of Health and Welfare, Oulu, Finland; 6) Oxford Centre for Diabetes Endocrinology and Metabolism, University of Oxford, Oxford, United Kingdom; 7) National Heart and Lung Institute, Imperial College London, London, United Kingdom.

Type 2 diabetes (T2D) is 2–4 times more common in South Asians than Europeans; this difference in susceptibility is not accounted for by differences in physical activity, diet, obesity or known genetic variants. We report a pilot epigenome-wide association study (EWAS) designed to explore epigenetic factors contributing to increased risk of T2D amongst South Asians. We investigated 45 South Asian men participating in the London Life Sciences Population (LOLIPOP) study, and recruited from GP lists in London (UK). At baseline (age=50.9±9.1 years), all individuals were normal glucose tolerant (NGT, glucose<7mmol/L and HbA1c<6.0%). During follow-up 7.0±0.9 years later, 15 subjects remained NGT, while 30 developed T2D (glucose≥7mmol/L and HbA1c≥6.5%). Of the 30 people with incident T2D, 15 were previously undiagnosed and not on treatment (U-T2D), while 15 had been clinically diagnosed and were on treatment (T-T2D). Whole blood was collected at baseline and follow-up: genomic DNA was extracted from all 90 samples in one batch. Percent methylation (beta) at 485,577 sites was profiled with the Infinium HM450 BeadChip, excluding signals with detection $p<0.01$ and markers with >5% failure rate. We investigated association of DNA methylation with incident T2D and potential confounders including age, body mass index, drug treatment, white blood cell count (WBC) and white cell subsets using linear models. We measured the inflation of associations $\lambda=\text{mean}\beta^2/\text{var}(\beta)$, where β is the regression coefficient. In the baseline samples, we found global differences in beta between incident T2D and controls (U-T2D vs NGT: λ=1.60; T-T2D vs NGT: λ=2.10). In contrast betas were similar between U-T2D and T-T2D (λ=1.01). To assess the impact of aging we compared baseline to follow-up samples as Δbeta per marker; we again found marked differences in Δbeta between people who developed T2D and those who remained NGT (U-T2D vs NGT: λ=1.51; T-T2D vs NGT: λ=1.51). Beta values closely associated with age (λ=2.29) and with white cell differentials (lymphocytes: λ=5.18; neutrophils: λ=2.30; eosinophils: λ=1.59), but not with total WBC or BMI. Associations of beta with T2D persisted after adjustment for these confounders (U-T2D vs NGT: λ=1.46; T-T2D vs NGT: λ=1.97). Our findings support the hypothesis that epigenetic markers may predict or influence risk of T2D in South Asians, but reveal important confounding effects that will need to be controlled for in the design of future T2D EWAS.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3436T

Epigenome-wide association analysis identified aging alteration of DNA methylation level in the genes related to metabolic pathways. *H. Kitajima, K. Yamamoto.* Division of Genome Analysis. Research Center for Genetic Information., Medical Institute of Bioregulation, Kyushu University, Fukuoka, Fukuoka, Japan.

Aging affects various physiological processes and increases susceptibility to diseases. Whether aging alters DNA methylation level, whether genetic polymorphism affects aging-related epigenetic change, and how the aging-related epigenetic change influences disease traits are of great interest. To settle this matter, we performed epigenome-wide association study by analyzing approximate 450K DNA methylation sites using the blood sample of total 384 subjects in Japanese general population. We conducted multistage association analysis between DNA methylation level and age by adjusting with sex and body mass index using general linear model for 96 subjects in each step. We identified 10 age-associated methylation sites which showed P-value < 0.001 in every four sets. The P-value of these sites in combined set (384 subjects) reached conservative Bonferroni threshold (P-value < 1.05×10^{-7} accounting for 473,864 tests). The location of these sites were as follows; the 2 sites (5.8×10^{-36} and 1.5×10^{-15}) in *ELOVL2* which is included in the n-3 polyunsaturated fatty acid (PUFA) metabolic pathway, the 2 sites (2.3×10^{-29} and 1.9×10^{-29}) in *FHL2*, the one site (1.2×10^{-20}) in *OTUD7A*, the 2 sites (9.3×10^{-20} and 4.6×10^{-16}) in *KLF14* which is associated with diabetes mellitus, the one site (2.5×10^{-18}) in *TRIM59*, the one site (4.1×10^{-15}) in *UBE2D2*, *CXXC5*, the one site (1.9×10^{-13}) in 16q24 locus. We further investigated disease association of the age-associated DNA methylation level in terms of hypertension, diabetes mellitus, and dyslipidemia. The one methylation site in *ELOVL2* significantly associated with hypertension (P-value = 4.4×10^{-4}). This association newly highlighted the relationship between the hypertensive state and the n-3 PUFA metabolic pathway which is related to the serum high-density lipoprotein cholesterol level. In addition, we found that the methylation levels of one site in *KLF14* and the one site in *OTUD7A* are significantly association with the genotype of nearby SNPs (linear regression P-value = 3.2×10^{-7} and 3.2×10^{-7}, respectively). Our findings imply that aging and its individual difference would be understood through an extent of epigenetic change of genes involved in metabolic pathways which are essential for maintenance of cellular activity.

3437T

Alterations in methylation status of immune response genes promoters in cell-free DNA during a hemodialysis procedure in patients with diabetic nephropathy and in non-diabetic subjects. *M. Korabecna[1, 2], E. Pazourkova[1, 2], A. Horinek[1, 2], M. Mokrejsova[3], V. Tesar[2].* 1) Faculty of Medicine in Pilsen, Charles University in Prague, Pilsen, Czech Republic; 2) First Faculty of Medicine, Charles University in Prague, Prague, Czech Republic; 3) Third Faculty of Medicine, Charles University in Prague, Prague, Czech Republic.

Background: Elevations of cell-free DNA (cfDNA) concentrations during hemodialysis (HD) sessions were reported in numerous studies. In this study, we focus on methylation status of promoters of 24 immune response genes at the level of plasma cfDNA and their alterations provoked by the HD process in patients with diabetic nephropathy (DN). 20%-40 % of patients with diabetes mellitus (DM) develop DN which belongs to one of the most frequent causes of HD therapy. Then the HD therapy itself could highly modify patient's immune system. Methods: We isolated plasma cfDNA from 20 patients with DN and from16 non diabetic patients before and after a HD session. We examined also cfDNA from 10 healthy volunteers two times per day with 4 h interval. The extent of promoter methylation of 24 genes involved in immune response was examined using the EpiTect Methyl qPCR Array Inflammatory Response and Autoimmunity and cluster analysis (SABiosciences, Qiagen). We compare results obtained in DN patients with methylation patterns in healthy controls and non- diabetic HD patients using STATISTICA software ver.10. Results: We described significant changes in promoter methylation of genes IL17C, IL7, IL13 and IL15 that were promoted in consequence of DN patient's blood contact with artificial surfaces of dialyzer. In interdialytic interval we found significant differences between DN patients and healthy subjects in three genes only (IL13RA1, EDG3 and IL7) but after a HD procedure significant alterations were found in 50% of examined genes. The only highly significant difference between DN and non-diabetic hemodialyzed subjects immediately after HD session was found in the promoter methylation in the IL12B gene which is one of candidate type I diabetes susceptibility gene. Conclusion: The method provides the new tools for evaluation of immune system activity in HD patients. The results can be further correlated with clinical data to contribute to understanding the complex pathogenesis of renal failure and its treatment. Supported by grants no. I/328 of Ministry of Industry and Trade and no. PRVOUK-P25/LF1/2 of Ministry of Education of the Czech Republic.

3438T

Genome-wide DNA methylation study of obesity and type 2 diabetes. *D. Zhou, D. Zhang, Z. Zhang.* Bio-X Institutes, Key Laboratory for the Genetics of Developmental and Neuropsychiatric Disorders (Ministry of Education), Shanghai Jiao Tong University, shanghai, China.

Obesity and type 2 diabetes are complex metabolic diseases, which affected by both environmental and genetic factors. Epigenetic effects, especially DNA methylation, may play very important role in the occurrence and development processes. In this study, we used high-fat diet induced obese and diabetic model mice as experimental material, and used two genome-wide methylation sequencing technology (MeDIP sequencing and MSCC sequencing) to investigate the changes of DNA methylation in the process of obesity and diabetes. We found that the genes, which involved in MAPK pathway, insulin pathway and cancer pathway, have significant changes in DNA methylation.

3439T

Density of inter-individual epigenome variation uncovered by integrative epigenome mapping at single base resolution. *J.J. Lambourne, S. Busche, V. Adoue, T. Kwan, M. Caron, G. Bourque, M. Lathrop, T. Pastinen.* McGill University, Montreal, Quebec, Canada.

Combination of human epigenomic data with transcriptome analyses has principles of tissue specificity and at increased pace assigned function to non-coding DNA elements. Large-scale efforts have mostly focused on revealing principles of tissue differentiation and demonstrated enrichment of phenotype associated genetic variants at non-coding functional elements. Previous studies have linked thousands of such variants also to expression quantitative traits and more recently to population variation in chromatin and transcription factor binding. Epigenome Mapping Centre at McGill (EMC-McGill) is pursuing understanding of tissue-specific epigenomic signatures applying next-generation sequencing (NGS) based approaches ChIP-seq. of chromatin marks, strand-specific RNA-seq. and whole genome bisulphite-seq. (methylC-seq.). To address contribution of inter-individual sequence and epigenomic variation to genome function, EMC-McGill is carrying out detailed epigenome mapping in primary cells related and unrelated donors. In pilot studies we compared low coverage (5x) methylC-seq. to Illumina 450K to interrogate inter-individual CpG-methylation variation and its correlation to other functional genomic data (ChIP-seq./RNA-seq). Overall, >100x more interindividual methylation variation was revealed by methylC-seq. Unbiased methylation seq. observed substantially higher genome methylation as compared to estimates based on targeted 450K analyses, while at all overlapping sites the methylation estimates were (r2 = 0.85 to 0.90) concordant even at low seq. coverage. Furthermore, expression phenotypes correlate substantially better with methylC-seq. as compared to Illumina 450K. On-going analyses are targeted to derive precise estimates of the additional functional genomic variation at chromatin at expression states achieved by methylC-seq. as well as extension of the single-base intergrative analyses to multiple human primary cells including longitudinal sampling in population. Early results suggest that studies applying targeted tools may not appreciate full spectrum of epigenomic variation in populations.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3440T

Allelic-dropout during PCR in the promoter of an imprinted gene potentially caused by interaction between G-quadruplex structures and DNA methylation. *M.A. Kennedy[1], A.J. Stevens[1], S. Stuffrein-Roberts[1], A. Gibb[1], K. Doudney[1], A. Bagshaw[1], A. Aitchison[1], M.R. Eccles[2], V.V. Filichev[3], P.R. Joyce[4].* 1) Dept Pathology, University of Otago, Christchurch, Christchurch, New Zealand; 2) Department of Pathology, University of Otago, Dunedin School of Medicine, Dunedin, New Zealand; 3) Institute of Fundamental Sciences, Massey University, Palmerston North, New Zealand; 4) Department of Psychological Medicine, University of Otago, Christchurch, New Zealand.

While exploring genetic variability in the maternally imprinted human gene MEST, which plays a role in development and maternal behaviour(1–3), we discovered three SNPs in a CpG island located at the 5' end of the gene(4). These SNPs are in total linkage disequilibrium, such that there are two haplotypes in the human population. When these SNPs were typed in many subjects we observed a non-Mendelian pattern whereby only one haplotype was detectable in each subject, never both, despite the use of multiple methods. Experiments in which genomic DNA from different subjects was mixed prove that the assays were capable of detecting both haplotypes simultaneously(4). The absence of observed heterozygotes was most likely due to allelic dropout of one allele, although dropout of different alleles in different subjects was puzzling. We tested the effect of methylation in PCRs where amplified products for both MEST haplotypes were mixed to mimic heterozygous templates, and found that artificially methylated templates dropped out of these reactions. This suggested that DNA methylation is important in dropout, but alone this seems unlikely to completely account for it. Sequences in the MEST promoter region appear capable of forming non-B DNA structures known as G-quadruplexes (G4). Circular dichroism and native gel electrophoresis with oligonucleotides confirmed that at least three MEST promoter regions are capable of forming marked G4 structures, and that the corresponding complementary strands for each region were capable of forming a different secondary structure known as an i-motif. These data lead us to suggest that methylated DNA causes stabilization of G4 structures in vitro, leading to polymerase blockage and allelic dropout of the imprinted allele. This would account for the loss of different alleles in different subjects. Proof of this awaits an assay that does not suffer allelic dropout, so the parental status and methylation pattern of lost alleles can be ascertained. If this model proves correct, it also raises the question of whether in vivo interactions between DNA methylation and G4 or i-motif structures can occur. 1. Lefebvre et al. (1998) Nat Genet 20, 163–169 2. Kobayashi et al. (1997) Human molecular genetics 6, 781-786 3. Riesewijk et al. (1997) Genomics 42, 236-244 4. Stuffrein-Roberts, S. (2008) Allelic expression patterns in psychiatric candidate genes. PhD Thesis in Pathology, University of Otago, Christchurch.

3441T

Genetics of Global DNA Methylation Patterns in Adipose Tissue from Twins and Its Effect on Gene Expression and Disease: The MuTHER Study. *E. Grundberg[1,2], E. Meduri[1,2], J.K. Sandling[1], A.K. Hedman[3], S. Keildson[3], J. Nisbet[1], M. Sekowska[1], A. Wilk[1], A. Barett[3], K.S. Small[1,2], J.T. Bell[2], E.T. Dermitzakis[4], M.I. McCarthy[3,5,6], T.D. Spector[2], P. Deloukas[1].* 1) Wellcome Trust Sanger Institute, Cambridge, United Kingdom; 2) Department of Twin Research and Genetic Epidemiology, King's College London, London, United Kingdom; 3) Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom; 4) Department of Genetic Medicine and Development, University of Geneva Medical School, Geneva, Switzerland; 5) Oxford Centre for Diabetes, Endocrinology & Metabolism, University of Oxford, Churchill Hospital, Oxford, United Kingdom; 6) Oxford National Institute for Health Research Biomedical Research Centre, Churchill Hospital, Oxford, United Kingdom.

Functional approaches to interpret disease-associations identified in GWAS are in high demand. Many of the disease-associated variants map to non-coding DNA and as a result functional association studies of transcriptome in human tissues targeted to the disease of interest have elucidated layers of impact for common SNPs.As DNA methylation is known to be an important regulator of gene expression the approach is now being expanded to include epigenomic data as well. To this end, we used the IlluminaHumanMethylation450 array to profile 662 adipose tissue samples derived from well phenotyped female twins from the MuTHER project - a major resource ofdetailed genomic (HapMap2) and transcriptomic(IlluminaHT-12)data from multiple tissues.As the majority of the profiled cytosine sites were invariable across individuals we restricted analysis to the 34,685 (10%) most informative sites. Both genome-wide heritability analysis and simple correlations of methylation levels within and between twin pairs (r_{MZ}= 0.96, r_{DZ}=0.94, $r_{unrelated}$=0.92) confirmed evidence for DNA methylation being heritable. To map the underlying genetic (regulatory) effect we performed methylationQTL (mQTL) analysis and found an abundance of common variants (MAF >5%) acting in *cis* with top associations enriched in close proximity to the site (median distance 9kb). The majority (71%) of the cytosine sites associated with an mSNP at 1% FDR were located in CpG islands, shores or shelves. By combining the *cis*-mQTL data with *cis*-eQTLs we noted that 7.6% of the associations were directly overlapping, indicating genetic regulation of gene expression being mediated through DNA methylation.This degree of overlap (~8%, 1% FDR) was further confirmed by association analysis directly comparing methylation patterns of CpG sites in gene promoters with gene expression scores of corresponding genes on the IlluminaHT-12 expression array where an excess of negative correlations ($beta_{median}$=-0.03) was noted. Finally, preliminary observations of a first attempt to integrate disease associations (NHGRI GWAS catalogue) with adipose-derived methylome data indicate an enrichment of mSNPs overlapping with disease loci associated with estrogen-dependent diseases and traits such as age of menarche, ovarian and breast cancer. On-going analysis includes utilizing the twin design to map familial (genetic and common environment) and unique environmental effect of the methylome.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3442T

A methylome-wide MBD-seq study followed by replication in a total of 3,000 schizophrenia case-control samples identifies new disease biomarkers. *K.A. Aberg[1], J.L. McClay[1], S. Nerella[1], S.L. Clark[1], G. Kumar[1], W. Chen[2], A.N. Khachane[1], G. Gao[2], L.Y. Xie[1], A. Hudson[1], J. Bukszar[1], C.M. Hultman[3], P.F. Sullivan[3,4], P.K.E. Magnusson[3], E.J.C.G. van den Oord[1], Swedish Schizophrenia Consortium.* 1) Sch Pharmacy, Ctr Biomarker Res & Personalized Med, Virginia Commonwealth Univ, Richmond, VA; 2) Department of Biostatistics, School of Medicine, Virginia Commonwealth University, Richmond, VA, USA; 3) Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden; 4) Department of Genetics, University of North Carolina at Chapel Hill, NC, USA.

We performed a methylome-wide association study (MWAS) of DNA extracted from whole blood in a schizophrenia case-control sample of 1,500 individuals. We applied the methyl-binding domain 2 sequencing (MBD-seq) approach to enrich for the methylated fraction of the genome followed by next-generation sequencing, which allows for investigation of 30.6 million CpGs in the human genome. To analyze this massive dataset, we developed an analysis pipeline specifically for MBD-seq that includes alignment, quality control (QC), novel estimators for calculating CpG coverage, data reduction combining highly correlated CpGs into blocks, principal component analysis, association testing, bioinformatics annotation and network analysis. The top findings were replicated in an independent schizophrenia case-control sample of 1,500 individuals using targeted pyrosequencing.

After QC and data reduction, 4.2 million high quality blocks were left for methylome-wide association (lambda of 1.013). In contrast to similarly sized GWAS for schizophrenia, our MWAS showed a number of highly significant findings with top p-values of 5.10E-10 and 110 blocks (450 CpGs) with p-values < 9.9E-7. Among the top results were differentially methylated blocks located in: *TBC1D22A* (p=5.10E-10), *PKDREJ* (p= 3.41E-09), *LRRC37A3* (p= 5.10E-09) and *SATB1* (p-value=3.27E-08).

Bioinformatic analyses of the top findings indicate that CpGs associated to schizophrenia are rarely located in CpG islands (<0.1%) but commonly are located in introns (>63.7%). Although replication is still ongoing, several genes on the top list have already replicated including *SATB1, ETS2*, *PHACTR3* and *FNDC3B*. Interestingly, when conducting network analysis, many of the genes among the top results clustered into the same networks. One such example includes the replicated genes *SATB1* and *ETS2*, which interact to regulate chromatin modeling and gene expression. A gene under regulation of the *SATB1/ETS2* network is *BCL2*, which recently was associated with schizophrenia in a large GWAS-meta analysis.

Our replicated results represent one of the first sets of CpG sites identified through MWAS that potentially can be used to improve treatment, diagnosis and disease etiology for schizophrenia patients.

3443T

Genome-wide estimation of DNA methylation heritability in CD4+ T-cells from 80 families. *D. Absher[1], L. Waite[1], H. Tiwari[2], D. Arnett[3].* 1) HudsonAlpha Institute for Biotechnology, Huntsville, AL; 2) Dept. of Biostatistics, University of Alabama at Birmingham, Birmingham, AL; 3) Dept. of Epidemiology, University of Alabama at Birmingham, Birmingham, AL.

DNA methylation is a potent regulator of transcription and is influenced by genetic and environmental factors. Genome-wide studies are now underway to identify DNA methylation patterns that are associated with common human diseases and their risk factors. However, the relative impact of genetic and environmental influences on locus-specific DNA methylation has not been well characterized, and controlling for these influences in epigenetic association studies is crucial to their success. To assess the heritability of CpG methylation and to distinguish allele-specific DNA methylation or methyl-QTLs from more complex epigenetic inheritance patterns, we measured DNA methylation at >480,000 CpGs in CD4+ T-cells from >400 individuals in >80 famiiles. The MHC locus on chromosome 6 was among several regions that displayed enrichment for heritable DNA methylation, and we identified strong cis effects from local genetic variants, as well as inheritance patterns that were independent of local DNA sequence.

3444T

Epigenetic Second Hit on Genebody of PKD1 Leads to Cystic Kidney in ADPKD. *J. Bae[1,2], Y. Woo[3], J. Lee[1], J. Park[3], Y. Kim[2].* 1) KNIH, Osong, ChungCheongbukdo, South Korea; 2) YGI, Seoul, South Korea; 3) Sookmyung W. University, Seoul, South Korea.

Autosomal dominant polycystic kidney disease (ADPKD) is a genetic disorder mostly caused by mutations in PKD1, possibly in a “two-hits” mechanism. To identify the disease-causing mutations of the PKD1 in cyst patients, both genomic and cDNA sequences were examined. Although the PKD1 genomic sequences of the patients revealed an inactivating mutation in some patients, no inactivating mutation was found in the expressed alleles of the PKD1. However, the levels of PKD1 expression were greatly diminished in the patients. To test whether epigenetic modification plays a role in ADPKD pathogenesis, we performed genome-wide methylation affinity purification sequencing (MIRA-Seq) analysis on 3 ADPKD patient cystic kidney samples and examined the disease-associated aberrant DNA methylations. Intriguingly, PKD1 gene was one of the most significantly hypermethylated genes in the patients along with other key regulators of Notch and Wnt signaling pathways, indicating an epigenetic silencing as the second inactivating mechanism of PKD1. The hypermethylation regions were observed as large blocks on several chromosomal locations and mostly enriched in the genebody areas, which showed negative correlation to the expression levels of the associated genes. The epigenetic regulation of the cyst development was further confirmed by the retarded cyst formation of MDCK cells in an in vitro 3D cyst culture system upon treatment of DNA methylation inhibitors, accompanied with the upregulation of the PKD1 expression. Taken together, DNA methylation changes of PKD1 and related regulatory genes in ADPKD appear to promote kidney cyst development.

3445T

Altered DNA Methylation in Preeclamptic Placentas. *J.D. Blair[1,3], R.K.C. Yuen[1,3], P. von Dadelszen[2,3], W.P. Robinson[1,2,3].* 1) Department of Medical Genetics, University of British Columbia, Vancouver, British Columbia, Canada; 2) Department of Obstetrics and Gynaecology, University of British Columbia, Vancouver, British Columbia, Canada; 3) Child and Family Research Institute, Vancouver, British Columbia, Canada.

Preeclampsia (PET) is one of the leading causes of maternal death in the world. It is clinically characterized by hypertension and proteinuria. Physiologically, it is suspected to be initiated by poor placental invasion and maternal artery remodeling. This creates a hypoxic environment, causing abnormal placental gene expression and increased apoptosis in syncytiotrophoblast, which then leads to vascular endothelial damage in the mother. Many of the molecular aspects of PET are still unknown. Epigenetic processes, such as DNA methylation, that are associated with regulation of gene expression may play a role in the molecular etiology of PET. We hypothesize there will be altered DNA methylation of biologically significant genes in placental samples from pregnancies affected by PET. Samples in this study included 10 chorionic villi samples from placentas affected by early-onset preeclampsia (EOPET) and 10 gestational age matched controls. Using the Illumina Infinium HumanMethylation450 BeadChip, we surveyed the DNA methylation status of 450,000 CpG sites across the genome. Using strict criteria of >15% absolute average DNA methylation difference and <5% false discovery rate, 267 CpG sites were significantly different. Interestingly, all of the significant CpGs were hypomethylated in the EOPET samples compared to controls. Most of the significant CpG sites were located in the gene body or in a CpG island shore, rather than in the CpG island of a promoter. Significant genes were followed up bioinformatically using MEME Suite to identify recurrent transcription factor binding motifs, DAVID to investigate enrichment of gene function and STRING to elucidate protein-protein interactions among the significant genes. DNA methylation and gene expression of several candidates are also being confirmed in a larger cohort. . Numerous genes with altered DNA methylation were biologically significant including those associated with i) trophoblast invasion and differentiation (ADAM12, TIMP3, CSF1R), ii) hypoxia (BHLHE40, STC2, PKM2, IDH2) and iii) genes that have previously been shown to be differentially expressed in preeclamptic placentas (CYP11A1, LEP, PAPPA2, GNA12). Hypomethylation of trophoblast-related and hypoxia-related genes suggest DNA methylation plays a significant role in the EOPET development and may be useful for diagnostic purposes. Future studies should focus on elucidating the underlying molecular mechanisms for these changes.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3446T

Highly abnormal methylation of the placental genome in the CBAXDBA model of pregnancy failure. *S. Brown, E. Bonney, L. Brown.* Dept OB/GYN, Univ Vermont, Burlington, VT.

When female mice of the inbred strain, DBA/2 (DBA/2F) are mated with CBA/J males (CBA/JM), few fetal resorptions are noted. However, when the opposite mating is performed (CBA/JF X DBA/2M), up to 50% of pregnancies are resorbed by gestational day 12–14. Because of this, CBA/JF X DBA/2M matings have served as a model system for miscarriage for many years; however, there is no molecular understanding of the cause for fetal loss. F1 embryos from CBA/JF X DBA/2M matings are genetically identical to F1 embryos from DBA/2F X CBA/JM matings, making it unlikely that genetic differences between embryos contribute to the differential rate of loss. Therefore, it is assumed that pregnancy loss is in some way mediated by the maternal environment. As an alternative hypothesis, we have considered the possibility that there may be parent of origin (imprinting) differences between the two classes of genetically identical embryos. In this model, parent of origin (epigenetic) differences (as opposed to maternal environment) would contribute to pregnancy failure. To test this model, we performed a microarray-based analysis of methylation of placental DNA of E9.5 embryos at ~15,000 sites on chromosome 7. We show that DBA/2F XCBA/JM (healthy) embryos have highly consistent methylation profiles, while CBA/JF XDBA/2M (loss-prone) embryos show a striking degree of variability of methylation at thousands of loci, reflecting a highly disordered epigenome. This was assessed by determining the variance of signal intensity across 7 microarray data sets for all ~16K loci. We then compared the mean variance of the two groups using a z test and confirmed a striking degree of variability of methylation in loss-prone pregnancies at E9.5. When the same experiment was performed at E8.5, we found less striking but still highly significant evidence for disordered methylation in the loss prone (CBAXDBA) embryos. Given what is known about the establishment of methylation during embryogenesis, the methylation abnormalities that we see are likely to have been present since implantation, and this suggests that the abnormalities are intrinsic to the embryo. Alternatively, the CBA/J maternal environment may act to alter methylation long before any overt evidence of pregnancy failure. Future experiments will be aimed at understanding whether methylation abnormalities are truly intrinsic to the embryo or depend on the maternal environment.

3447T

DNA Methylation Profiles of Paediatric Obsessive Compulsive Disorder (OCD). *D.T Butcher[1], G.L Hanna[2], D.R Rosenberg[3], R. Weksberg[1,4,5], P.D. Arnold[1,6].* 1) Genetics & Genome Biology, Sickkids Research Inst, Toronto, Ontario, Canada; 2) Department of Psychiatry, University of Michigan, Ann Arbor, Michigan, USA; 3) Psychiatry and Behavioral Neuroscience, Wayne State University School of Medicine, Detroit, Michgan, USA; 4) Department of Paediatrics and the Institute of Medical Sciences, University of Toronto, Toronto, Ontario, Canada; 5) Clinical and Metabolic Genetics, The Hospital for Sick Children, Toronto, Ontario, Canada; 6) Department of Psychiatry, The Hospital for Sick Children, Toronto, Ontario, Canada.

Obsessive-Compulsive disorder (OCD) is a complex, common and debilitating psychiatric disorder. Heritability estimates range from 40% to 65% suggesting that genetic factors play a role in this disorder. However, there are other potential important determinants, such as epigenetics contributing to the OCD phenotype. Epigenetic mechanisms, such as DNA methylation, have increasingly been shown to be important in the etiology of other complex disorders. We hypothesized that there are alterations in DNA methylation patterns in paediatric OCD compared to controls. We used a novel genome-wide approach to identify epigenetic variants in children with OCD. Although few epigenetic studies have been conducted using saliva DNA, epigenetic biomarkers have identified parallel DNA methylation alterations in blood and brain of patients with neuropsychiatric disorders.

Analysis of blood DNA methylation patterns in 10 paediatric OCD patients compared to 10 controls identified both gain and loss of methylation at a number of CpG sites. Candidate DNA methylation alterations identified in this initial investigation are highly relevant to the OCD phenotype and include altered DNA methylation in promoters of genes in glutamate signalling, myelin synthesis and lipid transporters. An expanded cohort of genome-wide DNA methylation profiles for 20 blood and 20 saliva of paediatric OCD patients compared to tissue, age and sex-matched controls is being generated using the Illumina 450K Infinium Methylation BeadChip array. Saliva DNA was collected using Oragene•Dx saliva kits (DNA Genotek). Data is analyzed using the IMA package in R and Genome Studio software from Illumina. Methylation differences of greater than or less than 15% between groups and $p<0.05$ defined candidate genes are being selected for validation using targeted pyrosequencing.

The identification of epigenetic modifications in children with OCD will lead to improved understanding of the pathophysiology and molecular mechanisms leading to this disorder. These epigenetic alterations could provide early diagnostic biomarkers for OCD which may influence treatment strategies, including both pharmacological and cognitive behavioural therapy. Such a personalized medicine approach has the potential to provide tremendous benefits to patients suffering from OCD.

3448T

Complete or partial loss of methylation at the *GNAS* locus identified in patients referred for molecular diagnosis of pseudohypoparathyroidism type 1a. *G.S. Charames[1], S.J. Steinberg[1,2], B.A. Karczeski[1], M.A. Levine[3,4], G.R. Cutting[1,5].* 1) Johns Hopkins DNA Diagnostic Lab, Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD; 2) Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, MD; 3) Division of Endocrinology and Diabetes, The Children's Hospital of Philadelphia, Philadelphia, PA; 4) Department of Pediatrics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA; 5) Department of Pediatrics, Johns Hopkins University School of Medicine, Baltimore, MD.

Pseudohypoparathyroidism (PHP) type 1 is a heterogeneous disorder characterized by resistance to parathyroid hormone. Patients with PHP-1a show impaired responsiveness to additional hormones, plus a constellation of physical findings called Albright Hereditary Osteodystrophy (AHO) that include brachydactyly, short stature and ectopic ossification. Some family members have only AHO and normal hormone responsiveness, a variant termed pseudopseudohypoparathyroidism (PPHP). By contrast, patients with PHP-1b have a more limited phenotype and typically have only PTH resistance and few, if any, features of AHO. PHP type 1 is caused by mutations that reduce expression or function of the imprinted *GNAS* gene, which encodes several proteins including the alpha subunit of the heterotrimeric G protein Gs that couples hormone receptors to activation of adenylyl cyclase. Gsα is biallelically expressed in most tissues, but in some tissues only the maternal allele is transcribed. While both PHP-1a and PPHP are caused by *GNAS* coding region mutations on the maternally-derived and paternally-derived coding regions, respectively, PHP-1b is associated with epigenetic defects, including complete loss of methylation at the *GNAS* exon 1A (e1A) differentially methylated region (DMR), which reduce transcription of the maternal Gsα allele in imprinted cells. Recently, partial loss of e1A methylation has been reported in some patients suspected to have PHP, but the implications of this molecular modification are not known. This report describes a cohort of 116 patients with presumed PHP-1a who were referred to our clinical lab for *GNAS* analysis and were mutation-negative. We examined *GNAS* e1A DMR in these patients by either bisulfite Sanger sequencing or bisulfite pyrosequencing to assess methylation status. Normal methylation is greater than 35% at ^{me}CpG, based on 1SD of the mean of normal controls; complete loss of methylation is less than 5%. Ten patients had a complete loss of e1A methylation suggesting that they actually had PHP-1b. Moreover, 5 subjects had a partial loss of e1A methylation (5–35% methylation). Our results suggest that e1A methylation defects are not uncommon in patients with presumed PHP-1a, and approximately 10% may have PHP-1b. Moreover, the spectrum of methylation defects in *GNAS*, including partial loss of e1A methylation, is more extensive than previously recognized. The clinical and molecular implications of these findings deserve further study.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3449T

DNA Methylation of *ADCYAP1R1* is associated with childhood asthma in Puerto Ricans. *W. Chen[1], N. Boutaoui[1], Y.Y. Han[1], J. Brehm[1], C. Schmitz[1], E. Acosta-Perez[2], M. Cloutier[3], D. Calvert[4], S. Thornton-Thompson[4], D.B. Wakefield[3], M. Alvarez[2], A. Colón-Semidey[2], G. Canino[2], J.C. Celedón[1].* 1) Division of Pulmonary Medicine, Allergy and Immunology, Children's Hospital of Pittsburgh of the University of Pittsburgh Medical Center, Pittsburgh, PA; 2) Behavioral Sciences Research Institute, University of Puerto Rico, San Juan, PR; 3) Department of Pediatrics, University of Connecticut Health Center, Farmington, CT; 4) Connecticut Children's Medical Center, Hartford, CT.

Background: Puerto Ricans share a disproportionate burden of asthma and are often exposed to high levels of stress and violence. Psychosocial stress has been associated with asthma morbidity in children, but little is known about the underlying biological mechanism. Ressler et al recently reported that methylation of the gene encoding the receptor for adenylate-cyclase activating polypeptide 1 (*ADCYAP1R1*) is associated with posttraumatic stress disorder (PTSD) in adults. We hypothesized that methylation of *ADCYAP1R1* is associated with childhood asthma in Puerto Rican (PR) children. **Methods**: Pyrosequencing assays were used to assess methylation of four CpG sites in *ADCYAP1R1* using DNA from white blood cells in 1,001 PR children (including children with [cases, n=550] and without [controls, n=451] asthma) recruited in Hartford (CT) and San Juan (PR). 1 ug of DNA for each sample was bisulfite-converted using the EZ-96 DNA Methylation-Gold kit ™ deep well format (Zymo research, Orange, CA). PCR was performed using diluted bisulfate-converted DNA and *ADCYAP1R1* primers. CpG analysis was performed using PyroMark Q24 software (Qiagen, Valencia, CA). T test and chi-square test were applied for bivariate analyses. Logistic regression was used for the multivariate analysis, which was adjusted for age, gender, household income and (for the combined cohort) study site. **Results**: Among all participants, higher methylation of a CpG site in the promoter of*ADCYAP1R1* is associated with an increased risk of asthma (odds ratio per 5% increase in methylation=1.011, 95% confidence interval=1.002–1.020, P=0.013). Similar results were obtained in a confirmatory analysis excluding two cases with extreme methylation levels (OR= 1.011, 95%CI=1.001–1.020, P=0.022). **Conclusions**: Epigenetic modification of *ADCYAP1R1* may play an important role in the etiology of asthma in PR children. The interaction between psychosocial stress, DNA methylation and childhood asthma merits further examination.

3450T

The Role of DNA Methylation in Cognitive Decline and Alzheimer's Disease. *L.B. Chibnik[1,2,3], B.T. Keenan[1,3], G. Srivastava[1,2,3], M.L. Eaton[3,4], A. Meissner[3,5], J.A. Schneider[6], M. Kellis[3,4], D. Bennett[6], P.L. De Jager[1,2,3].* 1) Program in Translational NeuroPsychiatric Genomics, Institute for the Neurosciences, Brigham & Women's Hospital; 2) Harvard Medical School; 3) Broad Institute of MIT and Harvard; 4) Computer Science and Artificial Intelligence Laboratory, MIT; 5) Department of Stem Cell and Regenerative Biology, Harvard University, Harvard Stem Cell Institute; 6) Rush Alzheimer's Disease Center, Rush University Medical Center.

The DNA methylome captures the transcriptional potential of a cell or tissue: hyper-methylation of a promoter region is typically a mark of a closed chromatin conformation, which prevents transcription. Differential methylation of validated Alzheimer's disease (AD) susceptibility loci could influence their effect. We assess the state of chromatin at the regions of the genome around validated loci to identify CpG sites that correlate with AD disease measures. We utilized data from two longitudinal cohorts, the Religious Order Study and Rush Memory and Aging Project. DNA methylation profiles were generated in samples of dorsolateral prefrontal cortex using Illumina HumanMet450K beadset. We analyzed CpG sites within 25 kb of 10 validated AD susceptibility genes. The outcomes of interest were episodic memory decline (EMD), quantified as residual slope from a linear mixed effects model, a count of neuritic amyloid plaques (NP) and diagnosis of AD (dxAD). We assessed the association between the outcomes and extent of CpG methylation at each site as well as between the AD SNPs and methylation using linear regression. 357 CpGs were tested with a significance threshold for the primary outcome (dxAD) of p<0.0001. A total of 748 subjects, with a mean (SD) age of death of 88.0 (6.6), were included in the analysis. Intriguing results were seen for AD genes BIN1 and EPHA1. Methylation level of the top CpG site at BIN1 (cg04019522), which lies in the gene body, is significantly associated with all three outcomes, dxAD (p=3.6×10-5), EMD (p=0.0006) and NP (p=3.4 x10-5). The best CpG site at EPHA1 (cg26960083) lies just 313 bp from the transcription start site and is associated with dxAD (p=3×10-5), marginally with EMD (p=0.03) but not NP (p= 0.29). In both cases, CpGs are hypermethylated in AD relative to non-demented subjects. Given our modest sample size, the validated SNPs within these genes are not associated with the outcomes in our data (ps>0.05). These observations suggest that, within known AD susceptibility genes, the level of methylation in the brains of older individuals may be related to pathophysiological processes associated with AD and may play a role in influencing gene expression from susceptibility loci.

3451T

Distinct DNA Methylation Signature in Ovarian Cancer Histological Subtypes. *M.S. Cicek[1], D.C. Koestler[2], B.L. Fridley[1], S.M. Armasu[1], K.R. Kalli[3], M.C. Larson[1], R.A. Vierkant[1], B.J. Winterhoff[4], J. Chien[5], J.B. Fan[6], M. Bibikova[6], B. Klotzle[6], G. Konecny[7], V. Shridhar[8], J.M. Cunningham[8], E.L. Goode[1].* 1) Dept. of Health Sciences Research, Mayo Clinic, Rochester, MN; 2) Dept. of Community and Family Medicine, Dartmouth Medical School, Lebanon, NH; 3) Dept. of Medicine, Mayo Clinic, Rochester, MN; 4) Dept. of Obstetrics and Gynecology, Mayo Clinic, Rochester, MN; 5) Dept. of Cancer Biology, University of Kansas, Kansas City, KS; 6) Illumina Inc, San Diego, CA; 7) Dept. of Medicine, UCLA, Los Angeles, CA; 8) Dept. of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN.

Ovarian cancer is a heterogeneous disease; however, it is considered clinically as one disease and treatment is dependent upon tumor stage and grade rather than histologic subtype. Among the most common subtypes of ovarian cancer (high grade serous (70%), low grade serous (5%), endometrioid (10%), clear cell (10%), and mucinous (3%)) there is increasing evidence that various histologic subtypes have different genetic risk factors and outcome. In this study, we tested the hypothesis that the ovarian histologic subtypes have different methylation profiles. We performed genome-wide DNA methylation assays (Infinium HumanMethylation450 BeadChip) on a discovery set of fresh frozen invasive epithelial ovarian cancer tumors from the Mayo Clinic (n=326) to identify methylation signatures that differentiate histological subtypes. In particular, we applied a semi-supervised recursively partitioned mixture modeling (SS-RPMM) strategy to identify methylation signature that can distinguish clear cell subtype from serous, mucinous and endometrioid subtypes. A clear cell methylation profile consisting of an optimal number of 22 CpG loci was identified based on a cross-validation procedure. The identified CpG loci in this methylation profile were correlated with gene expression (Agilent whole human genome 4×44K expression arrays), and CpG methylation was negatively correlated with gene expression at *FGFRL1, KCNH2, ATXN2, NDRG2,* and *SLC16A11*. Replication analysis is underway in two independent set of samples. Preliminary data suggest that we have identified an epithelial ovarian cancer clear cell methylation signature. Future studies will evaluate the relationship between methylation profile and outcome. This clear cell methylation signature will not only help us better understand heterogeneity in ovarian histological subtypes but may also open new treatment targets that might specifically help clear cell ovarian cancer patients.

3452T

Identifying new biomarkers for alcohol use behaviors in a methylome-wide MBD-seq study. *S.L. Clark[1], K.A. Aberg[1], S. Nerella[1], G. Kumar[1], J.L. McClay[1], W. Chen[2], L.Y. Xie[1], A. Hudson[1], G. Gao[2], J. Bukszar[1], C.M. Hultman[3], P.K.E. Magnusson[3], P.F. Sullivan[3,4], E.J.C.G. van den Oord[1], Swedish Schizophrenia Consortium.* 1) Center for Biomarker Research and Personalized Medicine, School of Pharmacy, Virginia Commonwealth University, Richmond, VA, USA; 2) Department of Biostatistics, School of Medicine, Virginia Commonwealth University, Richmond, VA, USA; 3) Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden; 4) Department of Genetics, University of North Carolina at Chapel Hill, NC, USA.

Methylation studies of alcohol phenotypes have historically been restricted to candidate genes in very small sample sizes. More recent studies have taken a methylome-wide approach but have only focused on <450k methylation sites and only in small samples. This study makes a huge leap forward by using next generation sequencing to screen the 30.6 million human methylation sites for association with drinking behaviors. We performed methylome-wide association studies (MWAS) in a sample of ~700 individuals using two phenotypes related to alcohol habits, including if participants ever used alcohol and in what frequency they drink. To investigate all CpGs in the human genome, we enriched for the methylated genomic fraction using a methyl-CpG binding domain (MBD) protein capture system (MethylMinerTM kit from Invitrogen) followed by next generation sequencing (MBD-seq). Each sample had, on average, 67.7 million reads, to yield a total of 4.19 trillion reads for all ~700 individuals. After quality control and collapsing highly correlated CpGs into smaller “blocks”, 4.2 million high quality blocks were left for methylome-wide association testing with the phenotypes of ever using alcohol (lambda of 1.04) and frequency of alcohol use (lambda of 1.01). The MWAS showed a number of highly significant findings with a top p-value of 3.30E-09 and 68 blocks (160 CpGs) with p-values < 9.9E-06. Among the top results were blocks located in, for example, ELOVL2 (p=5.47E-08) and PDE4B (p-value = 9.10E-07) for ever using alcohol and ATF1 (p-value = 1.95E-07) for frequency of alcohol use. PDE4B has previously been linked to alcohol response while ELOVL2 and ATF1 have been associated with comorbid psychiatric disorders. A network analysis was performed to examine if any of the top results cluster into the same network. For example, ATF1 and another top finding, RASGRF1, exist in a known Nerve Growth Factor (NGF) pathway. When this pathway is chronically exposed to alcohol the NGF-stimulated phosphorylation and activation of MAP kinases is enhanced. We are currently replicating the top findings in an independent sample of 560 individuals using targeted pyrosequencing. Once the results are replicated, we will have ascertained the first collection of CpG sites identified using MWAS for alcohol use behaviors.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3453T

Smoking reduces DNA methylation levels at multiple genomic loci. *P. Deloukas[1], L. Tsaprouni[1], T-P. Yang[1], J. Bell[2], K.J. Dick[3], S. Kanoni[1], C. Nelson[3], E. Meduri[1,2], F. Cambien[4], C. Hengstenberg[5], J. Erdmann[6], H. Schunkert[6], A.H. Goodall[3], W.H. Ouwehand[7,7], T. Spector[2], N. Samani[3].* 1) Human Genetics, Wellcome Trust Sanger Inst, Cambridge, United Kingdom; 2) Department of Twin Research and Genetic Epidemiology, King's College London, UK; 3) Department of Cardiovascular Sciences, University of Leicester & National Institute for Health Research Leicester Cardiovascular Biomedical Research Unit, Glenfield Hospital, Leicester, UK; 4) INSERM UMRS 937, Pierre and Marie Curie University and Medical School, Paris, France; 5) Klinik und Poliklinikfür Innere Medizin II, Regensburg, Germany; 6) Universität zu Lübeck, Medizinische Klinik II, & Deutsches Zentrum für Herz-Kreislauf-Forschung, Lübeck, Germany; 7) Department of Haematology, University of Cambridge and National Health Service (NHS) Blood and Transplant, Cambridge, UK.

Smoking is a major risk factor in most common diseases impacting mortality. Genome wide association studies have clearly linked nicotine dependence and smoking behaviour to increased risk of cardiovascular, pulmonary and malignant diseases whilst recent evidence show that global DNA methylation is influenced by smoking behaviour. We assessed methylation patterns in peripheral-blood DNA in 464 individuals (Cardiogenics study) of which 285 are current or ex-smokers with the use of the Illumina Human Methylation 450K BeadChip. A linear model analysis confirmed an effect of smoking on methylation in two previously reported genes: Coagulation factor 2 (F2RL3) and G-protein coupled receptor 15 (GPR15). F2RL3 was significantly associated with methylation in smokers (p=2.8×10-17 at FDR=1.65×10-12) which replicates the association reported earlier this year. GPR15 was also associated with smoking history ($p=2.51\times10^{-18}$ at $FDR=1.72\times10^{-13}$). We obtained 62 signals at a nominal p value 10^{-6} corresponding to 13 loci. Interestingly, one of the top signals is in a region of chromosome 2 (233283397–233285959) located close to the neuronal acetylcholine receptor subunit D (CHRND) gene. We then undertook replication in an independent sample of 356 Twins from the TwinsUK registry which comprises 145 current or ex-smokers. After applying Bonferroni correction 19 probes representing 12 loci which included F2RL3, GPR15 and the chromosome 2 region, replicated at $p < 8.06\times10^{-4}$. Levels of methylation were reduced in smokers in all replicating loci; for example all seven probes covering the chromosome 2 region ($p=6.48\times10^{-10}$ - 1.81×10^{-22}). Interestingly, this trend was reversed when we examined people that had stopped smoking (data from TwinsUK study only). Finally, we assessed the intersection between methylated QTLs (metQTLs; $p < 10^{-3}$) versus the top 62 probes from the linear model analysis described above. Both F2RL3 and the chr2 (233283397–233285959) region had a metQTL in blood associated with methylation. However, GPR15 did not have any metQTL at genome wide significance.

3454T

DNA hypomethylation of *INS-IGF2* and *BACH2* in affected individuals within discordant monozygotic twins suggests a role for epigenetic influences in type 1 diabetes susceptibility. *E. Elboudwarej[1], H. Quach[1], P. Ramsay[1], J. Lane[2], E. Sinclair[4], L. Criswell[4], P. Fain[3], J. Jeffrey[3], G. Eisenbarth[3], L. Barcellos[1], J. Noble[2].* 1) Genetic Epidemiology and Genomics Laboratory, Division of Epidemiology, School of Public Health, UC Berkeley, Berkeley, CA; 2) Children's Hospital Oakland Research Institute, Oakland, CA; 3) University of Colorado Health Sciences Center, Barbara Davis Center for Childhood Diabetes, Denver, CO; 4) University of California, San Francisco, Department of Medicine, San Francisco, CA.

Type 1 diabetes (T1D) has a substantial genetic risk component, with the strongest risk attributable to Human Leukocyte Antigen (HLA) genes in the Major Histocompatibility Complex (MHC). Traditional genetics cannot completely explain susceptibility, as evidenced by varying levels of disease discordance among monozygotic (MZ) twin pairs. In this matched case-control study, seven MZ twin pairs, T1D discordant for 8 to 20 years, were examined for differences in DNA methylation (DNAm) at more than 480,000 CG dinucleotides (CpGs) across the genome using DNA from peripheral blood samples. An additional set of five HLA matched, T1D discordant siblings, was analyzed as a replication dataset. Initially, measures of methylation were investigated as averages across all sites within three pre-selected regions: Whole-genome (global), T1D-associated genes (non-MHC), and MHC region genes. Among twin pairs, both T1D-associated and MHC region methylation were significantly lower in affected twins ($p < 0.05$). For site-specific analyses, CpGs of interest were analyzed individually, or were averaged across sites within a gene when appropriate. A total of 14 sites in 12 MHC region genes, and 7 sites in 5 T1D-associated genes (*INS-IGF2*, *BACH2*, *IL2*, *IL2RA*, and *TNFAIP3*) demonstrated significant differences in methylation between the affected and unaffected twin of each pair ($p < 0.05$). Differentially methylated genes were consistently hypomethylated in the affected twin when compared to the unaffected twin. To address the question of whether or not differences in the candidate loci might be attributable to specific cell populations, methylation was examined in the following cell subsets from a single unrelated individual: CD14+ monocytes, CD19+ B lymphocytes, and both CD4+ naïve and memory T lymphocytes derived from peripheral blood. While methylation of sites within *BACH2* and *IL2RA* showed high variability across the four cell types, *INS-IGF2*, *IL2*, and *TNFAIP3* profiles were identical. Results suggest that DNA methylation differences in T1D-associated genes distinguish affected from unaffected individuals in MZ twin pairs but not discordant siblings. Epigenetic influences may partially explain some of the missing heritability for T1D risk.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3455T

Migration and DNA methylation: a comparison of methylation patterns in type 2 diabetes susceptibility genes between Indians and Europeans. *H.R. Elliott[1], G.K. Walia[2], A. Duggirala[3], A. Groom[1], S.U. Reddy[3], G.R. Chandak[3], V. Gupta[2], M. Laakso[4], J.M. Dekker[5], M. Walker[1], S. Ebrahim[2,6], G. Davey Smith[7], C.L. Relton[1], The RISC consortium.* 1) Newcastle University, Newcastle-upon-Tyne, Tyne and Wear, United Kingdom; 2) South Asia Network for Chronic Disease, Public Health Foundation of India, New Delhi, India; 3) Centre for Cellular and Molecular Biology, Council of Scientific and Industrial Research, Hyderabad, India; 4) University of Eastern Finland, Finland; 5) VU University Medical Centre of Amsterdam, Netherlands; 6) Department of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, United Kingdom; 7) MRC Centre for Causal Analyses in Translational Epidemiology, Department of Social and Community Medicine, University of Bristol, Bristol, United Kingdom.

Introduction Type 2 diabetes is a global problem that is increasingly prevalent in low and middle income countries including India, partly attributed to increased urbanisation. Genotype clearly plays a role in type 2 diabetes susceptibility. However, the role of DNA methylation and its interaction with genotype and metabolic measures is poorly understood. This study aimed to establish whether methylation patterns of type 2 diabetes implicated genes differ between distinct Indian and European populations and/or change following rural to urban migration in India.**Methods** Quantitative DNA methylation analysis in Indians (Urban and Rural dwelling siblings recruited in Lucknow and Hyderabad) and Europeans using Sequenom® EpiTYPER™ technology was undertaken in three genes: ADCY5, FTO and KCNJ11. Metabolic measures and genotype data were also analysed to assess their relationships with DNA methylation.**Results** Consistent differences in DNA methylation patterns were observed between Indian and European populations. Mann-Whitney U tests showed DNA methylation was lower in the European group compared to the Urban Indian group for all three loci tested: ADCY5 median difference 6.3% (z= 17.72, p= 2.831e-70); FTO median difference 2.7% (z= 15.44, p= 9.495e-54); KCNJ11 median difference 1.5% (z= 14.60, p= 2.787e-48). Methylation differences were also observed between populations within India. In Europeans, SNPs were analysed for an additive effect using non parametric trend tests. An increase in HDL of 0.07 mmol/L was observed when the minor allele of ADCY5 rs17295401 was present (trend test: z= 2.66, p= 0.008). HDL showed a 0.16 mmol/L average increase per minor allele of FTO rs16952479 (trend: z= 2.62, p= 0.009). Fasting insulin was on average 0.45 mU/L higher (trend: z= 2.52, p= 0.012) and BMI was on average 0.53 units higher (z= 2.82, p= 0.005) per minor allele of FTO rs9939609. HDL showed a 0.16 mmol/L average decrease per minor allele of KCNJ11 rs1800467 (trend: z= -3.33, p= 0.001). However, these observations were not linked to local variation in DNA methylation levels. No differences in methylation patterns were observed in urban-dwelling migrants compared to their non-migrant rural-dwelling siblings. **Conclusions** Analysis of DNA methylation at three type 2 diabetes susceptibility loci highlighted geographical and ethnic differences in methylation patterns. These differences may be attributed to genetic and/or region-specific environmental factors.

3456T

DNA methylation alterations associated with autism spectrum disorders. *D. Grafodatskaya[1], R. Rajendram[1], Y. Lou[1], D.T. Butcher[1], L. Senman[2], C. Windpassinger[3], W. Roberts[2], S.W. Scherer[4], R. Weksberg[5].* 1) Gen & Genome Biol, Hosp Sick Children, Toronto, ON, Canada; 2) Autism Research Unit, Hosp Sick Children, Toronto, ON, Canada; 3) Institute of Human Genetics, Medical University of Graz, Graz, Austria; 4) The Centre for Applied Genomics, Hospital for Sick Children, Toronto, Canada; 5) Institute of Medical Sciences, University of Toronto, Toronto, Canada.

Autism spectrum disorders (ASD) are a group of childhood neurodevelopmental disorders typically diagnosed at 2–3 years of age and characterized by three areas of behavioral impairment: social interactions, communication and repetitive behaviors. Since known genetic factors account for only subset of ASD cases and environmental risk factors have been implicated, it is likely that epigenetics plays a significant role in ASD etiology. Epigenetics refers to heritable changes in gene expression occurring without altered DNA sequence. DNA methylation (DNAm) is the best studied form of epigenetic regulation. The objective of this study was to identify ASD-associated DNAm alterations in the blood of such individuals. Genome-wide DNAm was assessed in 12 blood samples from ASD cases and 12 controls at ~27,000 CpG sites using the Illumina 27K platform. To identify DNAm variants associated with ASD, we selected CpG sites with DNAm changes of at least 10%, compared to controls. Using this approach we identified 7 CpGs with loss and 5 CpGs with gain in ASD vs controls. Further, in order to assess the relevance of DNA methylation changes identified in blood to brain, we compared DNAm levels at the 13 CpG sites referred to above to 150 brain samples of neurologically normal individuals run on the same microarray platform (Gibbs et al. 2010). To narrow the list of candidates we selected CpGs with: 1) similar levels of DNAm in blood and brain in controls, 2) no outliers in the brain similar to the outliers levels in the blood of ASD cases. Two rare variants with gain of DNAm at CpG sites within promoters of arylsulfatase A (ARSA) and phosphatidylcholine transfer protein (PCTP) genes survived this comparison. DNAm at these two loci were validated and replicated in extended cohorts using targeted DNAm analysis. Our data support a role for epigenetic variants in ASD etiology and demonstrate the feasibility of identifying such epigenetic variants in clinically accessible tissues such as blood.

3457T

A population-based method for identification of differential methylation regions among samples. *C.L. Hsiao[1], C.S.J. Fann[1], C.J. Chang[2].* 1) IBMS, Academia Sinica, Taipei, Taiwan; 2) Graduate Institute of Clinical Medical Sciences, Chang-Gung University, Taiwan.

The epigenetic variation, specifically variation in DNA methylation, has been proven to be involved with gene regulation and may make a significant contribution to the missing heritability. Advances of genomic technologies have placed us in a position to initiate large-scale studies of DNA methylation. The epigenome-wide association studies present novel opportunities for understanding the role of DNA methylation in pathogenesis but with new challenges that are not encountered previously. The differential methylation regions (DMRs) represent genomic regions with different methylation statuses among multiple samples. Successfully recognizing DMRs without given any phenotypic information could help to understand variations of methylation status changes in a given population, either among individuals or between different conditions, e.g. different tissue types or tumor-normal samples. Previously many supervised methods have been studied, and we propose a novel unsupervised approach to directly analyze methylation intensity data across a large set of samples to identify DMRs, with no need for prior single-sample analysis. This approach is a region-based method that takes into account the correlations of methylation levels between nearby probes which may introduce a high rate of false positive for point-wise method in simulation studies. In a real data study, 25 methylation data profiling by genome tiling array (NimbleGen Human 2.1M array) are employed, including 5 samples each of normal liver, frontal cortex, spleen and colon, and 5 samples of colon tumor. More than 5,000 DMRs with at least 15% of samples having different methylation patterns in comparison with the rest of the samples have been identified. The changes of methylation status on these DMRs are highly associated with the specific tissue types and tumor/normal colon samples. The results show that our newly proposed method is useful in identification of differential methylation regions among samples.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3458T

Genetics and epigenetics of PON1 in Mexican-American children. *K. Huen, K. Harley, P. Yousefi, A. Bradman, L. Barcellos, B. Eskenazi, N. Holland*. Center for Environment al Research and Children's Health (CERCH), School of Public Health, University of California, Berkeley, CA.

Children's susceptibility to environmental toxicants may depend on multiple factors including early life exposures, genetics, and gene-environment interactions. CHAMACOS is a longitudinal birth cohort study of low-income Mexican-American farmworker families in California. We followed 601 pregnant women and their children and found that *in utero* exposures to pesticides were associated with abnormal birth and neurodevelopmental outcomes (Eskenazi et al,2004, 2007, 2010; Young et al, 2005; Bouchard et al, 2011). We also demonstrated that these health effects depend on phenotype and genotype of paraoxonase (PON1), a multifunctional enzyme involved in detoxification of organophosphate pesticides and oxidative stress (Holland et al, 2006; Huen et al, 2009;2010 Eskenazi et al, 2010; Harley et al, 2011). Previously, we sequenced the *PON1* gene in our Mexican-American subjects and identified several new functional polymorphisms (Huen et al, 2011). These polymorphisms as well as haplotypes containing multiple SNPs only explained an additional 2–4% of PON1 activity compared to well-known SNPs$PON1_{192}$ and $PON1_{-108}$ considered individually. Although coding SNP $PON1_{192}$ accounts for >70% of variation of substrate-specific paraoxonase activity, only a small portion of the variation (~25%) of PON1 enzyme quantity is explained by promoter SNP $PON1_{-108}$. We sought to determine whether epigenetic mechanisms such as DNA methylation in the *PON1* gene (assessed by Illumina 450K Methylation BeadChip) may also significantly influence PON1 quantity (measured by arylesterase activity) in CHAMACOS children at birth and 9 years of age (n=104). We interrogated 18 CpG sites in the *PON1* gene located in shores, shelves, and CpG islands. DNA methylation, particularly in several south shore and CpG island sites, was strongly associated with PON1 quantity ($p<0.005$) in 9 year olds but not in newborns. Furthermore, we observed significant allele-specific differences in methylation with the promoter SNP $PON1_{-108}$. Epigenetic changes were associated with PON1 quantity yet it remains unknown whether factors like environmental exposure influence these alterations. In future studies we will examine the relationship between early life exposures and DNA methylation in the *PON1* gene to determine whether epigenetics mediate the effects of exposure on PON1 variation.

3459T

Comprehensive DNA Methylation Profiling of 3.7 million CpGs with the SureSelect Target Enrichment System. *K. Jeong[1], M. Corioni[1], F. Useche[1], A. Shafer[3], E. Haugen[3], B. Novak[1], S. Happe[2], D. Roberts[1], J. Stamatoyannopoulos[3], E. LeProust[1]*. 1) Genomics Division, Agilent Technologies, Santa Clara, CA; 2) Genomics Division, Agilent Technologies, Cedar Creek, TX; 3) Dept. of Medicine, Univ. of Washington, Seattle, WA.

DNA cytosine methylation is a critical epigenetic modification involved in human diseases such as cancer and imprinting disorders. Various cellular processes including gene regulation, embryonic development, X chromosome inactivation, and chromatin remodeling are strongly associated with DNA methylation changes. Next-generation sequencing combined with sodium bisulfite treatment allows identification of methylation changes at single base resolution. However, whole genome bisulfite sequencing is prohibitively expensive. Additionally, many of the regions in whole genome bisulfite sequencing are in repetitive regions, and provide very little useful information. In many cases, the researcher is only interested in profiling a subset of biologically relevant regions. To address these needs, we have developed SureSelect Methyl-Seq, which combines Agilent's SureSelect Target Enrichment platform with bisulfite sequencing to detect methylation changes. Our Methyl-Seq design targets human genomic CpG sites within CpG islands/shores, promoters, known differentially methylated regions (DMRs) and previously determined regulatory regions. This comprehensive design covers 84Mb making it well suited to study cancer-, tissue-, and stem cell-specific DMRs. SureSelect Methy-Seq is also optimized for different sizes of custom designs. Here we describe the SureSelect Methyl-Seq workflow and demonstrate efficient target enrichment and precise methylation level detection. Further, we show high concordance with whole-genome bisulfite sequencing of known model systems. We also describe the detection of tissue and cancer specific DMRs in various samples including formalin fixed paraffin embedded tissues.

3460T

The development of a method for in-solution hybrid capture of bisulfite-converted DNA for targeted bisulfite sequencing of 175 ADME genes. *M. Kals[1], M. Ivanov[2], M. Kacevska[2], A. Metspalu[1], M. Ingelman-Sundberg[2], L. Milani[1]*. 1) Estonian Genome Center, University of Tartu, Tartu, Estonia; 2) Department of Physiology and Pharmacology, Karolinska Institutet, Stockholm, Sweden.

DNA methylation is one of the most important epigenetic alterations involved in the control of gene expression. Bisulfite sequencing (BS-seq) of genomic DNA is currently the only method to study DNA methylation patterns at the single-nucleotide resolution. Hence, next-generation sequencing (NGS) of bisulfite-converted DNA is the method of choice to investigate DNA methylation profiles at the genome-wide scale. Nevertheless, sequencing of hundreds of human methylomes is still a very expensive task. At the same time, researchers often would like to investigate not the whole methylome but only a subset of genes of interest. Thus, the use of DNA target enrichment techniques together with bisulfite conversion of DNA could allow researchers to restrict the analyzed region of human epigenome. We successfully developed a novel protocol for the hybrid capture of bisulfite converted gDNA using custom Agilent SureSelect Target Enrichment system. In the pilot experiment, we initiated the protocol with 4 gDNA samples isolated from adult human livers and prepared 4 bisulfite-converted NGS libraries, which were enriched for 4 Mb genomic region of interest covering non-repetitive intervals of coding and regulatory regions of 175 ADME genes which are involved in the metabolism and distribution of drugs. The proposed method was validated by pyrosequencing and the Illumina HumanMethylation450 assay, and both techniques were highly consistent with the NGS BS-seq results. Sequencing of these target enriched bisulfite libraries on Illumina HiSeq2000 allowed us to reliably quantify methylation levels of more than 60% of the CpG sites in the selected genes. The number of common CpG sites for all 4 samples was more than 45,000; the median coverage varied between 36–77x and the mean CpG methylation level was 57–60%. Data analysis revealed that specific CpG sites are variably methylated among individuals and this variability is found in almost all ADME genes assessed. Such variable DNA methylation patterns provide valuable information regarding ADME gene regulation and aid in understanding inter-individual differences in drug metabolism and transport.

3461T

Association of Fetal-Derived Hypermethylated RASSF1A Concentrations in Placenta-Mediated Pregnancy Complications. *S. Kim[1], M. Kim[1], S. Park[1], D. Lee[1], J. Lim[1], J. Han[2], H. Ahn[2], J. Chung[2], M. Ryu[1, 2]*. 1) Laboratory of Medical Genetics, Medical Research Institute, Cheil General Hospital and Women's Healthcare Center, Seoul, Korea; 2) Department of Obstetrics and Gynecology, Cheil General Hospital and Women's Healthcare Center, KwanDong University college of Medicine, Seoul, Korea.

Objective: The purpose of this study was to evaluate whether the fetal-derived hypermethylated RASSF1A concentration in maternal plasma is altered during pregnancy in pregnancies complicated with placental dysfunction manifested by intrauterine growth restriction (IUGR), preeclampsia (PE), or placental previa (PP) and whether any possible alternation can be detected in susceptible subjects before onset of the clinical disease. Methods: We performed real-time quantitative PCR to quantify RASSF1A concentrations before and after methylation-sensitive restriction digestion in maternal plasma at 7 to 40 gestational weeks of normal pregnancies (n = 161), IUGR (n = 43), PE (n = 22), PP (n = 10), and non-pregnant women (n = 20). Results: A positive correlation was observed between fetal-derived hypermethylated RASSF1A concentration and gestational age for all patient groups ($r = .624$; $P < .001$ for IUGR, $r = .381$; $P = .042$ for PE, $r = .606$; $P = .010$ for PP, $r = .560$; $P < .001$ for controls). The median concentration of hypermethylated RASSF1A was increased as early as at 10 to 13 weeks of gestation. Hypermethylated RASSF1A concentration in the second trimester (15 to 27 weeks) was significantly higher in patients who subsequently developed IUGR ($P = .002$), PE ($P < .001$) or PP ($P < .001$) than in controls. Conclusion: This study first demonstrated an increased concentration of fetal-derived hypermethylated RASSF1A sequences according to advancing gestation and before the onset of the clinical manifestation of pregnancy complications secondary to placental dysfunction such as IUGR, PE and PP. Hypermethylated RASSF1A in maternal plasma may be considered as a potential biomarker to detect placental-medicated pregnancy complications, regardless of fetal gender and polymorphism.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3462T

Epigenome Wide Association Study of Rheumatoid Arthritis. *I. Koh[1], J. Kim[1], J. Yoon[1], B. Kim[1], I. Park[1], K. Kim[1], Y. Kim[1], R. Huh[1], B. Han[1], S. Bae[2], J. Bae[1], J. Lee[1].* 1) KNIH, Osong, Chungcheongbukdo, South Korea; 2) Hanyang University, Seoul, Korea.

In addition to genetic factors, various environmental factors have been associated with the pathogenesis of autoimmune disease. One of major environmental factors related is epigenetic modification. To identify causative epigenetic aberration in autoimmune disease, we performed DNA methylation profiling of 76 Rheumatoid Arthritis (RA) patients against 14 healthy control by methyl binding protein affinity purification-Sequencing (MBD-Seq) and Illumina 450k DNA methylation beadarray in blood and synovium fluid CD4+ T cell. Epigenome profiling reveals that T cell development related pathway was most significantly enriched in differentially expressed gene set. Next, we identified Differential Methylation Site (DMS, p-value less than 10^{-7}). Top3 DMS include development related transcription factor, T cell development related transcription factor and angiotensin binding protein playing a role in autoimmune disease. In addition to DMS identification, we compared marker identified from Genome Wide Association Study (GWAS) data with epigenetic modification profile. we tested 7 top significant SNP loci confirmed from 3 independent GWAS data, which only revealed that the strongest marker HLA-DRB1(Odds Ratio=2.51, p-value=$4.9*10^{-23}$) showed significant difference in DNA methylation (p-value = 0.069), whereas other 6 marker didn't showed any significant methylation change. Thus, HLA-DRB1 locus was single common target for both genome and epigenome modification. Further study on common target of genome-epigenome will be discussed.

3463T

DNA methylation analysis of iPS cells using whole-genome bisulphite sequencing. *D.S. Lee[1,2,5], J.Y. Shin[1,4], Y.S. Ju[1,2,3,5], J.I. Kim[1,2,4,5], P. Tonge[6], M. Puri[6], A. Nagy[6,7], J.S. Seo[1,2,3,5].* 1) GMI, Seoul national university, Seoul, Seoul, South Korea; 2) Department of Biomedical Sciences, Seoul National University Graduate School, Seoul 110-799, Korea; 3) Macrogen Inc., Seoul 153-023, Korea; 4) Psoma Therapeutics Inc., Seoul 110-799, Korea; 5) Department of Biochemistry and Molecular Biology, Seoul National University College of Medicine, Seoul 110-799, Korea; 6) Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, Ontario M5G 1×5, Canada; 7) Department of Molecular Genetics, University of Toronto, Toronto M5S 1A8, Canada.

iPS cells have several advantages for stem cell research and regenerative medicine including a potentially limitless supply of patient-matched cells for cell-based therapies without the requirement for human embryonic donor tissue. iPS cells may be used for therapeutics, in vitro modelling of human complex diseases, and drug development. Somatic cell reprogramming involves epigenetic remodelling of chromatin architecture including methylation, conferring iPS cells with characteristics similar to embryonic stem (ES) cells. However, it remains unknown whether the re-establishment of an ES-cell-like DNA methylation pattern occurs throughout the genome, and how the methylation pattern changes during the reprogramming process. To address these questions, we utilized a secondary inducible reprogramming system developed in mouse using piggyBac-mediated transposition of the four reprogramming factors. Here we report whole-genome DNA methylation profiles at single-base resolution for mouse ES cells, secondary mouse embryonic fibroblast (MEF), secondary factor independent iPS cells induced from these MEFs, and cells undergoing reprogramming at Day 8, using whole genome bisulphite sequencing. Hiseq 2000 was used for the whole genome sequencing, and Bismark was used as an alignment tool. We defined differentially methylated regions (DMRs) and partially methylated domains(PMDs) in CpG, CHG, and CHH context. iPS cells show significant reprogramming variability, including somatic memory and aberrant reprogramming of DNA methylation.

3464T

Advanced genome-wide methylome profiling and analysis of leukemia reveals novel epigenetic dynamics. *S. Li, A. Akalin, F. Garrett-Bakelman, A. Melnick, C. Mason.* Weill Cornell Medical College, New York, NY.

Understanding the molecular mechanisms that lead cancer is one of the most important aspects of cancer research. Previous studies have shown that aberrant cytosine methylation patterning is a hallmark of acute lymphoblastic leukemia (AML). DNA cytosine methylation is a critical epigenetic modification that regulates gene expression and mediates cellular differentiation states. Current high-throughput methods for examination of the epigenome, such as reduced representation bisulfite sequencing (RRBS), ameliorate the cost of sequencing after bisulfite treatment, but limit the number of regions that can be examined. Here, we report biochemical and bioinformatics improvements in the RRBS method, creating a new technique called enhanced RRBS (ERRBS) and a R package methylKit which can quickly analyze genome-wide cytosine methylation profiles from high-throughput bisulfite sequencing experiments. We applied the protocol and performed differential methylation analysis on AML samples. Our method shows an improved conversation efficiency of the non-methylated cytosines (>99.5%). We found an increased number (11–61%) of CpG sites that can be detected and quantified in all areas: CpG islands, promoters, CpG shores, introns, and exons. We applied methylKit to stratify the AML patients based on the methylation dynamics. Moreover, we effectively determined the differentially methylated regions (DMRs) and CpG sites (DMCs) using a hidden markov model (HMM) that revealed new regions of the genome that are altered in AML.

3465T

Differential Methylation of Colorectal Adenomas of the Right Side. *J. Li[1], D. Koestler[1], J. Baron[1], A. Dietrich[1], G. Tsongalis[2], L. Butterly[2], M. Goodrich[2], C. Marsit[1], J. Moore[1], A. Andrew[1], A. Srivastava[3].* 1) Geisel School of Medicine, Dartmouth College, Lebanon, NH; 2) Dartmouth-Hitchcock Medical Center, Lebanon, NH; 3) Brigham and Women's Hospital, Boston, MA.

Colon polyps are growths on the surface of the colon, which are generally asymptomatic, but have the potential of progressing to cancer. Recent work has demonstrated that colonoscopy reduces mortality for cancers on the left, but not on the right side of the colon, suggesting that there are important biological differences between the right- and left-sided precursors of colon cancer. We hypothesized that the phenotypic differences between right- and left-colon polyps may be due to their underlying DNA methylation signatures. To evaluate side-specific differences in DNA methylation status, we obtained primary polyp tissue (n=65 left-sided, n=70 right-sided polyps), removed at the baseline screening colonoscopy from consenting New Hampshire Colonoscopy Registry (NHCR) participants who were diagnosed with a conventional adenoma. DNA extracted from the primary tissue was then subjected to epigenome-wide methylation assessment using the Illumina HumanMethylation450 BeadChip, which profiles the methylation status for >485,000 CpG loci. A principal component analysis (PCA) applied to the methylation data revealed that the first principle component, which explains the largest amount of variability in methylation across the array, was significantly associated with colon polyp side (p=0.007). Furthermore, examining methylation differences across the array revealed 57,795 (12.2%) loci that were associated with left/right colon polyp side (p<=0.05) and 168 loci below the Bonferroni threshold (p<=1.06e-7). Among significant loci, a greater than expected proportion was associated with genes regulating cell growth, mitotic cycle, oncogenic transformation, or had prior evidence of modification in cancer. Unexpectedly, the percentage of the probes that located on the CpG islands among the entire array was 30.8%, while percentage of CpG loci among the significant probes associated with side was only 18.2%. Our results suggest both gene-specific and genome wide differences in DNA methylation between left and right colon-adenomas. Molecular characterization of these lesions will help elucidate the basis of the side-specific differences in polyps and help to identify potential molecular markers of lesion subtypes that may require different screening intervals.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3466T

Methylation quantitative trait loci and epigenetic networks underlying asthma-related traits. *L. Liang[1], S.A.G. Willis-Owen[2], K.C.C. Wong[2], A. Binia[2], G. Davies[3], J.M. Hopkin[3], I. Yang[4], T.M. Pastinen[5], D.A. Schwartz[4], G. Abecasis[6], G.M. Lathrop[7], M.F. Moffatt[2], W.O.C.M. Cookson[2].* 1) Departments of Epidemiology and Biostatistics, Harvard School of Public Health, Boston, MA 02115; 2) National Heart and Lung Institute, Imperial College, London SW3 6LY, UK; 3) Institute of Life Science, School of Medicine, Swansea University, UK; 4) University of Colorado School of Medicine and National Jewish Health, Denver, CO; 5) Departments of Human and Medical Genetics, McGill University and Génome Québec Innovation Centre, Montréal, Canada; 6) Center for Statistical Genetics, Dept. of Biostatistics, SPH II, Ann Arbor, MI; 7) CEA / Centre National de Genotypage, 91057 Evry, France.

Epigenetic variation in the methylation of DNA at CpG islands (CGI) is related to the regulation of transcription. The extent of epigenetic variation in the human genome and its relevance to common diseases are not known. We studied methylation status in asthmatic nuclear families genome-wide at 27,000 CGI in DNA from peripheral blood leukocytes (PBL). We show here that major components of variation in methylation quantitative trait loci (meQTL) included age, sex and the shared family environment. We observed true genetic influences (meSNPs) on methylation status at 10% of loci. In order to establish the utility of genome-wide meQTL association studies (GMAS) to common diseases we show mapping of meQTL with genome wide significance to multiple phenotypes. We found and replicated meQTL associations to the total serum IgE concentration ($P<10^{-11}$) the top four of which accounted for 21.5% of the IgE variation, compared to 1–2% seen in GWAS studies of the same trait. Cigarette smoking showed replicated associations to loci influencing coagulation. Genome function is mediated through interactive networks of genes and regulatory elements, and we further show here the presence of strongly co-ordinated regulation of meQTL in the form of 30 scale-free meQTL correlation networks (meQTN). Enrichment analysis identified meQTN modules that could be attributed to peripheral blood neutrophils, lymphocytes, monocytes and eosinophils ($P<10^{-81}$ to 10^{-350}). These and other modules were enriched by loci associated to asthma and the total serum IgE ($P<10^{-16}$ to 10^{-104}). Hubs in these networks define potential therapeutic targets for manipulating cell-specific immune function in asthma and other common inflammatory diseases.

3467T

Activation of silenced tumor suppressor genes in prostate cancer cells by a novel energy restriction-mimetic agent. *H. Lin[1,2,4], Y. Kuo[2], Y. Weng[3], I. Lai[2], T. Huang[3], S. Lin[1], D. Niu[4,5], C. Chen[2].* 1) Department of Pediatrics, Mackay Memorial Hospital, Taipei, Taiwan and Mackay Medicine, Nursing and Management College, Taipei, Taiwan; 2) Division of Medicinal Chemistry, College of Pharmacy, The Ohio State University, Columbus, OH 43210, U.S.A; 3) Human Cancer Genetics Program, The Ohio State University, Columbus, OH 43210, U.S.A; 4) Institute of Clinical Medicine, National Yang-Ming University, Taipei, Taiwan; 5) Department of Pediatrics, Taipei Veterans General Hospital, Taipei, Taiwan.

BACKGROUND. Targeting tumor metabolism by energy restriction-mimetic agents (ERMAs) has emerged as a strategy for cancer therapy/prevention. Evidence suggests a mechanistic link between ERMA-mediated antitumor effects and epigenetic gene regulation. METHODS. Microarray analysis showed that a novel thiazolidinedione-derived ERMA, CG-12, and glucose deprivation could suppress DNA methyltransferase (DNMT)1 expression and reactivate DNA methylation-silenced tumor suppressor genes in LNCaP prostate cancer cells. Thus, we investigated the effects of a potent CG-12 derivative, CG-5, vis-à-vis 2-deoxyglucose, glucose deprivation and/or 5-aza-deoxycytidine, on DNMT isoform expression (Western blotting, RT-PCR), DNMT1 transcriptional activation (luciferase reporter assay), and expression of genes frequently hypermethylated in prostate cancer (quantitative real-time PCR). Promoter methylation in these genes was assessed by pyrosequencing analysis. SiRNA-mediated knockdown and ectopic expression of DNMT1 were used to validate DNMT1 as a target of CG-5. RESULTS. CG-5 and glucose deprivation upregulated the expression of DNA methylation-silenced tumor suppressor genes, including GADD45a, GADD45b, IGFBP3, LAMB3, BASP1, GPX3, and GSTP1, but also downregulated methylated tumor/invasion-promoting genes, including CD44, S100A4, and TACSTD2. In contrast, 5-aza-deoxycytidine induced global reactivation of these genes. CG-5 mediated these epigenetic effects through transcriptional repression of DNMT1, which was associated with reduced expression of Sp1 and E2F1. SiRNA-mediated knockdown and ectopic expression of DNMT1 corroborated DNMT1's role in CG-5-mediated modulation of gene expression. Pyrosequencing revealed differential effects of CG-5 versus 5-aza-deoxycytidine on promoter methylation in the genes examined. CONCLUSIONS. These findings reveal a previously uncharacterized epigenetic effect of ERMAs on DNA methylation-silenced tumor suppressor genes, which may foster novel strategies for prostate cancer therapy.

3468T

Patterns of SNP-based genome-wide heritability of methylation in four brain regions. *J. Listgarten, G. Quon, C. Lippert, D. Heckerman.* Microsoft Research, Los Angeles, CA., 90024.

We used a mixed model to investigate the extent to which genetic variation in DNA sequence explains variation in CpG dinucleotide methylation in data from 150 individuals for four brain regions. In particular, our goals were to investigate what role cis-DNA sequence plays in influencing methylation, what an optimum definition of cis (i.e., locality) is in this context, and whether CpG dinucleotides with heritable methylation were more likely to be associated with particular classes of genes. We found that 5–7% of CpG sites assayed were heritable, with a median narrow-sense heritability of 21% (and a mean of 3.5% over all sites) when using an optimal cis window of 50kb. We also found good concordance across brain regions. Our percentage of significant sites is similar to work by Bell et al., who found that 6.3% of CpG sites tested had significant QTLs, although they found a mean genome-wide heritability of 18%, higher than ours, unsurprisingly given that they used a twin-based correlation estimate. Finally, we show that the set of genes potentially targeted by these methylation loci are enriched in human leukocyte antigen (HLA) genes, which are important in a variety of developmental and immune processes. Our estimates of heritability are conservative, and we suspect that the number of heritable loci will increase in the near future as the methylome is assayed across a broader range of tissue and cell types and the density of the tested loci is increased.

3469T

Statistical Models to Predict DNA Methylation Level across Different Tissues in Human. *B. Ma[1], E. Wilker[2], S.A.G. Willis-Owen[3], K.C.C. Wong[3], A. Baccarelli[1,2], J. Schwartz[1,2], W.O.C.M. Cookson[3], K. Khabbaz[4], M.A. Mittleman[1], M.F. Moffatt[3], L. Liang[1,5].* 1) Department of Epidemiology, Harvard School of Public Health, Boston, MA; 2) Department of Environmental Health, Harvard School of Public Health, Boston, MA; 3) National Heart and Lung Institute, Imperial College, London SW3 6LY, UK; 4) Division of Cardiothoracic Surgery, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA; 5) Department of Biostatistics, Harvard School of Public Health, Boston, MA.

Variation in DNA methylation is important to cell differentiation and is known to be associated with aging, cancers and complex diseases. Although studies have shown that methylation patterns are highly conserved across tissues, there are substantial between-tissue variations that characterize tissue-specific DNA methylation patterns. In this investigation, we found that for individual loci the methylation difference between tissues is highly consistent and reliably reproducible across multiple individuals. Following this finding, we developed a statistical model for each CpG site to predict methylation level in one tissue based on the methylation level in another tissue. Using data from two ongoing studies: (1) In a study for childhood asthma, 39 individuals with methylation measured in both peripheral blood leukocytes (PBL) and EBV transformed lymphoblastoid cell lines (LCL) using the Illumina Infinium HumanMethylation27 array that covers 27,000 CpGs, and (2) In a study for atrial fibrillation, 14 individuals with methylation measures in 3 tissues (atrium, internal mammary artery and peripheral leukocytes) using Illumina Infinium HumanMethylation450 array that covers more than 480,000 methylation sites on the genome, we found that our method can greatly improve accuracy for cross tissue prediction (correlation square R2 increases from 0.81 (raw correlation R2) to 0.97 for blood->artery prediction, from 0.83 to 0.99 for blood->atrium prediction and 0.92 to 0.99 for LCL->PBL prediction based on cross-validation). These results indicate that methylation level in target tissue can be predicted using methylation measured in different tissues. We showed that methylation status at multiple CpG sites can be combined to further improve the prediction performance. This implies that using relatively easy accessible tissues (e.g. blood), large scale epidemiology investigations that target hard-to-access tissues could become feasible and non-invasive disease screening using epigenetic profile may also be achievable.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3470T

Comparative DNA methylation in TrkB gene across brain regions of suicide completers. *G. Maussion, B. Labonté, C. Ernst, A. Diallo, J. Yang, N. Mechawar, G. Turecki.*Department of Psychiatry, Douglas Hospital, McGill University, Montreal, QUEBEC, Canada.

TrkB gene codes for a receptor of BDNF. Previous studies by our group indicate that a subgroup of suicide completers have low expression levels of TrkBT1, an isoform which has no tyrosine kinase domain and is highly expressed in astrocytes. In previous studies, we observed epigenetic modifications in the TrkB promoter, which partially explained TrkB-T1 low expression level in brain tissue from suicide completers. The general aim of this study was to investigate DNA methylation throughout the TrkB-T1 gene sequence and adjacent genomic regions which could be responsible for the significant deregulation in TrkB-T1. Twelve low TrkB-T1 expressor suicides and 12 normal TrkB-T1 expressor controls were investigated. DNA was extracted from BA8/9, and the methylation fraction was pulled down by Methylated DNA immunoprecipitation. Microarray studies were performed using custom-made Agilent arrays tiling the whole TrkB gene and adjacent genomic sequence. After statistical correction for multiple testing, 9 DNA sequences located in TrkB gene were identified hypermethylated in frontal cortex of suicide completers. Four out of nine probes matched with intronic regions of TrkB gene whereas the five remaining probes corresponded to TrkB-T1 3'UTR sequence. These data are in agreement with the hypothesis of a transcript-specific deregulation through DNA methylation. We focused on the probes differentially methylated in this DNA region and validated them by clone sequencing of bisulfite treated DNA. Intergenic and extragenic regions present differential methylation between groups. Additional studies need to assess the functional significance of these findings. As DNA methylation is known to be region specific, we also aim to analyse the methylation pattern TrkB gene in the hippocampus where gene deregulations have been shown in brain from suicide completers.

3471T

Methylomic Variation In Alzheimer's Disease. *J. Mill[1], K. Lunnon[1], M. Volta[1], R. Smith[1], J. Powell[1], Y. Patel[1], N. Buckley[1], J. Cooper[1], C. Troakes[1], V. Haroutunian[2], S. Lovestone[1], L. Schalkwyk[1].* 1) SGDP Research Centre, Institute of Psychiatry, King's College London, London, United Kingdom; 2) The Mount Sinai School of Medicine, New York, USA.

Although the neuropathological changes associated with AD have been well characterized in post-mortem brain tissue, little is known about either the underlying etiology of the disorder or the precise mechanisms behind disease progression. In this project we aim to look beyond the traditional genetic and neuropathological etiological approaches to AD by testing the hypothesis that epigenetic phenomena play a crucial role in the development of the disorder. We have examined genome-wide patterns of DNA methylation across multiple brain-regions and blood, obtained from the same patients, using several complementary methylomic technologies (MeDIP-seq and Illumina 450K arrays) using samples from two well-characterized AD cohorts from the UK and US. We have identified several AD-associated differentially-methylated regions (DMRs); interestingly, some DMRs are observed in specific brain regions and others are common across tissues. We have also assessed cross-tissue patterns of allele-specific DNA methylation, identifying cis-acting genetic effects on epigenetic status in specific regions of the brain. Finally, we are using immunohistochemistry to quantify 5-methylcytosine, 5-hydroxymethylcytosine, 5-formylcytosine and 5-carboxylcystosine in tissue from AD patients and controls, and employing laser-capture microdissection to examine methylomic changes in purified populations of neurons and microglia from affected patients and controls. Ultimately we aim to integrate our methylomic data with other genomic, transcriptomic and proteomic data being collected from the same samples to undertake a truly integrative approach to the etiology of AD.

3472T

Cytosine modifications and cellular sensitivity to chemotherapeutics. *E.L. Moen[1], X. Zhang[2], S.M. Delaney[1], L.A. Godley[1], M.E. Dolan[1], W. Zhang[2].* 1) Department of Medicine, University of Chicago, Chicago, IL; 2) University of Illinois at Chicago, IL.

Covalent cytosine modifications, 5-methylcytosine and 5-hydroxymethylcytosine, occur in CpG dinucleotides throughout the human genome with genetic and environmental influences contributing to inter-individual differences. Our objective is to characterize inter-individual variation in cytosine modifications and investigate its contribution on gene expression and pharmacologic phenotypes, such as cellular sensitivity to chemotherapeutics. The International HapMap Project lymphoblastoid cell lines (LCLs) are an attractive model to address these questions at a genome-wide level due to publicly available genotype and gene expression datasets. In addition, our laboratory has gathered drug phenotype data for 14 chemotherapeutic agents. Thus, overlaying global cytosine modification data onto these existing datasets allows us to study how epigenetic variation affects cellular response to chemotherapeutic agents. We used the Illumina Infinium 450K HumanMethylation Beadchip array to measure cytosine modification levels at >485,000 CpGs. In this study, we included 73 unrelated African (YRI - Yoruba people from Ibadan, Nigeria) and 60 European ancestry (CEU - Caucasians from Utah, US) LCLs. After several filtering steps, we analyzed 283,540 CpGs and found that most of the inter-individual variation occurred outside of CpG islands, at CpGs in shores and shelves. Using gene expression data from the Affymetrix Human Exon Array, which measures the expression of 17,284 transcripts, we discovered 3,049 cis-acting modified cytosine-gene expression relationships at FDR<0.01. Interestingly, we saw both positive and negative associations between cytosine modification and gene expression. As expected, high levels of cytosine modification in gene promoter regions were more likely to be associated with negative gene expression. Conversely, high levels of cytosine modifications within the gene body were to more likely to be associated with positive gene expression. Importantly, cytosine modification heterogeneity and its consequences on gene expression contribute to variation in cellular response to chemotherapeutics. For example, a methylation locus in the gene body of PARD3 was found to significantly associated with both cisplatin and carboplatin IC50 in the YRI samples. These investigations will further our understanding of mechanisms of resistance to chemotherapeutic drugs.

3473T

Aberrant promoter hypermethylation of MMP16 gene in the colorectal cancer. *J. MOON[1], S. LEE[1], J. KIM[1], G. YOU[1], J. LEE[1], J. KIM[2], H. KIM[1], S. PARK[1].* 1) Anatomy, College of Medicine, Korea University, Seoul, South Korea; 2) Department of General Surgery, Korea University Medical Center, Seoul, South Korea.

We evaluated the methylation profile of 27,578 CpG sites spanning more than 14,000 genes in paired 20 colorectal cancer (CRC) tissues with methylation bead chip array-based technology. In array data, we confirmed 739 CpG sites as aberrant DNA methylation ($p<0.0001$) and 125 CpG sites was located in promoter region. There were strongly hypermethylated in CRC more than adjacent normal tissue ($p<0.0001$). Among these genes, matrix metallopeptidase 16 (MMP16, MT3-MMP) was one of the top-ranking methylation genes in 20 CRC tissues ($p<0.0001$). MMP16 involves ternary complex formation on the cell surface and possess the potential to promote melanoma invasion. To confirm the correlation between methylation and expression, we investigated the mRNA expression level of MMP16 by RT-PCR and western in CRC tissues compared adjacent normal tissues. The low expression level of MMP16 gene was correlated by about 30% on methylated CRC tissues. In addtion, MMP16 was down-expressed in three hypermethylated CRC cell lines. Futhermore, the mRNA expression of MMP16 in CRC cells was restored after 5-aza-2'-deoxycytidine treatment. In conclusion, our result revealed that the promoter hypermethylation of MMP16 is frequently present in CRC, may play role in colorectal cancer invasion. For understanding of MMP16 methylation effects, we are analyzing the correlation between hypermethylation of MMP16 onto CRC risk factor as migration and invasion.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3474T

Deep sequencing reveals allelic and parental inheritance patterns of genome-wide DNA methylation in large number of nuclear pedigrees with schizophrenia. *R.A. Ophoff[1,2,3], N. Plongthongkum[4], T. Wang[1], Y. Zhang[1], K.R. Van Eijk[3], R.S. Kahn[3], M.P.M. Boks[3], S. Horvath[2], K. Zhang[4].* 1) Center for Neurobehavioral Genetics, Semel Inst of Neuroscience, Univ California Los Angeles, Los Angeles, CA; 2) Department of Human Genetics, UCLA, Los Angeles; 3) Department of Psychiatry, University Medical Center Utrecht, The Netherlands; 4) Department of Bioengineering, UCSD, San Diego.

In an effort to study genome-wide epigenetic profiles and their inheritance patterns in the context of a neuropsychiatric disorder, we examined allelic DNA methylation in 22 nuclear pedigrees consisting of two parents, a proband affected with schizophrenia and one or two unaffected siblings. These subjects represent a relatively large number of potential informative meioses that can be used to robustly decipher epigenetic patterns of inheritance at a genome-wide scale, thereby facilitating genetic studies of disease susceptibility of complex traits. Using genomic DNA from whole blood we employed methylation sequencing with bisulfite padlock probes targeting >550,000 CpG sites genome-wide. After bisulfite treatment we generated libraries of 96 subjects and used barcodes for multiplexing of all samples simultaneously for data collection on a HiSeq2000 sequencer. We generated an average of 3Gbps of data per sample, and called an average of 504,825 CpG sites per subject, with >80% of the samples having over 450,000 CpG sites called. The sequence data identified more than 150,000 SNPs per subject allowing for genetic analysis and allelic-specific studies. As part of the initial data analysis we performed a mid-parent offspring analysis to estimate the heritability of CpG methylation. Our results suggest that 13.4% of CpGs with variable methylation profiles have strong evidence for a heritable component of which 60% could be explained by allelic variation at the CpG site itself. Thus far we have no evidence of chromosomal clustering of these CpGs. We also performed methylation QTL (mQTL) analysis and observed robust evidence of *cis* regulation of DNA methylation. We detected a general trend of increased *cis*-association signal closer to the CpG site with more than 50% of mQTLs within 2 kb of the methylated site. Overall, allele specific methylation (ASM) patterns within the pedigrees, show the strongest evidence for sequence-dependent ASM when SNPs are located within the CpG site itself. We are currently expanding our analyses to include patterns of imprinting, parent-of-origin and age effects. Examining the effect of disease status on CpG methylation to uncover possible biology underlying disease susceptibility completes our study.

3475T

Methylation Pattern of Twin Groups with Behcet's Syndrome. *M. Ozkilic[1], B. Shamloo[1], G. Celikyapi Erdem[1], E. Seyahi[2], H. Yazici[2], E. Tahir Turanli[1,3].* 1) Dr. Orhan Öcalgiray Molecular Biology-Biotechnology and Genetics Research Center, Istanbul Technical University, Istanbul, Turkey; 2) Department of Rheumotology, Cerrahpasa School of Medicine, Istanbul University, Istanbul, Turkey; 3) Molecular Biology and Genetics Department, Science and Letter Faculty, Istanbul Technical University, Istanbul, Turkey.

Objective: Behçet's Syndrome (BS) is a multisystem vasculitis of unknown etiology. HLA-B51 is the strongest marker associated with BS. Phenotypic discordance between BS monozygotic twins points to involvement of epigenetic or environmental factors in the etiology. We hypothesized that global methylation profile are different between discordant BS twins. Global methylation levels were compared among concordant and discordant 4 MZ and 4 DZ BS twins with age-gender matched healthy MZ and DZ twins (n=8). Methods: DNA was isolated from venous blood of BS and control twins. DNA was digested into nucleotides and then nucleotides converted to nucleosides. Global methylation value was obtained by injecting digested DNA samples to HPLC and separating nucleosides including dC and 5-mdC. The area below the peaks was recorded. Tests were triplicated, 5-mdC: dC ratio for each of the samples was calculated (%RSD= 7%). Authentic dc and 5-mdc was used as standard for the experiment to determine and compare dC:5-mdC ratios. The average values were compared between groups using student T test analyses. Results and Conclusions: The difference between 5-mdC: dC ratio was found to be more correlated with each other in healthy control twins compared to the discordant BS twins (P= 0.02847). However the global methylation level did not change with the disease status such that both BS (mean= 0.032 sd=0.006) and control MZ/DZ (mean= 0.030 sd=0.001) twins found to have similar total methylation levels. This finding is in agreement with previous studies on the level of global DNA methylation level in human genome, but it is in contrast with another study in which decreased global methylation was shown in SLE pairs of discordant MZ twins. We are further analyzing our sample for locus specific methylation and we are including other disease control twins such as SLE and FMF into our dataset.

3476T

Integrated analysis of genome-wide DNA methylation and gene expression in post-mortem schizophrenia brain samples. *R. Pidsley, J. Viana, C. Troakes, S. Al-Sarraj, L. Schalkwyk, J. Mill.* Institute of Psychiatry, London, London, United Kingdom.

Schizophrenia is a severe psychiatric disorder characterised by psychotic symptoms and cognitive deficits. Recent work suggests that epigenetic factors regulating gene expression, including DNA methylation, may play a role in the molecular aetiology of the disorder. Patterns of DNA methylation and gene expression are tissue-specific, so brain tissue is optimal for the identification of methylomic and transcriptomic changes associated with schizophrenia. In this study, DNA and RNA were extracted from a unique collection of well-characterised, neuropathologically dissected human post-mortem brain samples: frontal cortex (n=47) and cerebellum (n=47) samples, comprising 23 schizophrenia patients and 24 healthy controls. Genome-wide methylation and expression analyses were performed using the Illumina Infinium Human Methylation450 BeadChip and the Illumina HumanHT-12 v4 Expression BeadChip, respectively. Integrated analyses to test for genome-wide association between DNA methylation, gene expression and schizophrenia are ongoing. Initial results show tissue-specific, genome-wide significant differences in DNA methylation between cases and controls at numerous loci. I will report on these results, which give an insight into the role of epigenetics in schizophrenia, and on the results of the integrated analysis, which will provide additional information about the underlying biological mechanism of association.

3477T

Determination of methylation profile in patients with schizophrenia. *B. Rukova[1], R. Staneva[1], S. Hadjidekova[1], D. Nikolova[1], V. Milanova[2], D. Toncheva[1].* 1) Department of Medical Genetics, Sofia, Bulgaria; 2) Department of Psychiatry, Sofia, Bulgaria.

Schizophrenia is a severe chronic mental disorder that affects most of the higher brain functions. Its prevalence is estimated on about 1% worldwide. The classical symptoms occur in several other psychiatric disorders, which makes difficult the exact diagnoses. Schizophrenia is a serious social and economic healthcare problem. Currently, there is no etiological treatment. DNA methylation is a major epigenetic modification. It is a biochemical process that is important for normal organismal development and cellular differentiation. DNA methylation stably alters the gene expression pattern in cells. This modification can be inherited through cell division. Materials and methods: We analyzed age matched pools of 220 schizophrenia patients and 220 healthy controls. We've performed high-resolution genome-wide methylation array analysis (Agilent 1×244K). We've analyzed the methylation status of 27800 CpG islands of both groups to identify methylation profile differences. Results: Our experiments show significant difference in the methylation profile between patients and controls. In patients group we established 59 hypermethylated genes and 37 hypomethylated genes, compared to healthy controls. Some of these hypermethylated genes play a role in neuronal development, methylation and mediation of neuronal impulses like ESPNP, GNMT,BHLHE23,CRELD2. Some of the hypomethylated in patients genes like PROX1, FGF13 and GABBR1 take part in neuronal development, signal transmission and cell growth. Conclusions: Our data suggest that there is a major differences in methylation profile between patients and controls. This dysregulation can play a critical role in schizophrenia pathogenesis. Acknowledgements: funded by projects DO 02-12/2009 and DMY 03-36/2011, Ministry of Education and Science.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3478T

Association of CpG island shore methylation in peripheral blood with disease state and SNP variants in the MLH1 region in colorectal cancer patients. *A. Savio[1,2], M. Lemire[3], M. Mrkonjic[1,2], S. Gallinger[1,2,4,5], B. Zanke[6], T. Hudson[3], B. Bapat[1,2,7].* 1) Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, ON; 2) Samuel Lunenfeld Research Institute, Toronto, ON; 3) Ontario Institute for Cancer Research, Toronto, ON; 4) Ontario Familial Colorectal Cancer Registry, Cancer Care Ontario, Toronto, ON; 5) Department of Surgery, University of Toronto, Toronto, ON; 6) Ottawa Hospital Research Institute, Ottawa, ON; 7) Department of Pathology, University Health Network, Toronto, ON.

Epigenetic gene silencing via hypermethylation of CpG islands in tumor cells contributes to colorectal cancer (CRC) development. We previously demonstrated that SNPs (rs1800734, rs749072, and rs13098279) in the MLH1 (MutL homolog 1) gene region are associated with MLH1 promoter CpG island methylation, loss of MLH1 protein expression, and tumor microsatellite instability (MSI). Recent studies have identified less CpG-dense "shore" regions flanking many CpG islands. These CpG island shores often exhibit distinct methylation profiles between different tissues and normal versus tumor cells of patients. To date, most studies have focused on somatic methylation events within tumor tissue while few have examined the contributions of peripheral blood cell (PBC) methylation changes to CRC risk. To address whether MLH1 methylation in PBCs is correlated with tumorigenesis we utilized Illumina methylation microarrays to measure genome-wide methylation at 450,000 CpG sites in DNA extracted from PBCs of 884 CRC cases and 845 healthy controls. Analysis of chromosome 3p21 revealed that a CpG island shore 1 kb upstream of MLH1 incurs significantly lower methylation in CRC patients than controls. When stratified by SNP genotypes of rs1800734, rs749072, and rs13098279, individuals with wild-type genotypes incur significantly higher PBC CpG island shore methylation than heterozygous or homozygous variant individuals in both CRC cases and controls. CpG island shore methylation also decreases significantly with increasing age in cases and controls. This is the first study of its kind to integrate PBC methylation at a CpG island shore with SNP genotype status in CRC cases and controls. We have found decreased methylation in CRC patients, individuals possessing variant SNP alleles, and older individuals. This decreased methylation at the MLH1 CpG island shore may indicate genomic instability possibly predisposing to CRC.

3479T

The DNA methylation landscape of human placenta. *D. Schroeder[1], H. Yu[2], P. Lott[2], D. Hong[1], C. Walker[3,4], I. Korf[2], J. LaSalle[1,2,4].* 1) Medical Microbiology and Immunology, UC Davis, Davis, CA; 2) UC Davis Genome Center, UC Davis, Davis, CA; 3) Obstetrics & Gynecology, UC Davis School of Medicine, Sacramento, CA; 4) UC Davis M.I.N.D. Institute, UC Davis School of Medicine, Sacramento, CA.

As genome-scale DNA methylation sequencing technologies have improved it has become apparent that tissue-specific methylation can occur not only at promoters, enhancers, and CpG islands but also over larger genomic regions. In most human tissues, the vast majority of the genome is highly methylated (>70%). However, genomic sequencing of bisulfite-treated DNA (MethylC-seq) has revealed large partially methylated domains (PMDs) in some human cell lines. PMDs can cover as much as 40% of the genome in fetal lung fibroblasts and are associated with gene repression and inactive chromatin marks. We have previously demonstrated PMDs in human SH-SY5Y neuroblastoma cells and shown that tissue specific PMDs define a subset of neuronally expressed genes with synaptic functions. However, to date only cultured cells and some cancers have shown evidence for PMDs, suggesting that PMDs may not be observed in normal human tissues.

Here we performed MethylC-seq in full-term human placenta, discovering the first normal human tissue showing clear evidence of PMDs. Similar to our observations in cultured human cells having PMDs, RNA-seq analysis confirmed that genes in PMDs are repressed in placenta. Using a hidden Markov model, we mapped placental PMDs genome-wide and compared them to mapped PMDs of two other human cell lines (IMR90 and SH-SY5Y). GO analysis showed that genes in placental PMDs had tissue-specific functions and that combining PMD information from multiple cell and tissue types increased the statistical significance of the tissue-specific GO terms. A deeper analysis of DNA methylation patterns in placenta showed that methylation levels in promoter CpG islands were actually higher for genes within PMDs, despite the lower overall methylation of surrounding regions, perhaps due to protection from active demethylation of CpG islands in PMDs. These results suggest that PMDs are a developmentally dynamic feature of the methylome that is relevant for understanding not only cancer but also normal tissue development. Future investigations of human placenta by MethylC-seq may be informative for finding epigenetic biomarkers of diseases such as neurodevelopmental disorders.

This research was funded by NIH 2R01HD041462.

3480T

Methylation Quantitative Trait Loci (mQTL) are Consistently Detected across Ancestry, Age and Tissue Type. *A.K. Smith[1], V. Kilaru[1], M. Kocak[2], L.M. Almli[1], K.B. Mercer[3], E.B. Binder[4], K.J. Ressler[1,3], F.A. Tylavsky[2], K.N. Conneely[5].* 1) Psychiatry & Behavioral Sciences, Emory University, Atlanta, GA; 2) Department of Preventive Medicine, University of Tennessee Health Science Center, Memphis, TN; 3) Howard Hughes Medical Institute, Chevy Chase, MD; 4) Max-Planck Institute of Psychiatry, Munich, Germany; 5) Department of Human Genetics, Emory University, Atlanta, GA.

Individual genotypes can result in different patterns of DNA methylation. These methylation Quantitative Trait Loci (mQTLs) influence methylation across extended genomic regions and may underlie direct SNP associations or gene-environment interactions. mQTLs have been identified in multiple tissues and cohorts, but only a limited number are consistent across studies. We hypothesize that the detection of mQTLs will vary with ancestral population, age, and tissue type. To explore this, we analyzed 7 datasets that varied by race (Caucasian [CC] vs. African American [AA]), age (neonate vs. adult) and tissue type (blood vs. 4 regions of postmortem brain) with genome-wide DNA methylation (HumanMethylation27 BeadChip) and SNP data (Illumina or Affymetrix arrays). SNPs that were common to all arrays (~90k), had a MAF <.05 and did not deviate significantly from HWE ($p<.00001$) were included. Methylation probes that contained SNPs were excluded. We tested for mQTLs by performing linear regressions of methylation levels at each CpG site on SNP genotypes within 50kb under an additive model, controlling the false discovery rate at .05. We identified 591 mQTLs in CC umbilical cord blood (N=87), 116 in AA cord blood (N=87), 146 in AA adult blood (N=90), 380 in CC pons (N=125), 332 in CC cerebellum (N=119), 520 in CC frontal cortex (N=133) and 707 in the CC temporal cortex (N=127). Significant overlap of mQTLs was observed between all pairs of sample groups ($p<2.4\times10\text{-}13$). For example, rs2066992 (interleukin 6; IL6) associates with genotype-dependent DNA methylation of IL6 CpG sites in 6 of the 7 datasets evaluated ($1.18\times10\text{-}11<p<5.5\times10\text{-}5$). The highest rates of overlap were observed within the 4 brain regions (31–61%) and in comparisons of adult whole blood and cord blood samples (32–44%), even with the differences in race and age of samples. Overlap between blood and brain datasets was lower (15–34%) but still highly significant. The SNP-CpG relationships that were common between at least 1 blood and 1 brain dataset were enriched for biological functions including cell growth and signaling, cardiovascular development and disease, hematological development and disease, tissue and organism development, neurological and psychological disorders and hereditary disorders. These data provide insight into the mechanisms by which SNPs impact gene regulation and supports the notion that peripheral blood may be a reliable correlate of physiological processes in other tissues.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3481T

Whole-genome methylation analysis- new insights in BEN molecular ethiology. *R. Staneva[1], B. Rukova[1], I. Dimova[1], S. Hadzhidekova[1], O. Boyanova[1], M. Polenakovic[2], V. Stefanovic[3], R. Cukuranovic[3], V. Djonov[4], A. Galabov[5], D. Toncheva[1].* 1) Department of Medical Genetics, Medical University of Sofia, Sofia, Bulgaria; 2) Macedonian Academy of Sciences and Arts, Skopje, The Former Yugoslav Rrepublic of Macedonia; 3) Faculty of Medicine, Institute of Nephrology, University of Nis, Nis, Serbia; 4) Institute of Anatomy, University of Bern, Bern, Switzerland; 5) Institute of Microbiology, Bulgarian Academy of Sciences.

BACKGROUND: Balkan endemic nephropathy (BEN) represents a chronic progressive interstitial nephritis in striking correlation with uroepithelial tumors of the upper urinary tract. The disease has endemic distribution in the Danube river regions in several Balkan countries and in some endemic regions shows markedly predominance of affected females. DNA methylation is a primary epigenetic modification that is involved in major processes such as cancer, genomic imprinting, gene silencing etc. Epigenetic tests can prove to be the bridge between environmental factors and genetic background in BEN development. MATERIALS AND METHODS: Age matched pools of 19 female BEN patients and 33 healthy female controls from nonendemic regions were created as well as age matched pools of 30 male BEN patients and 31 healthy male nonendemic controls from Serbian ancestry. We've performed high-resolution genome-wide methylation array analysis on Agilent Human CpG island Microarray(1×244K). The data was analyzed by Agilent Genomic Workbench Lite v.6.5.0.18. RESULTS: We compared methylation status of the affected group and unaffected group by our own software, applying stringent criteria. In the female group 198 gene-related CpG islands showed disparity in methylation status between affected and unaffected individuals verus 228 CpG islands in male group. We further compared the males and females data and 26 CpG islands proved to be common for both groups. In both female and male patient group some transmembrane ion chanells, zinc-finger proteins, TNF-path related proteins,enzymes and DNA-modification proteins were hypomethylated. Hypermethylation in both patient groups showed some nuclear receptor proteins, cytoskeletal proteins and protein kinases. CONCLUSIONS: Data obtained from our experiments suggest that dysregulation of cytoskeletal proteins, transcription factors, transmembrane ion channels as well as proteins involved in DNA-splicing and cell proliferation can be key mechanism in BEN pathogenesis. These results are in unison with the key pathological alterations in BEN and further elucidate the precise molecular mechanism behind BEN development. Acknowledgements: funded by project DMU03/35 by Bulgarian Ministry of Education and Science, project IZ73Z0_ 127949 of SCOPES, SNSF, Switzerland.

3482T

Genome-wide differences in DNA methylation identified in children with Williams-Beuren Syndrome. *E. Strong[1], D. Butcher[2], C.B. Mervis[3], C.A. Morris[4], R. Weksberg[1,2], L.R. Osborne[1,5].* 1) Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada; 2) Program in Genetics and Genome Biology, SickKids Research Institute, Toronto, Ontario, Canada; 3) Department of Psychological and Brain Sciences, University of Louisville, Louisville, KY; 4) Department of Pediatrics, University of Nevada School of Medicine, Las Vegas, NV; 5) Department of Medicine, University of Toronto, Toronto, Ontario, Canada.

Williams-Beuren Syndrome (WBS) is a neurodevelopmental disorder caused by the hemizygosity of 26 genes on chromosome 7q11.23. People with WBS exhibit an array of cognitive and behavioral features including intellectual disability, social disinhibition, anxiety, deficits in visuospatial construction and attention deficit hyperactivity disorder, but the neurobiological basis for these symptoms remains unknown. Several genes within the deleted region have been associated with epigenetic complexes or mechanisms. BAZ1B is a component of the ISWI family of ATP dependent chromatin remodelling complexes WINAC, WICH and B-WICH. The general transcription factors GTF2I and GTF2IRD1 have been shown to interact with the histone deacetylase HDAC3. Furthermore, WBSCR22, an as yet uncharacterised protein, contains an S-adenosyl-methionine binding motif, typical of methyltransferases. Deletion of one or more of these genes could result in altered epigenetic pathways, which may contribute to the phenotypic presentation of WBS. To explore the possibility of disrupted epigenetic pathways in WBS, we assessed genome-wide DNA methylation levels using the Illumina HumanMethylation450 array. Comparison of whole blood DNA from a discovery cohort of 14 male children with WBS and from 10 age- and sex-matched controls, identified 635 statistically significant differentially methylated CpG dinucleotides (minimum beta value ±0.17 and adjusted $p<0.05$), corresponding to 263 unique genes, one large gene cluster and many inter-genic regions. We identified a significant gain of methylation across the protocadherin cluster, a 700kb gene cluster of neuronal cell-adhesion proteins, located on chromosome 5q31. Furthermore, we identified altered CpG methylation within several genes involved in synaptic plasticity and/or axon guidance. This is the first report of aberrant genome-wide DNA methylation in WBS and our results suggest that one or more genes within the 7q11.23 deleted region modify the epigenome at specific loci. We have identified aberrant DNA methylation across numerous genes with neurological functions, potentially implicating many additional genes and pathways outside the 7q11.23 region in the neurological phenotypes of WBS. Further validation of these findings through the analysis of additional DNA samples and correlation with gene expression is currently underway.

3483T

Genome-wide Profiling of 5-Formylcytosine in Embryonic Stem Cells. *K.E. Szulwach[1], C.X. Song[2], Q. Dai[2], Y. Fu[2], P. Liu[3], L. Li[1], Y. Li[1], G.L. Xu[3], C. He[2], P. Jin[1].* 1) Dept Human Gen, Emory Univ, Atlanta, GA; 2) Dept Chem and Inst Biophysical Dynamics, Univ of Chicago, Chicago, IL; 3) Group of DNA Metabolism, State Key Laboratory of Molecular Biology, Inst of Biochem and Cell Biology, Shanghai Inst for Biol Sciences, Chinese Academy of Sciences, Shanghai, China.

TET-family proteins oxidize 5-methylcytosine (5mC) stepwise to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC). 5fC and 5caC can be excised by the DNA glycosylase TDG, implicating 5mC oxidation in active DNA demethylation. Here we have developed and implemented a novel method for the enrichment of 5fC by coupling selective chemical reduction of 5fC with subsequent biotin tagging and streptavidin-mediated affinity capture, thereby enabling its genome-wide mapping. We applied this method to map the genomic distribution of 5fC in embryonic stem cells (ESCs) in parallel with 5hmC. Our results demonstrate that 5fC-enriched regions of the genome are depleted of 5mC concomitant with an increased frequency and abundance of 5hmC compared to 5hmC-marked regions, suggesting that specific mapping of 5fC allows for further refinement of genomic elements that are subjected to oxidation, more so than mapping 5hmC alone. Genomic and epigenomic annotation of 5fC-enriched regions demonstrated that they occur more frequently than expected at distinct gene regulatory elements, including known Tet1 binding sites, H3K4me1 marked putative enhancers, and sites of DNase I hypersensitivity, as compared to 5hmC. Additional profiling of 5fC in *Tdg*-null mouse ESCs revealed accumulation of 5fC at known Tet1 binding sites as well as p300 bound loci, CTCF binding sites, and a subset of transcription start sites. Examination of the sequence content within 5fC-enriched regions specific to *Tdg*-null ESCs revealed reduced CpG frequency and GC content, indicating that local sequence context influences oxidation of 5mC and TDG-dependant removal of 5fC at defined gene regulatory regions. Our results establish a novel methodology for genomic profiling of 5fC, providing the first genome-wide view of this mark, and suggest important roles for TET-mediated 5mC oxidation coupled with TDG-dependant removal of 5fC at defined gene regulatory elements.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3484T

Extremely preterm infants are epigenetically different from term infants at birth and an epigenetic legacy of preterm birth may remain at 18 years of age. *C. Theda[1,2], M. Cruickshank[2], A. Oshlack[2], P. Sheehan[1], R. Saffery[2], P. Davis[1,2], LW. Doyle[1,2], J. Craig[2].* 1) Royal Women's Hospital, Melbourne, Australia; 2) Murdoch Children's Research Institute, Melbourne, Australia.

Extremely preterm infants (born at less than 28 weeks gestation) have more short and longterm ill-health than term infants. Preterm birth and medical interventions are stressful for the infants affecting early postnatal nutrition and growth. Environment, stress and nutrition can persistently alter gene regulation, not through changes in DNA sequence but through epigenetic changes that are mitotically heritable. Epigenetic state is programmed early in development but can undergo modification during life. Our aim was to investigate whether (i) infants born premature are epigenetically different from those born at term and (ii) if prematurity leaves an epigenetic legacy detectable later in life. We examined genome wide DNA methylation profiles after extracting DNA from dried blood spots of 12 extremely preterm infants and 12 term probands obtained at birth and 18 years. DNA methylation was quantified using the Illumina Infinium Human Methylation 450 BeadChip array covering 485,000 genome wide CpGs. After stringent quality control, exclusion of X/Y probes and probes targeting known polymorphisms, 349,013 probes remained for statistical analysis. Combined analysis of birth and 18 yr samples identified 112 differentially methylated probes (DMPs; significant at $p<0.05$ after adjustment for multiple testing) comparing preterm and term probands. We identified 1,546 DMPs using birth samples only and detected no significant DMPs using only 18 yr samples. To explore a potential long term effect of premature birth on DNA methylation, we also analyzed the top 1,500-ranked sites (by odds) from the 18 yr samples: multidimensional scaling and clustering dendrograms revealed that this set of probes separates both birth and 18 yr samples according to proband groups. Of the 112 significant DMPs identified during the combined analysis, 16 probes overlapped with the top 1500-ranked probes using the 18 yr samples and 104 probes overlapped with the significant DMPs identified in birth samples. Our results indicate: (i) significant DNA methylation differences at birth comparing premature and term infants; possibly a gestational age effect; (ii) a distinct subset of methylation differences that are present at birth and persist to 18 years. While our results are preliminary and the numbers small, they are the first report of long term epigenetic effects of prematurity. Replication of the reported findings and functional analysis of genomic sites of interest are underway.

3485T

Genome Wide Detection and Analysis of CpG Dinucleotide Methylation in Cancer Fatigue Studies. *D. Wang[1], H. Lee[1], W. Xiao[2], C. Hsiao[1], H. Kim[1], R. Dionne[1], L. Saligan[1].* 1) NINR/NIH, Bethesda, MD, USA; 2) CIT/NIH, Bethesda, MD, USA.

Objectives: DNA methylation is an epigenetic mechanism involved in transcriptional regulation in development and cellular differentiation. Aberrant methylation patterns have been observed in a large number of human cancers. In this study, we investigated the role of DNA methylation in fatigue progression in men with non-metastatic prostate cancer receiving localized external beam radiation therapy (EBRT). Methods: We have applied Methyl Binding Domain (MBD)-seq method to peripheral blood DNA of 8 patients at the baseline (pre-treatment) and midpoint (21 days after receiving ERBT) when fatigue severity normally peaks. Methylated DNA was first precipitated by binding to the methyl—CpG binding domain of human MBD2 protein, then subjected to next-generation sequencing via 36-bp single-end module on the Genome Analyzer IIx (Illumina, San Diego, CA). The sequencing data were further analyzed including data processing, quality control, normalization and statistical analysis. The genes that were most differentially methylated after EBRT were investigated by comparing the expression profiles and changes of expression were confirmed with qPCR. Results: A total of 40 to 50 million reads were generated for each DNA sample, of which 73% to 80% could be mapped uniquely to the reference human genome (NCBI 37/hg19). Reads that were unambiguously mapped to a single location in the reference human genome were extended to 200 bp and hits were counted in every 25 bp bin which then led to peak identification. After filtering, we observed 350,000 to 490, 000 non-overlapping peaks with significant evidence of methylation in each DNA sample, of which 40,000 to 180,000 were found to be unique in the same patients between the baseline and midpoint of the EBRT. We further investigated the relationship between the methylation profiles and gene expression levels for the genes with most differentially methylated regions in at least half of the patients (4 out of 8). We could confirm the role of DNA methylation in transcriptional regulation of gene expression. Conclusion: This study is the first to provide a large scale and unbiased analysis of changes in DNA methylation occurring just 21 days after EBRT initiation in peripheral blood of men with non-metastatic prostate cancer. Based on the methylation profile changes along with their associated gene expression changes in these patients, we could suggest the role of DNA methylation in fatigue progression after EBRT.

3486T

Prediction of total serum IgE using genome-wide methylation profile. *X. Wang[1], S.A.G. Willis-Owen[2], K.C.C. Wong[2], A. Binia[2], G. Davies[3], J.M. Hopkin[3], G.M. Lathrop[4], M.F. Moffatt[2], W.O.C.M. Cookson[2], L. Liang[1,5].* 1) Department of Biostatistics, Harvard School of Public Health, Boston, MA; 2) National Heart and Lung Institute, Imperial College, London SW3 6LY, UK; 3) Institute of Life Science, School of Medicine, Swansea University, UK; 4) CEA / Centre National de Genotypage, 91057 Evry, France; 5) Departments of Epidemiology, Harvard School of Public Health, Boston, MA 02115.

DNA methylation at CpG sites is an important epigenetic modification that may regulate gene expression. There is a growing interest in understanding how the methylation inheritance contributes to the development of complex diseases or traits. It has been shown that methylation modification may influence individual asthma risk and related phenotypes. The primary purpose of this study is to comprehensively assess—by using genome-wide DNA methylation data as markers—the contribution of epigenetic effects on asthma related quantitative traits. Results of variance component analyses on our data indicated that a considerable proportion of phenotypic variation in total serum IgE (an important biomarker for asthma) can be explained by the variation in epigenetic markers. To evaluate the clinical utility of epigenetic markers, we then constructed and compared various prediction models by including top ranked methylation loci from the genome-wide association scan, together with selected sets of known genetic markers from published genome-wide association studies (GWAS). A new prediction model based upon Best Linear Unbiased Prediction (BLUP) was further proposed where all CpG sites (on the Illumina Infinium 27K methylation array) were simultaneously modeled. The overall prediction accuracies of the proposed methods were extensively evaluated via the cross-validation analysis. We observed a significant increase of correlation coefficient between actual and predicted IgE level when methylation markers were included (from 0.32 before to 0.57 after including methylation data). By using an independent sample based on Illumina 450K methylation array, we also assessed the performance of cross platform prediction using methylation markers. Taken together, results from this study suggest that DNA methylation has important influence in asthma and it explains much larger variability in IgE level than known genetic variants (2% due to top genetic markers in large GWAS vs. 15% due to top 3 CpG sites). Our comprehensive assessment suggests that methylation has great potential in prediction of clinical phenotypes.

3487T

Epigenomic Profiling of Autism Spectrum Disorder. *C.C.Y. Wong[1], L.C. Schalkwyk[1], E.L. Meaburn[1,4], A. Ronald[1,4], T.S. Price[1], R. Plomin[1], D.H. Geschwind[3], J. Mill[1,2].* 1) MRC SGDP Centre, King's College London, Institute of Psychiatry, London, U.K; 2) University of Exeter, U.K; 3) UCLA Department of Neurology. 710 Westwood Plaza. Los Angeles, CA 90095-1769. USA; 4) Birkbeck, University London, U.K.

Autism Spectrum disorder (ASD) is a group of common, complex neurodevelopmental disorders characterized by marked etiological heterogeneity. Evidence from twin and family studies suggest that ASD has a strong heritable component, supported by the identification of several susceptibility loci and the growing body of literature demonstrating the association of de novo and inherited copy number variations (CNVs) with disease. Recent evidence also implicates a role for epigenetic variation in the etiology of ASD. In this study, we first performed epigenomic profiling using Illumina 27K Methylation arrays in whole blood samples from a unique set of MZ twins discordant for ASD and ASD-related traits (n=53 twin pairs; 106 individuals). We identified a number of significant disease-relevant differentially methylated genes and enrichments of disease-related biological networks within discordant MZ pairs as well as between cases and controls. Second, we performed DNA methylation profiling in post-mortem brain tissues from multiple brain regions obtained from the Autism Tissue Program and the London Brain Bank. Illumina Infinium 450K Human Methylation arrays and Agilent Methylation SureSelect target enrichment combined with ultra-deep bisulfite sequencing were used. We observed significant DNA methylation differences between cases and controls using both approaches with several potential ASD-associated epigenetic biomarkers located in the vicinity of genes previously implicated in psychiatric disorder.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3488T

Microarray-Based Genome-Wide Analyses of DNA Methylation in Hematological Malignancies. *Y. Yang[1], H. Mei[1], L. Edelmann[1], A.J. Sharp[1], I. Peter[1], V. Najfeld[2], C.R. Geyer[3], J.F. DeCoteau[3], S.A. Scott[1].* 1) Department of Genetics and Genomic Sciences, Mount Sinai School of Medicine, New York, NY 10029; 2) Department of Pathology, Mount Sinai School of Medicine, New York, NY 10029; 3) Cancer Stem Cell Research Group, University of Saskatchewan, Saskatoon, SK, Canada, S7N 4H4.

Aberrant DNA methylation is directly involved in human imprinting disorders and is a hallmark of many cancers, including hematological malignancies. Genome-wide methylation analysis can be achieved using microarrays following either methylated DNA immunoprecipitation (MeDIP) or bisulfite modification. Our aim was to assess the MeDIP-based Human DNA Methylation Microarray (MeDIP-array; Agilent Technologies) using 16 hematological cancer cell lines by comparing results to those derived from the Infinium HumanMethylation450 BeadChip (BC-array; Illumina), a quantitative bisulfite-based array. A high correlation for all 16 samples was observed between MeDIP- and BC-array data when comparing overlapping 1kb genomic intervals (0.83±0.04). Bisulfite sequencing of 9 differentially methylated regions in AML193 cells showed stronger correlation with BC-array (0.99) than MeDIP-array (0.90); however, the BC-array correlation was different depending on which type of probe chemistry was assessed (Type I Infinium: 0.94; Type II Infinium: 0.84). Reassessment of BC-array data based on probe type indicated that Type II probes consistently performed with a reduced dynamic range compared to Type I probes, which necessitated additional normalization. Hierarchical clustering of the identified methylation profiles indicated that both MeDIP- and BC-array could generally distinguish disease subtypes (AML, CML, ALCL, BL, PCL) among the tested cell lines; however, the BC-array data was a more accurate predictor, particularly when clustering only by promoter region methylation. After subtracting the methylated genes detected in peripheral blood specimens from 36 healthy multi-ethnic controls, 373 to 3139 genes with promoter methylation were identified among the five different disease subtypes, and 92 were methylated across all 16 cell lines. Additionally, treatment of AML193 cells with a DNA methylation inhibitor, 5-aza-2'-deoxycytidine, resulted in a prominent reduction in detected methylation by both MeDIP- and BC-array (MeDIP log2 ratio median: 0.50 vs 0.34; BC beta value median: 0.86 vs 0.67). Taken together, these data indicate that both MeDIP- and BC-arrays are useful for studying whole-genome methylation profiles and identifying genes aberrantly silenced by promoter methylation in hematological and other malignancies. However, the inability to accurately quantitate DNA methylation levels by MeDIP-array may be a limitation depending on study design and research question.

3489T

Discovery of Cross-Reactive Probes in the Illumina Infinium HumanMethylation450 Microarray: A Cautionary Tale. *Y. Chen[1,5], M. Lemire[2], S. Choufani[1], D.T. Butcher[1], D. Grafodatskaya[1], B.W. Zanke[2,3], R. Weksberg[1,4].* 1) Genetics and Genome Biology Program, Hospital for Sick Children, Toronto, Ontario, Canada; 2) Ontario Institute for Cancer Research, Toronto, Ontario, Canada; 3) Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, Canada; 4) Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada; 5) Institute of Medical Science, University of Toronto, Toronto, Ontario, Canada.

DNA methylation, an important type of epigenetic modification in humans, participates in crucial cellular processes such as embryonic development, X-inactivation, genomic imprinting and chromosome stability. Several platforms have been developed to study genome-wide DNA methylation. Many investigators in the field have chosen the Illumina Infinium HumanMethylation microarrays for its ability to reliably assess DNA methylation following bisulfite conversion. We have recently reported that 6–10% of probes on the Illumina Infinium HumanMethylation27K microarray are cross-reactive, co-hybridizing to alternate sequences that are highly homologous to the intended targets.

Here we report our findings on the new Illumina array, the Infinium HumanMethylation450K microarray which has >17-fold increase in coverage spanning multiple genomic regions, compared to the 27K array which covers only promoter regions. We analyzed data from the 450K methylation profiles of 489 males and 357 females from the control cohort of the Assessment of Risk for Colorectal Cancer Tumours in Canada (ARCTIC) project. Among the autosomal CpG sites that displayed significant methylation differences between males and females, we observed an enrichment of cross-reactive probes co-hybridizing to the sex chromosomes. This can lead investigators to mistakenly infer the existence of significant autosomal sex-associated methylation differences. Based on the sex methylation analysis, we concluded that ~5% of the probes are potentially cross-reactive. In addition, we found a number of probes targeting polymorphic CpGs (i.e. SNPs overlapping CpG sites). The methylation levels detected by these probes are simply the reflection of underlying genetic polymorphisms and thus could be misinterpreted as true signals.

The existence of cross-reactive probes and polymorphic CpGs in the Illumina microarrays reflecting homologous/repetitive sequences and SNPs in the human genome can confound data obtained from these microarrays. Therefore, investigators should exercise caution when significant associations are found on these platforms. A list of all the cross-reactive probes will be made available to the scientific community. Finally, our data suggest that biological interpretation of significant microarray findings continue to be validated by a second independent method such as sodium-bisulfite pyrosequencing.

3490T

Epigenetic changes in response to common environmental compounds. *V. Labrie[1], I. Rezaian[2], S. Konigorski[3], N. Miller[4], Y. Li[2], C. Austin[4], R. Tice[5], M. Xia[4], A. Ngom[2], L. Rueda[2], R. Kustra[3], A. Petronis[1].* 1) Krembil Family Epigenetics Laboratory, Centre for Addiction and Mental Health, Toronto, Ontario, Canada; 2) School of Computer Science, University of Windsor, Windsor, Ontario, Canada; 3) Division of Biostatistics, University of Toronto, Toronto, Ontario, Canada; 4) National Institutes of Health Chemical Genomics Center, Rockville, Maryland, USA; 5) National Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina, USA.

Environmental chemicals have the potential to impact disease susceptibility, and thereby pose an important public health concern. The mechanisms through which environmental contaminants influence human health are not well-defined and studies examining the biological effects of environmental chemicals will facilitate disease prevention and management. We have initiated a large-scale study to examine the possibility that environmental agents target the epigenome. As a preliminary investigation, we established the experimental parameters that are necessary for our study. Genome-wide changes in DNA methylation were examined in human cells exposed to several agents that are known to induce epigenetic alterations. A series of times points and concentrations were tested to identify the optimal design for the study of epigenetic effects of chemical agents. Furthermore, certain genomic regions were found to be consistently more affected by the chemical compounds. This study is a first-step in understanding how environmental chemicals can impact the epigenome and affect genomic functions relevant to the etiology of complex diseases.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3491T

Regulatory network of microRNA action - miR-199a in health and diseases. *S. Gu, Y.K. Suen, W.Y. Chan.* School of Biomedical Sciences, The Chinese University of Hong Kong, HKSAR, Hong Kong.

It is well-known that microRNAs (miRNAs) affect gene expression by regulating the translation of mRNAs. However, how miRNAs are regulated is less studied. The well-conserved human miR-199a plays important roles in different systems. MiR-199a is encoded by two loci in the human genome, miR-199a-1 on chromosome 19 and miR-199a-2 on chromosome 1. Both loci encode the same miRNA. Past studies showed that miR-199a behaves quite differently in different systems and diseases. The cause of this diverse action of the miRNA is not clear. However, it is likely caused by different regulation of the two genomic loci and variable targets of the miRNA in different cells and tissues. We used miR-199a as a model in order to elucidate the mechanism(s)that underpins miRNA actions. Previous studies showed that the promoters of both miR-199a-1 and -2 were hypermethylated resulting in silencing the expression of miR-199a in testicular germ cell tumors (TGCTs). In addition, transcription factor REST was observed to bind to the promoter of both miR-199a genes, raising the possibility that it might also be involved in controlling the expression of the miRNA in TGCTs. Contrary to the observations in TGCTs, miR-199a-2 was showed to have lower methylation levels in glioma patient samples comparing to the normal brain, while the methylation status of miR-199a-1 remained unchanged. The decreased methylation level of miR-199a-2 might explain the higher expression levels of miR-199a in glioma patients. In mesenchymal stem cells, expression of miR-199a was up-regulated during the differentiation process. This might be controlled by transcription factor Twist1. Previous studies showed that Twist1 binds to the promoter of DNM3OS, the intron of which contains the miR-199a-2 locus, and the elevated levels of miR-199a during stem cell differentiation may be due to the activating effect of Twist1 binding. Besides its upstream regulators, we also studied the activities of miR-199a by identifying its downstream targets. Observations indicated that the targets of miR-199a in different tissues are likely to be different resulting in different effects of the miRNA. By studying the mechanisms that control the expressions of miR-199a and its various downstream targets in different systems, we hope to illustrate the complexity of miRNA biology and indications of future directions for miRNA research.

3492T

Gold nanoparticles (AuNPs) cause variation in the level of methylation in H3K4 and H3K9. *A. Polverino[1], A. Iuliano[1,7], L. Cristino[1,4], R. Imperatore[2,4], A. Longo[5], G. Carotenuto[5], S. De Nicola[6], M. Piscopo[3], L. Fucci[3], E. Vitale[1].* 1) Cybernetics, CNR, Pozzuoli, Naples, Italy; 2) Institute of Biomolecular Chemistry CNR, Pozzuoli Naples, Italy; 3) Department of Structural and Functional Biology University of Naples, Naples Italy; 4) Endocannabinoid Research Group; 5) Institute of Biomaterials, CNR of Portici, Naples, Italy; 6) National Institute of Applied Optics (INOA) of the CNR Pozzuoli, Naples Italy; 7) Department of Institutional Studies Parthenope University Naples.

Histones are small basic proteins that play important roles in DNA packaging. A 10 nm DNA fiber is wrapped around a core of histones H2A, H2B, H3 and H4, forming a nucleosome, the first level of DNA condensation in chromatin. Histone methylation of specific lysine residues regulates the opening and closing of chromatin and has been proposed as a stable epigenetic marker that directs biological functions. Mitotically heritable epigenetic modifications have been reported to affect transcriptional regulation and to play key roles in human pathology, including inflammatory and neoplastic disorders. Here we characterize the effects of gold nanoparticles (AuNPs) and their uptake by human neuroblastoma cells SH-SY5Y in order to characterize their effects and optimize their use as potential drug delivery vehicles. We explored alterations in chromatin structure by analyzing the level of H3K4 and H3K9 methylation. We selected colloidal gold as our nanoparticle of choice, primarily because of its characteristic stability as well as its advantageous spectral properties. AuNPs have a diameter of approximately 2 nm and an absorption spectrum with a peak at 350 nm and their emission shifts from green to red upon forming clusters of approximately 200 nm in diameter. The clustered particles have a broad emission spectrum ranging from 400 to 600 nm. SH-SY5Y cells were incubated with AuNPs from 1 to 100 ng/ml for one hour, histones were extracted and dimethyl H3K4 and dimethyl H3K9 were quantitated by Western immunostaining. Histone H3 Lys 4 dimethylation positively correlates with transcription rate and Lys 9 dimethylation is associated with transcriptional repression. Cells treated with AuNP 100 ng/ml, 10 ng/ml and 5 ng/ml nanoparticles had a lower level of H3K4 dimethylation than controls (48%, 29,4% and 15%, respectively). Nanoparticle concentrations of 2 ng/ml and 1 ng/ml had no effect. Conversely, we observed a 50% increase of H3K9 dimethylation upon treatment with 100 ng/ml nanoparticles but no effect with 10 ng/ml and 5 ng/ml. Of note, treatment with 2 ng/ml and 1 ng/ml nanoparticles resulted in lower levels of H3K9 dimethylation (25,3% and 29,2%, respectively). Our data suggest an epigenetic modification of the genome, which correlates with a global gene expression decrease. However, cellular effects due to AuNPs appear to decrease proportionally to nanoparticle concentrations.

3493T

Gene methylation and allelic expression imbalance of the CYP1A2 and CYP3A4 genes in human livers. *J. Shi[1], W. Shou[1,2], Z. Niu[1], Y. Wang[1], K. Zhang[1], W. Huang[1,2].* 1) Department of Genetics, Shanghai-MOST Key Laboratory of Health and Disease Genomics, Chinese National Human Genome Center at Shanghai, Shanghai, 201203, China; 2) Rui Jin Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai 200025, China.

The CYPs play a critical role in the drug metabolism and contribute to the therapy outcome. CYP1A2 has an important function in the metabolism of caffeine as well as of several clinically important drugs. CYP3A4 are involved in the metabolism of 45–60% of all currently used drugs and a variety of other compounds. The large inter-individual difference in CYP1A2 and CYPP3A4 expression was well known. However, the underlying causes for these variations were not known. Epigenetic modifications and allelic expression imbalance have been found to influence the expression and activity of human CYP genes. Thus, we investigated the allelic expression imbalance and the methylation status of CYP1A2 and CYP3A4 genes in Chinese population. 25.0% of the rs2470890 heterozygote in CYP1A2 gene and 9.7% of the rs33972239 heterozygote in CYP3A4 gene showed allelic expression imbalance. The total hepatic CYP1A2 mRNA level correlated with the allelic ratio, the individual with a low ratio tended to have a low mRNA expression level. The lower extent of methylation associated to a higher CYP3A4 mRNA expression level was observed in four CpG sites located 25kb upstream of the translational start codon of CYP3A4. Our study suggested that the cis-acting variants might be involved in the regulation of CYP1A2 and the regulation of CYP1A2 might be different in populations, and provided more evidence of epigenetics playing a role in regulating the expression of CYP3A4. Our results present a new insight into the large inter-individual difference in CYP1A2 and CYP3A4 gene expression.

3494T

6q24 Transient neonatal diabetes mellitus: 16 years of data collection. *I. K. Temple[1,2], S. Kabwama[1], L. Doherty[1,3], S. Ennis[1], E. Cook[1], L. Harrison[2], S. Ellard[4], S. Flanagan[4], A. T. Hattersley[4], J. P. H. Shield[5], D. J. G. Mackay[1,3].* 1) Human Genetics and Genomic Medicine, Faculty of Medicine, University of Southampton, Southampton, Hampshire, United Kingdom; 2) Wessex Clinical Genetics Service, Princess Anne Hospital, University Hospital Southampton NHS Foundation Trust, UK; 3) Wessex Regional Genetics Laboratory, Salisbury District Hospital, Salisbury NHS Foundation Trust, Salisbury, UK; 4) Institute of Biomedical and Clinical Science, Peninsular College of Medicine and Dentistry, University of Exeter, Exeter UK; 5) Institute of Child Life and Health, University of Bristol, Bristol, UK.

Introduction: TNDM due to genetic aberrations at 6q24 is the commonest cause of diabetes presenting within the first week of life. TNDM predisposes to diabetes mellitus in later life. We report on the clinical presentation of 164 confirmed cases, the largest worldwide cohort of 6q24 TNDM cases. Data were collected from referring physicians by questionnaire completed at the time of referral. Results: The mean age of presentation was 8 days with a mode on day 1. On average, babies recovered by 4.5 months with a mode of 2 months. Birth weight had a normal distribution with a mean of 1997g but ranged from 3370g to 1050g with most babies born at term. Data analysis showed a positive correlation between age of presentation and both higher birth weight and longer gestation (p=0.04 and p = 0.007 respectively). 34% had paternally inherited duplications at 6q24, 37% had paternal uniparental disomy 6 (pUPD6) and 29 % had a maternal hypomethylation at 6q24 (44% isolated at 6q24, 46% with loss of methylation at multiple loci (28% due to ZFP57 mutations and 18% with hypomethylation at multiple loci of unknown cause)). No significant difference between the three genetic abnormalities (duplication, pUPD6 and methylation) was identified in terms of presentation of diabetes. However, analysis of variance (ANOVA) identified a significant difference (p=0.009) between the three genetic subgroups in terms of number of congenital abnormalities with babies with a duplication at 6q24 having a statistically lower number of congenital abnormalities compared to those with pUPD6 and hypomethylation (p=0.022 and p=0.020 respectively). For individual anomalies there were statistically significant differences for macroglossia and umbilical hernia. We also showed that patients with a loss of methylation at multiple imprinted loci (TND HIL (ZFP57/unknown)) had significantly more congenital abnormalities compared to patients with methylation loss at only 6q24(p=0.005). The IVF rate for babies with an epigenetic cause was 12% (2/17) compared with 0/46 in the other genetic subgroups. Conclusions: These results define the clinical presentation of 6q24 TNDM. The results also indicate a genotype phenotype correlation that patients with TNDM due to duplications at 6q24 have fewer associated congenital abnormalities. Furthermore congenital abnormalities are more common in those with loss of methylation affecting multiple imprinted loci than those with isolated hypomethylation at 6q24.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3495T

Genome-wide profiling of human uniparental tissues to search for novel imprinted genes. *S. Wen[1,2], W. Wan[1], A. Balasa[3], S. Mahadevan[4], Z. Liu[3,5], IB. Van den Veyver[1,2,3,4].* 1) OB/GYN, Baylor College of Medicine, Houston, TX; 2) Mol Hum Gen, Baylor College of Medicine, Houston, TX; 3) Pediatrics, Baylor College of Medicine, Houston, TX; 4) TBMM, Baylor College of Medicine, Houston, TX; 5) Neurology, Baylor College of Medicine, Houston, TX.

Purpose: Imprinted genes are preferentially/exclusively expressed from one parental allele. Many known imprinted genes are essential for embryonic and placental development. Their imprinting status is often associated with allele-specific CpG methylation of nearby differentially methylated regions (DMR). In this study we searched for novel imprinted genes by genome-wide DNA methylation profiling comparing DNA from androgenetic complete hydatidiform moles (AnCHM) and normal placenta. Methods: Genomic DNA was extracted from six AnCHM and six term placentas, bisulfite treated and hybridized to an Illumina Infinium Methyl27 array, which allows simultaneous characterization of the methylation status of 27,578 CpGs. For data analysis, we first examined a training data set of 129 CpGs located within 18 well-characterized imprinted DMRs. Based on behavior of these CpGs in AnCHMs and placentas, we defined criteria for candidate novel CpGs as follows: (1) p<0.01119; (2) beta values in placenta between 0.35 to 0.77; (3) beta-value difference between AnCHM and placenta is at least 0.15; (4) CpGs are located on autosomes. After identification of candidate imprinted genes, their allele-specific expression patterns were validated in four human trios by identifying informative expressed single-nucleotide polymorphisms (SNPs) in parental DNAs and assessing them for mono-allelic expression in cDNA of offspring placentas by RT-PCR. Results: We identified 592 candidate paternally methylated CpGs and 207 candidate maternally methylated CpGs. After defining a DMR as a genomic region smaller than 4kb, 576 candidate paternally imprinted genes and 142 candidate maternally imprinted genes remained. From these, we selected 24 genes for validation based on their p-value and a known or putative function relevant to human placenta. We identified mono-allelic expression of the SMPD3 gene, which encodes sphingomyelin sphingomyelinases 3. SMPD3 plays an important role in embryonic development and inactivation of Smpd3 in mice causes dwarfism and other developmental defects. Conclusion: Differential methylation profiling of AnCHM is an efficient method to identify imprinted loci in human placenta. We are continuing to validate other identified candidate imprinted loci in human trios.

3496T

Next generation bisulfite analysis around a 2.2 kb ICR1 deletion in 11p15.5 reveals variable hypermethylation explaining reduced penetrance in BWS families. *J. Beygo[1], V. Citro[2], D. Prawitt[3], M. Heitmann[1], K. Rademacher[1], B. Horsthemke[1], A. Riccio[2], K. Buiting[1].* 1) Institut für Humangenetik, Universitätsklinikum Essen, Essen, Germany; 2) Department of Environmental Science, Second University of Naples, Caserta 81100, Italy; 3) Zentrum für Kinder und Jugendmedizin, Universitätsmedizin Mainz, Mainz, Germany.

The congenital overgrowth disorder Beckwith-Wiedemann Syndrome (BWS) is caused by epigenetic or genomic alterations affecting one or both of the two imprinting control regions ICR1 or ICR2 on chromosome 11p15.5 which are normally methylated only on the paternal or the maternal allele, respectively. In rare cases BWS is caused by a maternally transmitted microdeletion affecting CTCF binding sites in the ICR1, which results in loss of IGF2 imprinting and BWS. Using MLPA (multiplex ligation-dependent probe amplification) we identified a microdeletion in peripheral blood from a female patient presenting with classical features of BWS. The deletion was also found in the healthy mother, maternal aunt and maternal grandfather of the patient. Interestingly, sequence analysis of the deletion breakpoints revealed that the 2.2 kb microdeletion is identical on the nucleotide level to a familial case reported by Prawitt *et al.*(2005). The deletion abolishes the CTCF binding sites 4 and 5 and fuses CTCF3 and 6 within the repeats B5 and B1. In order to find out how this 2.2 kb deletion and other deletions of the ICR1, which have been reported before (Sparago *et al.* 2004 and 2007; Cerrato *et al.* 2005; De Crescenzo *et al.* 2011), affect the DNA methylation in affected and unaffected deletion carriers we performed high resolution methylation analysis using next generation bisulfite sequencing on the 454 GS Junior System in six families to obtain highly quantitative DNA methylation. In three families with a 1.8 kb deletion and one family with a 1.4 kb deletion, which all show complete penetrance, we observed strong hypermethylation of adjacent CpGs and CTCF binding sites after maternal transmission of the deletion. In contrast, affected and unaffected family members with a maternally inherited 2.2 kb deletion had slight or no hypermethylation, respectively. These studies suggest that the 2.2 kb deletion can, but must not lead to a gain of methylation on the maternal allele, thus explaining the reduced penetrance in these two families.

3497T

A noncoding imprinted RNA, MESTIT1 is essential for the repression in cis of KLF14. *S. Horike, M. Meguro-Horike.* Advanced Science Research Center, Kanazawa University, Kanazawa, Ishikawa, Japan.

Imprinted genes on human chromosome 7 are thought to be involved in Russell-Silver syndrome (RSS). However, involvement of the known imprinted genes in RSS has yet to be established. Interestingly, previous study identified a noncoding imprinted RNA, *MESTIT1*, which is located intronic of *MEST* at 7q32. This gene is transcribed from the paternal allele only in the opposite direction to the *MEST* gene. However, the function of this noncoding imprinted RNA, *MESTIT1*, has remained unclear. In this study, we investigated some of its basic feature. We found that *MESTIT1* was exclusively localized in the nucleus, like other nuclear-retained RNAs, *45S* precursor rRNAs, imprinted ncRNA *Airn*. In addition to the basic analyses of *MESTIT1*, functional testing would be very important to determine whether the *MESTIT1* can be crucial imprinting elements and act in cis throughout the 7q32 imprinted cluster to silence protein-coding genes. Towards the dissection of the molecular basis of imprinted gene clusters, we have previously established monochromosomal hybrids containing individual human chromosomes of defined parental origin. In our study, a human chromosome 7 in the monochromosomal hybrids was tagged with pSV2*neo*, so that the marked chromosome could be successfully transferred into mouse A9 cells. Furthermore, the paternal or maternal copies of human chromosome 7 were transferred independently from A9 cells into DT40 cells by microcell fusion. DT40 cells exhibit high frequency homologous recombination between exogenous DNA templates and their chromosomal counterparts. This enables the efficient targeted modification of individual human chromosomes, too. In this study, we created modified human chromosome 7 carrying a targeted deletion of the *MESTIT*1 with the use of DT40 cells. We showed that deletion of the *MESTIT1*on the paternal chromosome results in the de-repression in cis of the normally silent paternal allele of *KLF14.* Our finding provided an important aspect for a detailed understanding of the mechanisms of long-range control gene activity within chromosomal domain.

3498T

MeCP2 is required for chromatin higher-order structure and dynamics at the imprinted 15q11-q13 locus. *M. Meguro-Horike[1], D.H. Yasui[2], W. Powell[2], J.M. LaSalle[2], S. Horike[1].* 1) Advanced Science Research Center, Kanazawa University, Kanazawa, Ishikawa, Japan; 2) Dep. of Medical Microbiology and Immunology, UC Davis, USA.

Deletions of 15q11-q13 on the paternal chromosome 15 cause Prader-Willi syndrome, whereas maternal deletions cause Angelman syndrome. Parental differences in DNA methylation at PWS-IC have a crucial role in genomic imprinting of 15q11-q13, as well as differences in DNA replication timing, histone modifications, chromosome nuclear organization, and mitotic recombination frequencies. However, despite progress in molecular characterization of 15q11-q13 rearrangements and imprinting mechanisms, the molecular pathogenesis of the autism phenotype resulting from maternal 15q11-q13 duplications remains largely unknown. Currently, no cell culture model exists for maternal 15q11-q13 duplication observed in autism cases, so our microcell-mediated transfers of a maternal chromosome 15 into human SH-SY5Y neuronal cells are expected to be a novel and essential experimental system for further understanding 15q11-q13 epigenetics in autism. Our 15q duplication neuronal model revealed that by quantitative RT-PCR transcript levels of *NDN*, *SNRPN*, *GABRB3*, and *CHRNA7* were reduced compared to expected levels despite having no detectable alteration in promoter DNA methylation. Since 15q11-q13 alleles have been previously shown to exhibit homologous pairing in mature human neurons, we assessed homologous pairing of 15q11-q13 by fluorescence in situ hybridization (FISH). Homologous pairing of 15q11-q13 was significantly disrupted by 15q duplication. To further understand the extent and mechanism of 15q11-q13 homologous pairing, we mapped the minimal region of homologous pairing to a ~500 kb region at the 3' end of*GABRB3* which contains multiple binding sites for chromatin regulators MeCP2 and CTCF as well as long-range interactions with the imprinting control region. Using RNA interference, we show that MeCP2 and CTCF are required for the homologous pairing of 15q11-q13 during neuronal maturational differentiation. In addition to the 15q11-q13 homologous trans interactions, we demonstrate that parental allele-specific chromatin structure at 15q11-q13 are regulated by PWS-IC and methyl-CpG binding proteins such as MeCP2. These data support a model where 15q11-q13 genes are regulated epigenetically at the level of both inter- and intra-chromosomal associations and that chromosome imbalance disrupts the epigenetic regulation of genes in 15q11-q13.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3499T

DNA methylation analysis of reciprocal genome-wide UPDs to define imprinted differentially methylated regions in the human genome. *K. Nakabayashi[1], A.M. Trujillo[2], C. Tayama[1], V. Romanelli[2], P. Lapunzina[3], M. Kagami[4], H. Soejima[5], H. Ogata[6], F. Court[2], D. Monk[2], K. Hata[1].* 1) Dept Maternal-Fetal Biol, Natl Res Inst for Child Health and Development, Tokyo, Japan; 2) Cancer Epigenetic and Biology Program, Institut d'Investigació Biomedica de Bellvitge, Barcelona, Spain; 3) Instituto de Genética Médica y Molecular, Universidad Autónoma de Madrid, Madrid, Spain; 4) Dept of Mol Endocrinol, Natl Res Inst for Child Health and Development, Tokyo, Japan; 5) Div of Mol Genet & Epigenet, Facult of Med, Saga Univ, Saga, Japan; 6) Dept of Pediatrics, Hamamatsu Univ, School of Med, Hamamatsu, Japan.

Genomic imprinting, an epigenetic phenomenon whereby genes are differentially expressed according to their parental origin, is shown to be crucial for placental development and fetal growth in mammals including humans. Most of well-characterized imprinted loci have been found to contain a germline differentially methylated region (DMR) harboring allelic DNA methylation inherited from the male or the female gamete. Several germline DMRs have been shown to govern imprinted gene expression and methylation of post-zygotic (secondary) DMRs as a cis-acting imprinting-control region (ICR). Epimutation at ICRs in humans has been shown to cause a number of imprinting disorders. We and others have previously identified rare mosaic genome-wide paternal and maternal uniparental disomy (UPD) cases among cohorts diagnosed to be Beckwith-Wiedemann and Silver-Russell syndromes, respectively. Profiling and comparison of the DNA methylation patterns of such UPD cases successfully detected the majority of known DMRs and also lead to identify novel DMRs such as FAM50B-DMR and ZFP597-DMR. In this study, we subjected the blood genomic DNA of genome-wide paternal and maternal UPD cases to illumina 450K methylation BeadChip analysis to further define the regions of known DMRs in the human genome. Among over 45,000 probes, we identified 382 probes for 32 DMRs in 20 known imprinted loci. This information will facilitate to further elucidate the genomic and epigenomic signatures of DMRs as well as to characterize the extent of epimutation(s) at DMRs in imprinting disorder cases.

3500T

The imprinted *C15orf2* gene in the Prader-Willi syndrome region encodes a nuclear pore complex associated protein. *L.C. Neumann[1], Y. Markaki[2], E. Mladenov[3], D. Hoffmann[4], K. Buiting[1], B. Horsthemke[1].* 1) Institut für Humangenetik, Universitätsklinikum Essen, Essen, Germany; 2) LMU München, Biozentrum Martinsried, Department Biologie II, Martinsried, Germany; 3) Institut für Strahlenbiologie, Universitätsklinikum Essen, Essen, Germany; 4) Abteilung für Bioinformatik, Zentrum für Medizinische Biotechnologie, Universität Duisburg-Essen, Essen, Germany.

The Prader-Willi syndrome (PWS) region in 15q11q13 harbors a cluster of imprinted genes expressed from the paternal chromosome only. Whereas loss of function of the *SNORD116*genes appears to be responsible for the major features of PWS, the role of the other genes is less clear. One of these genes is *C15orf2*, which has no orthologs in rodents but appears to be under strong positive selection in primates.

C15orf2 encodes an 1156 amino acids protein with six nuclear localization sequences. To find out more about the function of C15orf2, we used protein BLAST analysis and the InterProScan signature recognition search. With both tools we found a highly significant similarity of C15orf2 to the nuclear pore complex (NPC) protein POM121. Using the Invitrogen Flp-In system we generated a stable cell line that expresses FLAG-tagged C15orf2 upon doxycycline induction. By immunocytochemistry in expression-induced cells we found C15orf2 located at the nuclear periphery, where it co-localized with NPCs and nuclear lamins. At very high expression levels we observed invaginations of the nuclear envelope. Extending these observations to three-dimensional structured illumination microscopy (3D-SIM), which achieves an eight-fold improved volumetric resolution over conventional imaging and can resolve single NPCs, we saw that C15orf2 is located at the inner face of the nuclear envelope where it strongly associates with the NPC. Additionally, in nuclear envelope isolation and fractionation experiments we detected C15orf2 in the NPC and lamina fractions.

These experiments for the first time demonstrate that C15orf2 is part of the NPC or its associated molecular networks. After consultation with the Human Genome Nomenclature Committee we propose to assign a new name to *C15orf2*: *Nuclear pore associated protein 1* (*NPAP1*).

3501T

The `Language Gene' FOXP2 is not Imprinted. *A.C. Thomas, J.M. Frost, F. Khadem, P. Stanier, G.E. Moore.* University College London, Institute of Child Health, London, United Kingdom.

The Forkhead-box protein P2 (FOXP2) was identified as the first gene to be linked to an inherited form of language and speech disorder, namely developmental verbal dyspraxia (DVD) in the three generational `KE' family. DVD causes impaired selection and sequencing of coordinated orofacial movements that are required for speech and individuals show deficits in linguistic and grammatical processing. FOXP2 is a transcriptional repressor that is widely expressed in the fetal and adult brain (as well as other tissues) where it is thought to regulate the expression of genes in the cortical, basal ganglia and cerebellar circuits. We wanted to test the theory that FOXP2 may be an imprinted gene. It's location on chromosome 7 (7q31) makes it within fairly close proximity to a known imprinted region and additionally there is evidence for a differential parent-of-origin expression effect with regard to FOXP2, although recent genetic evidence has come to light that conflicts with this data. FOXP2 is a highly conserved protein and so finding an informative SNP for imprinting analysis was challenging. Initially we tried to perform imprinting analysis using the heterozygous point mutation found in members of the KE family (R553H) but FOXP2 is not expressed in blood so we were unable to do so. Subsequently we carried out imprinting analysis in brain and liver using rs12705977 in the 3' Untranslated Region of FOXP2. Methods used involved sequencing of genomic DNA to identify individuals informative (heterozygous) for this SNP followed by RNA extraction, DNase treatment, Reverse Transcriptase-PCR and subsequent sequencing of the corresponding tissues. In all cases FOXP2 was bi-allelically expressed therefore verifying that it is not imprinted in these tissues. This finding would imply that the origin of speech is too important to have evolved through imprinting or through haplo-insufficiency.

3502T

Evaluation of gene expression profile of chromatin modification enzymes in ACP02 and ACP03 gastric cell lines. *J.C. Santos[1,2], R.R. Burbano[3], M.L. Ribeiro[1,2].* 1) Sao Francisco University Medical School, Braganca Paulista, Sao Paulo, Brazil; 2) Universidade Estadual de Campinas, Genetics and Molecular Biology Post Graduation Program; 3) Human Cytogenetics Laboratory, Institute of Biological Sciences, Federal University of Pará, Belém-PA,Brazil.

Gastric cancer is one of the leading causes of cancer-related deaths worldwide. Two main gastric cancer histotypes, intestinal and diffuse, are recognized. Although most of the described genetic alterations have been observed in both types, it has been proposed that these two tumor types might result from different epigenetic pathways. Regarding this subject, in the recent years, much attention has focused on modifications of chromatin because their critical role in regulation gene expression and their active involvement in a number of cellular processes. Additionally, it is well known that epigenetic alterations such as aberrant expression of histone-modifying enzymes have been implicated in tumorigenesis. Since the knowledge of the events that lead to these alterations in cancer cells in vivo is still limited, the study of cancer cell lines is useful in solving this problem. Thus the aim of this work was to evaluate the mRNA expression profile of 84 chromatin modification enzymes in two Brazilian cell lines obtained from primary gastric adenocarcinoma. The cell lines were previously obtained from two patients with gastric cancer, including one diffuse type (ACP02) and one intestinal type (ACP03). After cellular growth (48h), the samples were evaluated by real-time PCR using the Human Epigenetic Chromatin Modification Enzymes RT^2 Profiler™ PCR Array. The expression was normalized with mRNA obtained from a normal tissue. Our data revealed that 26% of the genes were up-regulated and 24% down-regulated, in both ACP02 and ACP03. The up-regulation was exclusively observed among 55% of histone methyltransferases. On the other hand, the down-regulation was found among 46% of histone deacetilases, 43% histone phosphorylation and 34% of histone demethylases. The data from the present study contribute to characterize these new cell lines, and also indicate both shared the same pattern of expression of chromatin modification enzymes.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3503T

Expression patterns of histone acetylation modifier genes in gastric cancer. *F. Wisnieski[1], D.Q. Calcagno[1], M.F. Leal[1], T.B. Pontes[1], C.O. Gigek[1], E.S. Chen[1], S. Demachki[2], L.G. Lourenço[3], P.P. Assumpção[2], R. Artigiani[4], R.R. Burbano[5], M.A.C Smith[1].* 1) Department of Morphology and Genetics, UNIFESP, São Paulo, São Paulo, Brazil; 2) Service Surgery, University Hospital João de Barros Barreto, Belém, Pará, Brazil; 3) Department of Surgical Gastroenterology, UNIFESP, São Paulo, São Paulo, Brazil; 4) Department of Pathology, UNIFESP, São Paulo, Brazil; 5) Laboratory of Human Cytogenetics, UFPA, Belém, Pará, Brazil.

Gastric cancer is the fourth most prevalent cancer and the second cause of cancer-related deaths worldwide. Brazil has the fourth highest rate of gastric cancer in Latin America. The degree of acetylation of specific lysine residues in the tails of nucleosomal core histones plays an important role in the regulation of gene expression through chromatin structure. The acetylation of histone residues by histone acetyltransferases (HATs) leads to increased gene activity, whereas deacetylation by histone deacetylases (HDACs) results in chromatin condensation, making gene sequences inaccessible to transcription factors. Alterations in the pattern of histone modification can affect genome structure and integrity and, therefore, disrupt normal patterns of gene expression. Moreover, such histone alterations could be causal factors in cancer. The aims of this study were to compare the expression of two classes of chromatin-modifier genes (HDACs - *HDAC1*, *HDAC2*, *HDAC3* and HATs - *GCN5* and *PCAF*) between primary gastric adenocarcinoma and adjacent non-tumor tissues and to identify possible associations between the expression of these genes and clinicopathological features. Gene expression was evaluated by RT-qPCR in 38 gastric tumor and the matched non-tumor tissues from individuals of Brazil. Before evaluate HDACs and HATs expression, four bioinformatics tools (NormFinder, geNorm, BestKeeper and DataAssist) were used to assess the suitability of five possible reference genes (*ACTB*, *B2M*, *GAPDH*, *18SrRNA* and *RPL29*) for gastric samples. The best combination of reference genes was *GAPDH-B2M*. *HDAC1*, *HDAC3* and *GCN5*expression did not differ between tumor and non-tumor gastric tissues. *HDAC2* expression was significantly higher in tumor tissue compared to adjacent non-tumor tissue (p=0.04), whereas*PCAF* expression was significantly lower in tumor tissue compared to adjacent non-tumor tissue (p=0.01). In addition, *HDAC2* expression was higher in intestinal-type than diffuse-type gastric cancer (p=0.04). These findings suggest that *HDAC2* and *PCAF* may be possible biomarkers for the diagnosis of gastric cancer or useful targets for gene therapy.

3504T

The enzyme holocarboxylase synthetase mediates biotin-independent gene-silencing through the recruitment of histone deacetilases in the nucleus of human cells. *I. Trujillo, T. Barrios-García, S. Reyes-Carmona, I. Meneses-Morales, A. Leon-Del-Rio.* Biologia Molecular, Universidad Nacional Autonoma de Mexico, Mexico, Mexico City, Mexico.

The enzyme Holocarboxylase synthetase (HCS) plays a major role in metabolism catalyzing the biotinylation and activation of five biotin-dependent carboxylases in human cells. We demonstrated that HCS is also an obligate participant in biotin-mediated transcription control of multiple genes through the activation of a cGMP-dependent signal transduction pathway. The finding of HCS in the cell nucleus lead to suggest it could control gene expression through a mechanism based on histone biotinylation. However, this hypothesis has been recently challenged by the finding that biotin is not a natural histone modification. Using polytene chromosomes of D. melanogaster as a model we found HCS is a heterochromatin-associated protein that co-localizes with the transcription repressive epigenetic mark H3K9met3. In this work we explore the hypothesis that nuclear HCS is involved in gene silencing by determining the transactivation effect of a GAL4-HCS hybrid on the expression of a luciferase reporter gene in human HepG2 cells. Our results show that GAL4-HCS acts as a transcriptional repressor whose activity can be reversed by the non-specific HDAC inhibitor trichostatin A (TSA). To test the hypothesis that HCS could mediate transcriptional repression by interacting with HDACs we performed pull-down assays with recombinant GST-HCS and 35S-Met-HDACs. The results showed that HCS can interact with different histone deacetilases in vitro. Co-inmunoprecipitation of HepG2 nuclear proteins using an antibody anti-HCS allowed us to demonstrate that this enzyme is associated to HDACs in the nucleus of human cells in vivo. To explore if the role of HCS as a repressor of transcription requires of its biotinylation activity we introduced amino acid sequence changes in the active site of the enzyme, mimicking the mutations found in patients with severe HCS deficiency. The mutant HCS proteins promoted transcriptional repression of the reporter gene as efficiently as the wild type HCS. We propose that nuclear HCS mediates events of transcriptional repression through a novel biotin-independent mechanism which involves the recruitment of HDACs to the promoter of target genes.

3505T

Epigenetic Variation Among Humans. *M. Kasowski[1], F. Grubert[1], S. Kyriazopoulou-Panagiotopoulou[1], A. Kundaje[2], J. Li[1], D. Spacek[1], M. Snyder[1].* 1) Stanford University, Stanford, CA; 2) MIT, Cambridge, MA.

We previously showed that variation in transcription factor binding is an important mechanism leading to global differences in gene expression among humans. It is currently unknown to what extent differences in histone marks also contribute to expression variation between people. Histone proteins possess a variety of covalent modifications (e.g., methylation and acetylation) that correlate with the activation state of genes and functions of noncoding elements. Epigenetic modifications influence expression in part by altering chromatin structure, including accessibility of enhancers to transcription factors. Therefore variations in these marks may give rise to variations in transcription factor binding and gene expression. We are investigating global variation in epigenetic marks among individuals, its genetic basis, and influence on expression. Specifically, we have employed ChIP-Seq to map histone marks associated with a variety of functional elements genome wide in 12 HapMap individuals, and are correlating variations in these marks with transcription factor binding and RNA-Seq data. This study is expected to provide fundamental insights into the role of epigenetic variation in human phenotypic diversity.

3506T

Expression of mouse Lin28 gene is epigenetically regulated by histone modification. *A. Pang, A. Title, O.M. Rennert.* Laboratory of Clinical and Developmental Genomics, NICHD/NIH, Bethesda, MD.

The RNA-binding protein Lin28 is considered to be a “stemness” marker. It has a positive regulatory role on cell proliferation, which is partly mediated by the enhancement of translation of genes involved in cell growth. Additionally it blocks the maturation of microRNA let-7 that functions in tumor suppression. The expression of Lin28 gene predominantly occurs in pluripotent cells and cancer cells. Nevertheless, the regulation of Lin28 gene expression has not been thoroughly studied. We hypothesized the cell-limited expression of Lin28 is regulated epigenetically. We examined whether the absence of Lin28 expression in mouse cells resulted from the methylation of Lin28 promoter DNA. Bisulfite-sequencing analysis did not reveal a correlation between its promoter DNA methylation and expression of Lin28 gene. Consistent with this observation, Lin28 expression was not reactivated in non-Lin28 expressing cells, such as NIH/3T3 cells, after treatment with 5-aza-2'-deoxycytidine. We next examined if Lin28 expression was regulated by histone acetylation. Treatment of NIH/3T3 cells with the histone deacetylase inhibitor trichostatin A (TSA) led to the reactivation of Lin28 expresssion. In addition, the promoter region of Lin28 gene became more sensitive to nuclease digestion after TSA treatment, suggesting that Lin28 expression was mediated by the relaxation of chromatin. In agreement with these findings, we observed a significant enrichment of H3K9ac and reduced occupancy of H3K9me2 at the Lin28 promoter region in TSA-treated NIH/3T3 cells and in undifferentiated P19 embryonal carcinoma cells. The opposite pattern of histone modification was found in untreated NIH/3T3 cells and differentiated P19 cells. Taken together, our results identify that Lin28 expression in mouse cells is epigenetically regulated by histone modification.

3507T

Transcriptional and epigenetic variation in human induced pluripotent stem cells. *N. Kumasaka[1], F. Rouhani[1], L. Vallier[2], A. Bradley[1], D. Gaffney[1].* 1) Wellcome Trust Sanger Institute, Cambridge, United Kingdom; 2) Stem Cell Institute, University of Cambridge.

Human induced pluripotent stem cells (hIPSCs) are a potentially powerful system for developing cellular models of human disease. However, significant epigenetic and transcriptional differences may occur between hIPSCs derived from different tissues and different individuals. Identifying the sources of this variation, and their relative importance in determining the transcriptional and epigenetic landscape of hIPSCs is important for understanding the biology of cellular reprogramming, and for informing efforts to improve the quality and differentiation potential of hIPSC lines. Here we perform RNA-seq, ChIP-seq for a panel of 4 histone modifications and FAIRE-seq on hIPSCs derived from skin fibroblasts, keratinocytes and endothelial precursor cells across a number of different, healthy donor individuals and from a human embryonic cell line (H9ESC) for comparison. We compared epigenetic and transcriptional variability between and within tissues and individuals using a novel statistical framework based on Bayesian hierarchical negative-binomial model. Our approach allows us to estimate over-dispersion for each transcript using the entire data set while adjusting for known experimental biases, such as GC content, as covariates. Our results suggest that, although a small number of genes are consistently differentially expressed between hIPSCs and hESCs, very few of these are specific to their tissue of origin. We suggest that transcriptional memory of the progenitor cell type in hIPSCs is a relatively rare phenomenon.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3508T

The ratio of *IGF2/IGF2R* in chorionic villus samples as a biomarker for birth weight. *C. Demetriou[1, 2], A. Thomas[1], M. Ishida[1], S. Abu-Amero[1], R. Aggarwal[3], D. Peebles[3], A. Syngelaki[4], K. Nicolaides[4], L. Regan[2], G.E. Moore[1].* 1) Clinical and Molecular Genetics Unit, Institute of Child Health, University College London, UK; 2) Department of Obstetrics and Gynaecology, St. Mary's Campus, Imperial College London, UK; 3) Institute for Womens Health, University College London, UK; 4) Harris Birthright Research Centre for Fetal Medicine, King's College Hospital, London, UK.

Fetal growth and development is a complex process driven by both genetic and environmental factors. Although many genes are involved, of particular interest are imprinted genes that are monoallelically expressed in a parent-of-origin fashion. Insulin-like growth factor 2 (*Igf2*) and one of its receptors, (*Igf2r*) are examples of direct interaction between paternal and maternal imprinted genes in mice involved in growth regulation. *Igf2* promotes growth enhancement while *Igf2r* sequesters IGF2 and therefore restricts the growth promoting actions of IGF2. In this study, the expression levels of *IGF2* and *IGF2R* were investigated using real-time PCR in 260 human chorionic villus samples (CVSs) obtained at 11–13 weeks of gestation. The expression levels of *IGF2*, *IGF2R* and *IGF2/IGF2R* ratio were correlated with fetal growth parameters such as birth weight (BW) and crown-rump length (CRL). Expression of *IGF2* and *IGF2R* was standardised against the *L19* housekeeping gene and fitted to appropriate regression models. The analysis of the data was performed correcting for confounding factors such as gender, maternal weight and gestational age. *IGF2* CVS expression showed a significant positive correlation with both BW (p=0.004) and CRL (p=0.007). A similar positive association was observed between the expression levels of *IGF2R* and BW (p=0.006) and CRL (p=0.03). Babies that were small for gestational age (N=50) had a significantly lower *IGF2* expression levels compared to normal babies (N=145) (p=1.1e-07) and a significantly lower *IGF2/IGF2R* ratio (p=3.8e-09). This suggests that the low IGF2 levels could be causing the small size at 11–13 weeks and is predictive of <10% birth weight babies at term. Babies that were large for gestational age (N=65) had significantly higher *IGF2R* expression levels compared to normal babies (p=0.02), suggesting that these babies are releasing more IGF2R in order to remove IGF2 and hence avoid further increase in size. This is the first study that has looked at the expression of *IGF2* and*IGF2R* in human CVS and the first report of a positive correlation with growth parameters such as BW and CRL.

3509T

Genome Wide Studies of Methylation in the Mouse Frontal Cortex Reveals Novel Imprinted Differential Methylated Regions and Non-CG Methylation. *C. Barr[1,2], W. Xie[3], Y. Feng[1], K. Wigg[1], E. Dempster[1,4], L. Gomez[1], J. Eubanks[1], B. Ren[3].* 1) Genes & Development, Toronto Western Hosp, UHN, Toronto, ON, Canada; 2) Programme in Neurosciences and Mental Health, The Hospital for Sick Children, Toronto, ON, Canada; 3) Ludwig Institute for Cancer Research, University of California San Diego, CA; 4) Institute of Psychiatry, Kings College London, London UK.

Current data support approximately 90 validated imprinted genes in mice, but a recent study indicated that this number may be well over 1,000. Many of the genes are imprinted in a tissue/cell type specific pattern and little is known of these genes. Previous evidence indicates that imprinted genes influence the development of specific brain regions and we sought to identify imprinted genes contributing to the development of the frontal cortex. We created crosses and reverse crosses of mouse strains to determine the parent-of-origin of expressed genes. We used next generation sequencing of RNA (RNA-seq), chromatin immunoprecipitation to histone modifications (ChiP-seq), and bisulfite sequencing (MethylC-Seq) of DNA from the frontal cortex of the adult F1 mice to identity transcripts and epigenetic markers with parent-of-origin effects. Using a FDR of .01, and requiring that the same parent-of-origin bias was evident in 6 biological replicates, we identified known imprinted genes as well as novel patterns of expression for known imprinted genes. We also identified 109 novel transcripts with parent-of-origin effects. Using Methyl-seq we correctly identified known differentially methylated regions (DMRs) as well as 23 novel DMRs most within known imprinted regions. Interestingly, we found significant amounts of non-CG methylation in frontal cortex (35% of methylation), some showing evidence of imprinting. We confirmed the finding of non-CG methylation in 6 additional brain regions and spinal cord but not in other differentiated tissues. None of the novel imprinted transcripts that we identified using RNA-seq were marked by a DMR or by chromatin patterns that were indicative of imprinting. Further, pyrosequencing of a subset of the novel candidates did not replicate their imprinted expression. The results thus do not support the existence of many more imprinted genes than previously identified in frontal cortex and suggest a potential bias in RNA-seq data. The finding of significant amounts of non-CG methylation was surprising given that this epigenetic mark previously documented only in ES cells, preimplantation embryos and oocytes, and was thought to be a marker of pluripotency. The role of this epigenetic mark in human brain is currently under investigation.

3510T

An integrated analysis of DNA methylation, histone modifications and gene expression in monozygotic twins discordant for psoriasis identifies disease-associated genes. *R. Lyle[1], K. Gervin[1], G.D. Gilfillan[1], M. Hammerø[1], H.S. Hjorthaug[1], A.O. Olsen[2], T. Hughes[1], J.R. Harris[3], D.E. Undlien[1].* 1) Department of Medical Genetics, Oslo University Hospital, Oslo, Norway; 2) Department of Dermatology, Oslo University Hospital and University of Oslo, Oslo, Norway; 3) Division of Epidemiology, Norwegian Institute of Public Health, Oslo, Norway.

Psoriasis is a common, chronic inflammatory immune-mediated disease, which affects mainly the skin, but also the joints. The worldwide prevalence ranges between 1–11.8% depending on ethnicity and geographical area. Psoriasis has a strong genetic component with an estimated heritability of 66%. However, additional environmental and/or epigenetic factors are thought to be important, as concordance rates among MZ twins are only 35–72%. The aim of this study is to identify epigenetically dysregulated genes which are associated with psoriasis. We investigated the epigenetic component of psoriasis using high-throughput bisulfite sequencing (RRBS), ChIPseq (H3K4me3 and H3K27me3) and RNAseq to explore the extent of DNA methylation, histone modifications and gene expression differences in CD4+ cells isolated from 20 pairs of monozygotic (MZ) twins discordant for psoriasis. This approach enables an integrated analysis of the interplay between DNA methylation, histone modifications and gene expression and identifies disease-associated epigenetic patterns and dysregulated genes. This pinpoints potential susceptibility genes and pathways dysregulated in psoriasis. This is the first study to systematically integrate DNA methylation, histone modifications and gene expression in discordant MZ twins in order to reveal and understand the epigenetic component of disease.

3511T

Monozygotic male twins with loss of methylation at DMR2 but discordant for clinical features of Beckwith Wiedemann syndrome. *J. Lazier, M.A. Thomas.* Department of Medical Genetics, University of Calgary, Calgary, Alberta, Canada.

Beckwith Wiedemann Syndrome (BWS) is characterized by overgrowth, hemihyperplasia, neonatal hypoglycemia, abdominal wall defects, macroglossia, ear pitting, and an increased risk of embryonal tumours. Several cases of discordant monozygotic (MZ) twins for BWS have been reported, with a disproportionate number of female twins with a loss of methylation (LOM) at DMR2. Understanding the mechanisms underlying the phenotypic differences in MZ twins will help understand variability in this condition. We present the tenth known case of male monozygotic twins who are discordant for clinical features of BWS. Pregnancy was by natural conception with no complications. One twin was born larger than the other, with a weight of 2.72 Kg (<97%ile) and was noted to have a large umbilical cord. He was treated for hypoglycemia for days 3–5, with spontaneous recovery. He had macroglossia and a nevus flammeus on his forehead. He was later noted to have hemihyperplasia affecting his right arm. The other twin had a birthweight of 2.3kg (50%ile). He had a fainter nevus flammeus than his brother, but did not have neonatal hypoglycemia, macroglossia or hemihyperplasia. Genetic testing confirmed monozygosity and showed that both had LOM at DMR2 in their serum.

Several mechanisms have been proposed to account for the discordant phenotypes, including non-random X-inactivation in females, although this has been questioned. Differences in either new mutations or methylation changes in other genes likely play a role. Weksberg *et. al*, (2002) proposed that DMR2 LOM either predisposes to twinning and discordance, or discordance is caused by unequal splitting of the inner cell mass during a twinning event. DMR2 LOM has also been seen repeatedly in the blood of an unaffected twin, but usually not also in the buccal swab. This may be caused either by allo-transfusion (due to shared blood flow in placenta) or migration of affected blood cell precursors from the yolk sac to both twins (Bliek, 2009).

We will perform further investigation from their buccal swabs which may help provide new evidence into modifying factors affecting expression.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3512T

Acquired Epigenetic and Chromosomal Instability Alterations are Present in Adults Who Experienced Childhood Sexual Abuse: A Discordant Monozygotic (MZ) Twin Study. *T. York[1], J. Brumelle[2], J. Juusola[2], K. Kendler[3], E. Eaves[1,3], A. Amstadter[3], S. Aggen[3], K. Jones[4], S. Latendresse[3], A. Ferreira-Gonzalez[2], C. Jackson-Cook[1,2].* 1) Human and Molecular Genetics, Virginia Commonwealth Univ, Richmond, VA; 2) Pathology, Virginia Commonwealth Univ, Richmond, VA; 3) Psychiatry, Virginia Commonwealth Univ, Richmond, VA; 4) Neodiagnostix, Inc, Rockville, MD.

Childhood sexual abuse (CSA) is a traumatic event associated with an increased lifetime risk for developing psychiatric disorders, as well as several age-related health problems such as cardiovascular and respiratory disease. We hypothesize that the latent health consequences of CSA exposure are biologically mediated by acquired somatic epigenetic and chromosomal alterations that persist into adulthood. To test this hypothesis we collected blood specimens from female adult MZ twins (17 pairs, 34 individuals) who were DISCORDANT for a history of CSA. Acquired chromosomal changes were measured using the cytokinesis-block micronucleus (MN) assay on stimulated lymphocytes, which allows for scoring of numerical and structural acquired chromosomal anomalies. Epigenetic alterations were quantified by methylation profiles in bisulphite converted DNA using the Illumina Infinium Methylation 450K BeadChip, which interrogates over 480K genome-wide cytosines. Female MZ twins exposed to CSA (CSA+) exhibited a 1.63-fold average increase in their frequency of MN compared to their nonexposed cotwins (Paired t-test, t_{16}=2.65, P=0.017). This increase primarily reflected the impact of CSA exposure since no additional effects of familial or environmental factors were noted after controlling for the effect of CSA exposure. The observed co-twin divergence in MN frequencies increased with age, suggesting that the biological effects of CSA on adult MN formation may be cumulative. Extrapolating from normative twin data, the observed increase in MN in the CSA+ twins was calculated to be consistent with a 9.9 year biological age increase compared to their unabused MZ co-twins. Divergence in methylation profiles were also seen between the discordant identical co-twins. Acquired changes were detected for 465 sites (FDR of 15%). The majority of these differences (393 sites or 84.5%) were attributable to hypomethylation in the CSA+ twins, suggesting that these abused twins may have activation of normally inactivated genomic regions. The hypomethylated cytosines were preferentially located in sites outside of CpG islands, in CpG shores or “open sea” sequences (Binomial test P=0.063). These data support a direct link between CSA exposure and both epigenetic and chromosomal instability frequencies in adults and provide a foundational basis for future investigations to understand whether and how these DNA-based changes might mediate the association between CSA and adult disease outcomes.

3513T

Genome-wide DNA methylation and gene expression analyses of monozygotic twins discordant for intelligence levels. *CC. Yu[1], M. Furukawa[1], K. Kobayashi[1], C. Shikishima[2], PC. Cha[1], J. Sese[3], H. Sugawara[4], K. Iwamoto[5], T. Kato[4], J. Ando[6], T. Toda[1].* 1) Division of Neurology/Molecular Brain Science, Kobe University Graduate School of Medicine, Kobe, Japan; 2) Keio Advance Research Centers, Keio University, Tokyo, Japan; 3) Department of Computer Science, Graduate School of Information Science and Engineering, Tokyo Institute of Technology, Tokyo, Japan; 4) Laboratory for Molecular Dynamics of Mental Disorders, RIKEN Brain Science Institute, Saitama, Japan; 5) Department of Molecular Psychiatry, Graduate School of Medicine, the University of Tokyo, Tokyo, Japan; 6) Faculty of Letters, Keio University, Tokyo, Japan.

Human intelligence, as measured by intelligence quotient (IQ) tests, demonstrates one of the highest heritabilities among human quantitative traits. Nevertheless, studies to identify quantitative trait loci responsible for intelligence face challenges because of the small effect sizes of individual genes. Phenotypically discordant monozygotic (MZ) twins provide a feasible way to minimize the effects of irrelevant genetic and environmental factors, and should yield more interpretable results by finding epigenetic or gene expression differences between twins. Here we conducted array-based genome-wide DNA methylation and gene expression analyses using 17 pairs of healthy MZ twins discordant intelligently. ARHGAP18, related to Rho GTPase, was identified in pair-wise methylation status analysis and validated via direct bisulfite sequencing and quantitative RT-PCR. To perform expression profile analysis, gene set enrichment analysis (GSEA) between the groups of twins with higher IQ and their co-twins revealed up-regulated expression of several ribosome-related genes and DNA replication-related genes in the group with higher IQ. To focus more on individual pairs, we conducted pair-wise GSEA and leading edge analysis, which indicated up-regulated expression of several ion channel-related genes in twins with lower IQ, and up-regulated expression of IGF-1 in twins with higher IQ. Our findings implied that these groups of genes may be related to IQ and should shed light on the mechanism underlying human intelligence.

3514T

Analysis of expressed SNPs identifies new genes escaping X-chromosome inactivation and variable extents of expression from the inactive X. *A. Cotton[1], B. Ge[2], T. Pastinen[2], C. Brown[1].* 1) University of British Columbia, Department of Medical Genetics, Life Sciences Institute, Molecular Epigenetics Group, Vancouver, British Columbia, Canada; 2) McGill University and Genome Québec Innovation Centre, Department of Human and Medical Genetics, McGill University Health Centre, McGill University, Montréal, Québec, Canada.

X-chromosome inactivation (XCI) results in the transcriptional silencing of most, but not all, genes on one X chromosome. Genes which escape XCI are candidates to explain the phenotype of X-aneuploides and sex-specific disease susceptibilities. Cis-acting silencing is also observed for imprinted genes and recent genome-wide studies of expressed SNPs have revealed considerable allelic imbalance (AI), in which two alleles of the same gene show large differences in the level of expression. As the majority of X-linked genes are subject to XCI, these genes would be expected to show AI in females with non-random XCI which would in turn allow for the identification of genes subject to or escaping from XCI. X-linked AI was examined using the Illumina Human1M-duo BeadChip in 190 samples (99 female and 91 male) from three populations (CEPH, YRI and primary fibroblasts). By combining multiple probes across genes, a single AI score was calculated for 725 X-linked genes. In the female fibroblast samples, an average 97% of genes showed bi-allelic expression leading to the conclusion that these samples had random XCI. In 31 lymphoblast samples (27 CEPH and 4 YRI), more than half of the X-linked genes were found to have mono-allelic expression, suggesting XCI in these females was not random. Using only those females with substantial non-random XCI, AI scores were converted into an XCI status for 374 X-linked genes, 81% of which showed the same XCI status as previously established. Of the 52 genes with no previous data, 36 were demonstrated to be subject to XCI and 11 escaped from inactivation. An additional 5 novel genes were found to escape from X-chromosome inactivation in some females but were subject in others demonstrating a pattern of expression known as variable escape from XCI. Data from different cell types, different populations and families, allows investigation of potential sources of such variability. Interestingly, in the populations where all females showed random XCI, genes with mono-allelic expression were still detected. X-linked AI analysis has not only provided new data on which genes escape from XCI, but also direct information about the wide-range of expression level on the inactive X relative to the active X chromosome. Overall, the study of AI for X-linked genes provides the opportunity to increase our understanding of the silencing of one X chromosome during XCI, and thus how cis-acting silencing can contribute to variations in disease.

3515T

Investigating the role of the WICH chromatin remodeling complex in maintaining facultative heterochromatin at the human inactive X chromosome through targeted deletion of the BAZ1B gene using zinc finger nuclease technology. *A.E. Culver-Cochran, B.P. Chadwick.* Biological Science, Florida State University, Tallahassee, FL.

X chromosome inactivation (XCI) is the mammalian from of dosage compensation that balances the levels of X-linked gene expression between the sexes. Early in female development, transcription from one X is shut down as the chromosome is repackaged into facultative heterochromatin. Once established, the same X is maintained as the inactive X chromosome (Xi) throughout subsequent somatic cell divisions. While much is known about the initiation of XCI, our understanding of how the Xi chromatin is faithfully maintained through each cell cycle is limited. We have identified the multifunctional Williams syndrome transcription factor (WSTF) as a protein that transiently associates with the Xi as it is replicated in late S-phase, and, therefore, is a prime candidate for maintaining XCI. WSTF acts at the Xi as a component of the WSTF-ISWI chromatin remodeling complex (WICH). Interestingly, we have found that WICH and two other proteins implicated in heterochromatin maintenance, γ-H2A.X and BRCA1, sequentially associate with the Xi, suggesting that each plays a unique role in maintaining the chromosome's epigenetic signature. In order to assess the role of WICH at the Xi, we used zinc finger nuclease (ZFN) technology to disrupt the WSTF encoding BAZ1B gene on chromosome 7q11. A pair of active ZFNs were designed to exon-7 and were subsequently introduced into a female diploid hTERT-immortalized cell line along with a promoter-trap repair template that completely replaced the exon. Two clones were isolated that had successfully removed both alleles of exon-7, and three independent clones were identified that had replaced one allele and disrupted the open reading frame of the second allele by non-homologous end joining. We are actively investigating the impact of WSTF loss on the Xi by examining changes to chromatin structure and X-linked gene expression.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3516T

DNA methylation profiling in X;autosome translocations supports a role for repeat elements in the spread of X chromosome inactivation. *A. Sharp, N. Bala, M. Brahmachary, P. Garg, C. Borel.* Genetics & Genomics Sci, Mount Sinai Sch Med, New York, NY.

X chromosome inactivation (XCI) is an epigenetic mechanism that silences in cis the majority of genes on one of the two X chromosomes in females. However, a significant fraction of genes remain expressed from the inactive X and thus `escape' XCI. Several previous studies have shown differential densities of certain types of common repeat elements associated with genes subject to versus those that escaping XCI, suggesting that the spread of XCI might be regulated by these repeats. However, due to the non-random distribution of genes that escape XCI and the unusual evolutionary history of the X chromosome, it has been unclear if the correlation between repeat elements and escape from XCI is a functional one. In order to test the hypothesis that the spread of X chromosome inactivation shows sequence specificity, we have analyzed the spread of XCI into autosomal chromatin by performing DNA methylation profiling in six unbalanced X;autosome translocations using Illumina 450k beadchips. Using promoter hypermethylation as a robust epigenetic signature to measure the spread of XCI, we have determined the inactivation status of 1,189 autosomal genes after translocation onto the inactive X. Subsequent sequence analysis comparing the prevalence of common repeats and DNA motifs in translocated autosomal genes that are either subject to or escape the spread of XCI shows that the density of many repeat elements, including LINEs, Alus, ERVs and MIRs, is significantly correlated with local susceptibility to the XCI signal. This association between repeat density and silencing of autosomal genes by XCI recapitulates similar observations made on the X chromosome, suggesting a causal relationship between DNA sequence features such as common repeats and the spread of XCI. These data support Mary Lyon's `repeat hypothesis', suggesting a functional role for repetitive elements in the XCI process.

3517T

Tools for epigenetic research. *L. Apone, P. Liu, G. Lohman, E. Cantor, B. Langhorst, D. Munafo, C. Sumner, E. Yigit, L. Merrill, F. Stewart, T. Evans, E. Dimalanta, T. Davis.* New England Biolabs, Inc., Ipswich, MA.

Epigenetics is the study of heritable changes in gene expression that occur without a change in the primary DNA sequence. These changes, which include DNA methylation and histone protein modification, are reversible, responsive to environmental stimuli and add layers of complexity to the human genome. Epigenetic changes occur predominately through protein DNA interactions. Identifying and mapping these interactions is key to understanding how changes in the epigenome affect human health and disease. The chromatin immunoprecipitation assay (ChIP) is one of the most powerful tools available to map DNA protein interactions. When coupled with high throughput DNA sequencing (ChIP-Seq), genome wide, protein interaction maps can be generated. Although a powerful technique, the use of ChIP-Seq has been hampered by the difficulty in obtaining sufficient quantities of input DNA. To overcome this challenge we have developed a library construction method using novel NEBNext® reagents and adaptors with a low DNA input requirement. The method is quick, easy and compatible with multiple high throughput sequencing platforms. In addition, the use of NEBNext® multiplex oligos allows the sequencing of multiple ChIP-Seq samples in a single sequencing run. Taken together, the improvements to library construction, the versatility of platform choice and the ability to multiplex, will aid researchers in overcoming some of the hurdles to using ChIP-Seq in epigenetic research.

3518T

Sex-influenced epigenetic effects in chromosomal rearrangements. *J. Kapalanga[1,2,3], D. Wong[2,3], N. Nwebube[1], A. Gandy[2,3].* 1) Dept Pediatrics, Genetics, Grey Bruce Health Services, Owen Sound, ON, Canada; 2) Dept Pediatrics, Dalhousie University, Halifax, NS, Canada; 3) Summerside Medical Center, Summerside, PE, Canada.

Balanced chromosomal rearrangements can be associated with an abnormal phenotype; conceivably due to break-point involvement of loci and regions of the genome that are critical in gene or genomic function. Abnormal phenotypes in balanced translocation carriers have been attributed to microdeletions, microduplications, gene disruptions and position effects at the break points. The aforementioned mechanisms may not be the only explanation for abnormal phenotypes in chromosomal rearrangements. Epigenetic influences could be another explanation. We report the effect of sex on phenotype in individuals with familial chromosomal rearrangements in different regions of the genome. The chromosomal rearrangements, viz; 46, XY, inv(11)(p15.4q12).ish inv(11)(pter+qter+), t(14;17)(q32.32;qter) and 47, +mar(12) were observed in 3 families and across three generations. An analysis of clinical data resulted in the identification of 9 individuals with balanced chromosomal rearrangements or supernumerary marker chromosomes (SMC). Four of the 9 individuals had abnormal phenotypes and all 4 were male, while all 5 individuals with normal phenotypes were female. While studies have shown that there are more mothers than fathers in translocationor SMC families, no studies have reported a sex effect on microdeletions, microduplications, gene disruptions and position effects associated with balanced chromosomal rearrangements. We, therefore, hypothesize that sex-influenced epigenetic mechanisms are an explanation for the observed sex related differences in phenotypic effects of balanced chromosomal rearrangements. DNA methylation and histone modification studies are being undertaken to characterize the epigenetic mechanisms.

3519T

The relationship between paternal uniparental disomy and clinical features in patients with Beckwith-Wiedemann syndrome. *O. Yasufumi, K. Jozaki, T. Maeda, H. Yatsuki, K. Higashimoto, H. Soejima.* Division of Molecular Genetics and Epigenetics, Department of Biomolecular Sciences, Faculty of Medicine, Saga University, Saga, Japan.

(Background and purpose) Beckwith-Wiedemann syndrome (BWS) is an imprinting disorder characterized by macroglossia, macrosomia, abdominal wall defect and a predisposition to embryonal tumor in childhood. Approximately 20% of patients with BWS show the paternal uniparental disomy (UPD) of chromosome 11. The patients with UPD were mosaic of normal cells and UPD cells, leading to hypertmethylation at H19-DMR and hypomethylation at KvDMR within 11p15. 5 imprinted region. UPD of BWS has genetic diversity such as the various ratio of mosaicism and the various extent of UPD region. The aim of this study is to clarify the relationship between the genetic diversity and clinical manifestations in BWS patients caused by UPD. (Subjects and methods) We analyzed 25 of BWS patients with UPD (male 9, female 16). The mosaic ratio was evaluated by tetranucleotide repeat markers from 11p15.5 and calculated as follows: (k-1)/(k+1) × 100; where k is the ratio of the peak height of ratios of paternal allele to maternal allele. The extent of UPD was evaluated by Affymetrix® Human SNP array 6.0. MassARRAY analysis was performed to search the methylation of 35 imprinted differentially methylated regions (DMRs) scattered throughout genome. Based on these results, we statistically analyzed the relationship between the diversity of UPD and clinical information. (Result) The mosaic ratio ranged from 28% to 99%;, and was 60% or more in 16 cases. As for the extent of UPD, the recombination breakpoints existed within chromosome 11p in 19 cases, and spread into the 11q in two cases. The smallest UPD region was 2.7 Mb, including both H19-DMR and KvDMR. The largest UPD region was whole genome, i.e., genome-wide paternal UPD (GWU), which was detected in four patients. In the methylation analysis, the hypermethylation at H19-DMR and the hypomethylation at KvDMR were seen in all cases according to the mosaic ratio. In addition, the majority of other imprinted DMRs showed aberrant methylation in all GWU cases. There was no correlation between the mosaic ratio and the extent of UPD. We found negative correlations between the mosaic ratio and fetal-placental weight ratio, and between the extent of UPD and gestational age. Furthermore, an increased incidence of tumor development correlated with the extent of UPD. (Conclusion) The genetic diversity of UPD may affect the BWS phenotypes including gestational age, placentation, and tumor development.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3520T

Streamlined chromosome conformation capture workflow for the studies of human β-globin gene control by locus control region. *C. Woo, C. Liu, S. Dong.* Life Technologies, 850 Lincoln Centre Drive, Foster City, CA.

The human genome is organized non-randomly in the nucleus. The three-dimensional chromosome structure and interaction represent the final outcome of genetic and epigenetic interplays and provide the basis for gene regulation. We developed a streamlined workflow to convert such physical interactions between genomic loci to quantifiable DNA sequences by real time PCR. The improved workflow has shortened sample to readout time from five to two days, yielding an average of 22ug of 3C library from 5×10e6 cells. Real time PCR also provides interaction profiles in high resolution. The human β-globin gene family on Chromosome 11, which consists of five isoforms, is expressed sequentially during development, as enhanced by the locus control region (LCR) 60kb away. Using the improved workflow, we measured the detailed interactions between LCR region and β-globin isoforms and correlated the interaction with gene expression levels in two cell lines: a β-globin expressing erythroleukemia cell line (K562) and non-expressing lymphoblast cell line (GM06990) in relation to LCR conformation. The results confirmed the different LCR interaction in expressing and non-expressing cell lines. Quantitative interaction analysis of the data further refined the enhancer functioning domain and revealed the strong correlation of high expression with strong interaction with LCR, indicating that β-globin gene expression was controlled by the looping interaction between the LCR enhancer and the target genes.

3521T

Micro-RNAs in the Sclera: Role in Ocular Growth Regulation and Implications for Myopia. *R. Metlapally[1], P. Gonzalez[2], F.A. Hawthorne[3], K. Tran-Viet[3], C.F. Wildsoet[1], T.L. Young[2,3].* 1) School of Optometry, University of California at Berkeley, CA; 2) Duke Eye Center, Durham, NC; 3) Duke Center for Human Genetics, Durham, NC.

Purpose: The enhanced ocular enlargement underlying myopia is driven by a signaling cascade that originates in the retina and ultimately determined by scleral (outer coat) extracellular matrix remodeling. miRNAs regulate gene expression by pairing with the 3' UTR of target sequences and serve as nodes of signaling networks. The role of miRNAs in scleral remodelling is unknown. We hypothesized that the sclera, like most tissues, expresses miRNAs and that some play active roles in eye growth regulation. **Methods:** Scleral samples were obtained from normal human fetal donor eyes (24 wk), representing rapidly growing eyes, as well as adult (age-matched) donor eyes (n=6, each group). RNA was extracted using the miRVANA kit and genome-wide miRNA profiling performed using the Agilent Human miRNA Microarray platform. To obtain miRNA target predictions, Microcosm, TargetScan and PicTar algorithms were used. Select collagen-specific miRNAs were validated using Taqman® MicroRNA Assays in samples from posterior and peripheral scleral regions (n=7, each group). Microarray data were analyzed using R (http://www.r-project.org), and quantitative PCR data with 2^-deltaCt method. **Results:** Human sclera expressed several miRNAs and many were differentially expressed, higher in rapidly growing fetal eyes ($p<0.01$, min $p= 6.5\times10^{11}$). Increased expression of collagen-specific mir-214, let-7c, let-7e, mir-103, mir-107, and mir-98 in fetal tissue was also demonstrated (fold changes 1.5 to 4, $p<0.01$) in follow-up experiments. In the validation experiments no significant differences were observed in miRNA expression between posterior and peripheral regions, for either age group. **Conclusions:** To our knowledge, this is the first comprehensive study of miRNA expression in human sclera. The sclera expresses several miRNAs, some of which show age-related differential regulation, consistent with a role in ocular growth regulation. Corresponding genome-wide mRNA profiling and analyses of correlations with micro-RNA signatures are underway to characterize the pathways involved, by way of exploring the potential role for miRNAs in ocular growth regulation in general and mechanisms underlying myopia development more specifically.

3522T

DNA combing assay for detection of contraction-dependent facioscapulohumeral muscular dystrophy (FSHD1). *F.Z. Boyar[1], P. Chan[1], V. Sulcova[1], D. Tsao[1], R. Owen[1], P. Walrafen[2], C.D. Braastad[3], M. Jocson[3], W. Sun[1], A. Anguiano[1], C.M. Strom[1].* 1) Department of Genetics, Quest Diagnostics Nichols Institute, San Juan Capistrano, CA; 2) Diagnostics Division, Genomic Vision, Paris, France; 3) Athena Diagnostics Inc., Worcester, MA.

FSHD, the third most common muscular dystrophy, is an autosomal dominant disorder with a prevalence of 4–10 per 100,000. It is characterized by often asymmetric progressive weakness of facial, scapular and upper arm muscles. FSHD1 is associated with the contraction of the D4Z4 region (3.3 kb repeat motif that contains DUX4, double homeobox 4, gene) at the subtelomeric region of 4q35 and activation of the normally epigenetically silenced DUX4. Patients carry 1–10 copies of D4Z4 on the 4qA allele, while normal individuals harbor 11–100 copies. Contractions of the partner 4qB allele or a homologous region on chromosome 10q26 are non-pathogenic. Detection of truncated D4Z4 and diagnosis of FSHD1 have been based on Southern blot which provides only an approximate size of the fragment of interest. Here we report a method for analysis of the D4Z4 region using DNA combing for direct visualization, recognition and measurement of the pathogenic 4qA allele. DNA combing and southern blot results showed complete concordance in a series of DNA plugs from normal and abnormal peripheral blood samples tested with the former providing a specific number of D4Z4 repeat elements. DNA combing assay was also performed directly on fresh blood and FSHD1 cell lines. Our findings underscore the utility of DNA combing assay for more accurate measurement of the D4Z1 locus on the pathogenic 4qA allele and diagnosis of FSHD1. It may also be applicable in the study of other disorders associated with copy number variation or repeat expansion and contraction. This represents the first successful validation of a molecular combing assay in a United States commercial laboratory.

3523T

RNA editing of the *SDHB* transcripts (C to U) in peripheral blood monocytes. *R.T. Taggart, J.D. Tario, K.H. De Jong, P.K Wallace, B.E. Baysal.* Dept Pathology, Roswell Park Cancer Institute, Buffalo, NY.

RNA editing is an epigenetic mechanism of gene regulation involving programmed enzymatic recoding of a RNA transcript. Targeted C to U RNA editing is performed primarily by APOBEC3 family of cytidine deaminase enzymes. Only a few endogenous gene transcripts of C to U editing are known. We described previously a recurrent stop codon mutation, C136T (R46X), in a fraction of succinate dehydrogenase subunit B gene, *SDHB*, transcripts in peripheral blood mononuclear cells (PBMCs). Similar mutations in the SDH subunits in germ line cause hereditary paraganglioma tumors. We describe here screening of cell types where *SDHB* editing occurs using a novel quantitative PCR assay for the measurement of mutated 136T versus wild-type 136C transcripts in cold-aggregated or CD14-column purified monocytes from peripheral blood and in various cell lines. We also measured expression of APOBEC3 family of enzymes.

Significant fraction of *SDHB* RNAs carries the 136T mutation in monocyte enriched populations (mean=6.3%; range 2.5–9.2%; n=4). *SDHB* editing is very low in lymphocyte enriched populations (mean=1.7%; range=1.3–2.2%; n=4) and absent in fibroblast line X269/8 and lymphoblastoid cell lines (<1%). Monocytic differentiation of HL60, THP1 and U937 cell lines by vitamin D exposure, as confirmed by flow cytometry, induces no *SDHB* editing. qRT-PCR analysis of APOBEC3 enzymes reveals that APOBEC3A, APOBEC3C and cytidine deaminase enzymes are expressed at least 5-fold more in monocyte-enriched PBMCs than in monocyte-depleted cells. The expression of other APOBEC3 enzymes is either very low in monocyte-enriched PBMCs (APOBEC3B, -3DE,) or decreased relative to monocyte-depleted PBMCs (APOBEC3F, -3G, -3H).

The *SDHB* gene is downregulated by site-specific C to U RNA editing in peripheral blood monocytes. SDH (mitochondrial complex II) is subject to tissue-specific dosage reduction by epigenetic mechanisms including genomic imprinting of *SDHD* and RNA editing of *SDHB*. Constitutive hypoxic phenotype of paraganglioma tumors suggests that these epigenetic changes may enhance hypoxia sensitivity and survival in normal cells.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3524T

FTO levels affect RNA modification and the transcriptome. *T. Berulava[1], M. Ziehe[2], L. Klein-Hitpass[3], E. Mladenov[4], J. Thomale[3], U. Rüther[5], B. Horsthemke[1].* 1) Institut für Humangenetik, Universitätsklinikum Essen, Hufelandstrasse 55, 45122 Essen, Germany; 2) Analytical and Environmental Chemistry, Humboldt-Universität zu Berlin, Brook-Taylor-Str.2, 12489 Berlin, Germany; 3) Institut für Zellbiologie (Tumorforschung), Universitätsklinikum Essen, Hufelandstrasse 55, 45122 Essen Germany; 4) Institut für medizinische Strahlenbiologie,Universitätsklinikum Essen, Hufelandstrasse 55, 45122 Essen, Germany; 5) Institut für Entwicklungs- und Molekularbiologie der Tiere (EMT) Heinrich-Heine-Universität, Universitätsstr.1, 40225 Düsseldorf Germany.

A block of single nucleotide polymorphisms (SNPs) within the intron 1 of the *FTO* (fat mass and obesity associated) gene is associated with variation in body weight. Previous works suggest that increased expression of *FTO* leads to increased body weight, although the underlying mechanism has remained unclear. The gene encodes a 2-oxoglutarate-dependent RNA demethylase, which has been shown to demethylate 3meU, 3meC, 1meA *in vitro* and 6meA *in vitro* and *in vivo*. To elucidate the function of FTO we examined the consequences of altered FTO levels in cultured cells and murine brain. Here we show that a knockdown of *FTO* in HEK293 cells affects the transcripts levels of genes involved in the response to starvation, whereas overexpression of FTO affects the transcript levels of genes related to RNA processing and metabolism. Changes in transcripts levels may reflect changes in transcript stability depending on the degree of adenine methylation. Subcellular localization of FTO further strengthens the notion that FTO is involved in RNA processing and metabolism. Using immunocytochemistry and confocal laser scanning microscopy, we detected FTO in nuclear speckles and - to a lesser and varying extent - in the nucleoplasm and nucleoli of HEK293, HeLa and MCF-7 cells. Moreover, RNA modification analyses of total brain RNA revealed that loss of Fto affects the 3-methyluridine/uridine and pseudouridine/uridine ratios. We conclude that altered levels of FTO have multiple and diverse consequences on RNA modifications and the transcriptome.

3525T

Prioritizing candidate functional SNPs at GWAS loci using epigenomic datasets from ENCODE. *K.S. Lo[1,2], C.D. Palmer[3,4], J.N. Hirschhorn[3,4], G. Lettre[1,2].* 1) Montreal Heart Institute, Montreal, Quebec, Canada; 2) Université de Montréal, Montréal, Québec, Canada; 3) Children's Hospital Boston, Boston, MA, USA; 4) Broad Institute, Cambridge, MA, USA.

INTRODUCTION: Genome-wide association studies (GWAS) have identified hundreds of coding but mostly non-coding single nucleotide polymorphisms (SNPs) associated with complex human diseases and traits. At most GWAS loci, the SNPs that causally influence gene products and thus phenotypic variation remain unknown. Epigenomics could help clarify some of these findings by prioritizing functional follow-up through the discovery of regulatory elements.

METHODS & RESULTS: To identify functional variants at the 180 loci associated with adult height by the GIANT Consortium (Lango Allen et al., Nature, 2010), we used epigenomic data compiled by the Encyclopedia of DNA Elements (ENCODE) Consortium. Because there is no clearly relevant cell line in this dataset to study the biology of height, we decided to explore 24 epigenomic tracks available from the UCSC genome browser for human umbilical vein endothelial cells (HUVEC). We tested the overlap between the 180 height SNPs (and their linkage disequilibrium (LD) proxies) and features such as chromatin accessibility, DNA methylation, histone modification and transcription factor binding sites. To assess statistical significance, we repeated the analysis with 1000 sets of 180 SNPs matched on allele frequency, gene proximity and LD. We observed a strong enrichment ($P<0.001$) for most epigenomic features, except for the histone H3 lysine 27 tri-methyl (H3K27me3) mark associated with transcriptional repression ($P=0.25$) and for the hematopoietic lineage-specific GATA2 transcription factor binding sites ($P=0.58$). These results are consistent with the previously reported enrichment of expression quantitative trait loci (eQTL) among the height loci.

CONCLUSION: We show that there is a strong enrichment of epigenomic features that overlap with adult height association signals. Our approach is applicable to other complex human phenotypes and should help fine-map causal variants identified by GWAS.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3526W

Bioinformatic Parallel Processing Tools Development for Mutation Identification from Whole Exome Data Following Homozygosity Mapping for Autosomal Recessive Disorders. *Y. Al-Sarraj[1], A. Abouzehry[1], H. ElShanti[1,2], M. Kambouris[1,3].* 1) Shafallah Medical Genetic Center, Doha, Qatar; 2) University of Iowa | Iowa, Pediatrics, Iowa City, IA, USA; 3) Yale University School of Medicine, Genetics, New Haven, CT, USA.

Eight consanguineous Arab families with novel autosomal recessive disorders were mapped with illumina OmniExpress 700K SNP genotyping & Homozygositymapper software. All relevant positional candidate genes were screened for pathogenic mutations. None were identified. Multiple homozygosity intervals were obtained for each family since no significant LOD scores were possible. Whole exome target enrichment sequencing was on ABI SOLiD4 for one affected individual from each family. Sequence mapping and annotation was on LifeScope software. Data validation was done manually for each linkage interval, by visual inspection of read depth and bead number coverage. On average 30,000 sequence variations were detected in each sample including novel variants, known polymorphisms & exome sequencing errors. For each chromosome with a linkage interval, data was isolated and filtered by exportation to Excel spreadsheets and visual inspection to exclude non-linkage interval data. The number of variants in the linkage intervals for each family was between 400 and 1300. Homozygous sequence variations within the linkage intervals were between 50 and 300 with 15–30 novel variants. Determination if a variant was homozygous or heterozygous, novel or annotated was done manually upon visual inspection of data on Excel spreadsheets. For each novel variant it was manually determined if it were exonic, splice site specific or intronic. For each annotated variation it was manually determined if it is associated to a disease phenotype relevant to the family disease. Minor genotype frequency ware investigated for annotated variants if they represent disease states. All novel exonic variants were tested in silico with PolyPhen and Sift Protein Modeling software to access the effect on protein function. All damaging variants (novel or annotated exonic, and splice site) were validated by Sanger sequencing and tested for co-segregation to disease. An identical approach is essential to access pathogenic effects of insertion/deletion variants within each linkage interval. This approach is tedious, involves a tremendous amount of manual work and is prone to oversight errors. Software tool development for automating next generation sequencing data analysis is essential to eliminate manual work and identify pathogenic mutations among the plethora of existing variants. Such automation is applicable in cases without linkage intervals to limit the number of variants under consideration.

3527F

Kailos BlueTM: A Complete Cloud-Based Bioinformatics Solution for Management and Analysis of Targeted Sequencing Across Next Generation Sequencing Platforms. *R.C. Bachmeyer[1], D.A. Kloske[1], B. Stone[1], K.E. Varley[1,2], D.T. Moore[1].* 1) Kailos Genetics Inc., Huntsville, AL; 2) HudsonAlpha Institute for Biotechnology, Huntsville, AL.

Next-generation sequencing (NGS) applications are developing at unprecedented speed, enabling the large-scale production of sequence data for whole genome, exome and targeted sequencing. This scale of data generation presents significant computational challenges for most laboratories in storage, retrieval, analyses, and equipment run cost management. There exist many differing software solutions designed and optimized to handle each individual step in the overall NGS workflow. However, these solutions often require a high degree of IT knowledge to implement, configure, support and operate. Additionally, connectivity and storage are generally limited to local systems that were not conceived of with NGS in mind. We developed Kailos Blue, a highly secure, easily accessible cloud environment, for investigators to manage their NGS workflow, analyze sequences post-target enrichment and actively utilize their stored data. In combining the Amazon cloud (secure & scalable), the industry leading alignment and genotyping systems (BWA & GATK), statistical resources and an intuitive user interface, investigators from any size lab can perform complex analysis of samples within a run and across multiple runs over time, independent of the sequencing equipment utilized.

Kailos Blue provides an end-to-end informatics solution for targeted next-generation sequencing. The system can be used to plan and organize flowcells and configure sequencing runs. This feature is especially important for targeted sequencing approaches, such as Kailos Genetics' Target Rich system, where dozens of samples are barcoded and pooled to efficiently and optimally utilize sequencer capacity. Raw data can be uploaded from sequencing instruments to the cloud for fast and secure alignment and analysis. Analysis, including variant identification, can be performed on DNA samples and compared against a reference sample (1000 Genomes, HapMap, User Defined) or across multiple samples run over any period of time, enabling entire studies to be compiled and analyzed in one system. Here we will present the overall design of the cloud-based bioinformatics solution and its performance in analyzing targeted sequencing data generated on the Illumina MiSeq, GAIIx and the Ion Torrent PGM.

3528W

Streamlined, accurate low level detection of variants in deep, next generation sequencing data from amplicons. *C. Boysen, M. Matvienko, N. Thomson, B. Turner, J. Bendtsen, H. Sandmann, J. Jakobsen, P. Nielsen, A. Joecker, A. Joecker.* CLC bio, Cambridge, MA.

Purpose: Next generation sequencing is being added to research, diagnostics, and clinical labs to identify disease linked DNA variants. However, NGS applications and platforms each have various issues that can confound accurate detection of variants. Here we address issues associated with amplicon-based deep re-sequencing data analysis for identification of low level variants as warranted in the detection of somatic mutations in potentially mixed cancer/normal tissue. Study and Results: Using CLC bio Genomics Workbench, we have made an efficient re-sequencing workflow for the end-to-end analysis of NGS data. This workflow has been optimized to accommodate the characteristics associated with short amplicon, deep sequencing. Particular mapping and variant calling algorithms have been deployed to handle known NGS platform dependent issues, such as homopolymer problems and other consistent sequencing errors. Variants can be ranked and filtered according to a number of criteria, incl. quality values, forward/reverse reads, comparisons to variant databases and known function or disease associations. To help ensure sufficient coverage of all amplicons and flag potential drop-outs, an extensive suite of quality control measurements have been incorporated, and provides the user with an easily interpreted overview of all samples in a given run. Results regarding sensitivity and specificity are presented as is discussion on frequently encountered issues and the provided solutions for best practices.

3529F

De novo Assembly by Positional Sequencing. *B. Bready, P. Goldstein, W.H. Heaton, P. Ianakiev, H-Y. Li, J.S. Oliver, J. Thompson.* Nabsys Inc, Providence, RI.

Because of the highly repetitive nature of human DNA and the relatively short length scales over which DNA sequencing platforms obtain information, assembling the data produced by these platforms is computationally intensive and results in contig lengths that are very short compared to the lengths of chromosomes. Complete genomes require significant additional finishing to unambiguously place repeated or difficult regions and to order scaffolds and contigs. In contrast, Positional Sequencing as developed by NABsys using nano-scale detectors and specific hybridization probes will provide information over hundreds of kilobases and even megabases of contiguous sequence. Specifically, the platform locates, with sub-diffraction-limit resolution, the positions of oligonucleotide probes that have bound to long DNA fragments. This information can be assembled into contigs whose lengths approach the lengths of chromosomes. M13 has been used as a model system for sequencing. Hybridization probes of 8 to 20 long were found effective in providing data that allowed resequencing of parts of the M13 genome in preparation for larger genomes. An error model based on this effort is used to determine the performance of the system on more complex genomes.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3530W

Translational research in newborn screening: development of informatics tools to support longitudinal research and accelerate transformations in clinical practice. *A. Brower[1], B. Bowdish[1], M. Porter[2], J. Pennington[2], S. Wrazien[2], J. Loutrel[2], P. White[2,3], M. Watson[1].* 1) Newborn Screening Translational Research Network, American College of Medical Genetics, Bethesda, MD, 20814, USA; 2) Center for Biomedical Informatics, The Children's Hospital of Philadelphia, Philadelphia, PA, 19104, USA; 3) Department of Pediatrics, University of Pennsylvania School of Medicine, Philadelphia, PA, 19104, USA.

Background: The goal of newborn screening is the early identification, diagnosis and treatment of disease to improve health outcome for affected infants. Over the past decade innovations in analytical technology have enabled the rapid expansion of the number of conditions included in screening efforts, and a system of evidence review has led to further expansion. All states have implemented screening for a core panel of conditions, resulting in nearly 12,500 newborns diagnosed each year. While the majority of the conditions require life-long medical care and management, there is no nationwide system to collectively assess health outcomes and further scientific understanding of the conditions. Informatics tools that can be incorporated into clinical workflows in a variety of care settings are required. These tools will facilitate the translation new discoveries to clinical practice and improve scientific understanding. Methods: In 2008, the NICHD established the Newborn Screening Translational Research Network (NBSTRN) Coordinating Center to develop tools and infrastructure to support research in newborn screening. The NBSTRN established a broadly constituted workgroup comprised of content experts from laboratory-based specialties, clinical care teams and state-based newborn screening programs to identify the data from routine clinical care that would be important for research efforts. A data capture tool and management system is under development and leverages the Research Electronic Data Capture (REDCap) system to create user friendly clinical case report forms that incorporate national information standards to facilitate information sharing, data mining and data aggregation. Development of this informatics tool and an accompanying almanac of definitions and standards are focused initially on the forty-two metabolic disorders as well as conditions that have not been recommended for routine newborn screening. Results: A series of clinical case report forms that incorporate a consensus data set defined by an NBSTRN workgroup and a data management system are in development. Condition-specific clinical case report forms for currently screened and candidate conditions for newborn screening will be presented. Conclusions: The creation of informatics tools to support the longitudinal gathering, sharing and analysis of the health information of diagnosed individuals supports research and facilitates the translation of new discoveries to clinical practice.

3531F

Web-based tool for target sequence capture assay design. *J. Cai, B. Morrow.* Department of Genetics, Albert Einstein College of Medicine, Bronx, NY.

Re-sequencing of selected genes of interest within the human genome has become a popular strategy to discover variants associated with disease. Compared with whole exome sequencing, the customized target capture approach of known genes related with a certain phenotype/disease has the advantage that of achieving deeper coverage with less cost. It also will enable putative regulatory regions or alternatively spliced exons to be included that may not be present in generic exome capture kits. We developed a user friendly web-based tool that create sequence capture assay design from a gene list suited for clinicians or basic researchers, who are not experts in mining genomic data. The tool has two components. One of the components is named as P2G (from Phenotype to Genes). Given terms of disease name or phenotype, such as hearing loss, deafness, the program will generate a list of known genes related to the input biological terms. The known gene information regarding a phenotype term is based on both OMIM database and the Genetic Association Database. The other component is named G2L (from Genes to Locations). Given a list of official gene symbols (can be automatically generated from the P2G component), the program will generate a bed file with genomic location according to Ref_Gene Annotation from UCSC database. This component also provides flexibility in selection of different regulatory regions, such as UTR, up/down-stream regions.

3532W

wANNOVAR: annotating genetic variants for personal genomes via the Web. *X. Chang[1], K. Wang[1,2].* 1) Zilkha Neurogenetic Institute, Keck School of Medicine, University of Southern California, Los Angeles, USA; 2) Department of Psychiatry and Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, USA.

High-throughput DNA sequencing platforms have now become widely available. As a result, personal genomes are increasingly being sequenced in research and clinical settings. However, the resulting massive amounts of variants data pose significant challenges to average biologists and clinicians without bioinformatics skills. To fill the urgent demands, we built a web server called wANNOVAR to address the critical needs for functional annotation of genetic variants from personal genomes. The server was developed based on the ANNOVAR (ANNOtate VARiation) software. It provides simple and intuitive interface to help users determine the functional significance of variants without the burden of typing command lines and configuring database files. wANNOVAR has two major features including “annotating variants” and “filtering variants”. For the submitted variants list, wANNOVAR will first annotate single nucleotide variants (SNVs) and insertions/deletions for their effects on genes, report their conservation levels (such as PhyloP and GERP++ scores), calculate their predicted functional importance scores (such as SIFT and PolyPhen scores), retrieve allele frequencies in public databases (such as the 1000 Genomes Project and NHLBI-ESP 5400 exomes). Optionally, a “variants reduction” protocol will be applied on the variants list to identify a subset of potentially deleterious variants/genes. We provide several default pipelines for different disease models such as “rare recessive Mendelian disease” and “rare dominant Mendelian disease”, but users can also use “advanced options” to specify custom filtering strategy. We next illustrated the utility and functionality of wANNOVAR on an exome sequencing and a whole-genome sequencing data set on the “Example” page. We conclude that wANNOVAR will help biologists and clinicians take advantage of personal genome information to expedite scientific discoveries. The wANNOVAR server is available at http://wannovar.usc.edu, and will be continuously updated to reflect the latest annotation information.

3533F

Simultaneous analysis of common and rare variants in complex traits. *G. Chen[1], A. Yuan[2], A. Bentley[1], D. Shriner[1], A. Adeyemo[1], C. Rotimi[1].* 1) NHGRI/NIH, National Institutes of Health, Bethesda, MD; 2) National Human Genome Center, Howard University, Washington DC, USA.

Genome-wide association studies (GWAS) have facilitated the detection of common genetic variants underlying common traits and diseases. However, significant proportions of the underlying genetic variance for the traits investigated to date remain unexplained. Also, the predictive power of common variants identified by GWAS has not been encouraging. Given these observations along with the fact that the effects of rare variants are often unaccounted for by GWAS and the increasing availability of sequence data, we developed a new method that enables the simultaneous analysis of the association between individual common and rare variants while controlling for the potential confounding effects of covariates. We refer to this method as SCARVAsnp. SCARVAsnp is implemented in four stages: first, all common variants in a pre-specified region (e.g. gene) are evaluated individually; second, a union procedure is used to combine all rare variants (RVs) in the index region and the ratio of the log likelihood with one RV excluded to the log likelihood of a model with all the collapsed RVs is calculated. On the basis of simulation studies, a likelihood ratio ≥ 1.3 is considered statistically significant; third, the direction of the association of the removed RV is determined by evaluating the change in λ values with the inclusion and exclusion of that RV. Lastly, significant common and rare SNPs along with covariates are included in a final regression model to evaluate the association between the index trait and variants in a pre-specified genomic region. We use simulated and real data sets to show that the method is simple to use, computationally efficient and that it can accurately identify both common and rare risk variants. This method overcomes several limitations of existing methods; for example, SCARVAsnp limits loss of statistical power by not including rare variants that are not associated with the trait of interest in the final model. The combined analysis of rare and common variants may be important in explaining the missing heritability from GWAS analyses. Also, SCARVAsnp takes into consideration the direction of association by effectively modelling positively and negatively rare associated variants.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3534W

On Combining Reference Data to Improve Imputation Accuracy. *J. Chen[1], J.G. Zhang[1], J. Li[1], Y.F. Pei[1,2], H.W. Deng[1,2].* 1) Center for Bioinformatics and Genomics, Department of Biostatistics and Bioinformatics, School of Public Health and Tropical Medicine, Tulane University, New Orleans, LA 70112, USA; 2) Center of System Biomedical Sciences, University of Shanghai for Science and Technology, Shanghai 200093, P. R. China.

Background: Genotype imputation is an important tool in human genetics studies, which uses reference sets with known genotypes and prior knowledge on linkage disequilibrium and recombination rates to infer un-typed alleles for human genetic variations at a low cost. The reference sets used by current imputation approaches are based on HapMap data, and/or based on recently available next-generation sequencing (NGS) data such as data generated by 1000 Genomes Project. However, with different coverage and call rates for different NGS data sets, how to integrate multiple NGS data sets along with previously available reference data as references in imputation is not an easy task and has not been systematically investigated. Results: In this study, we perform a comprehensive assessment of three strategies on using NGS data and previously available reference data in genotype imputation for both simulated data and empirical data, in order to obtain guidelines for optimal reference set construction. Briefly, we consider 3 strategies: strategy 1 uses one NGS data as a reference; strategy 2 imputes samples by using individual available data sets as independent references and then combines the imputed samples with samples based on high accuracy reference selected when overlapping occurs; and strategy 3 combines multiple available data sets as a single reference after imputing each other. Our results show that strategy 2 and strategy 3 has higher imputation accuracy than strategy 1, and strategy 2 is the best strategy across all the conditions that we have investigated. Especially, strategy 2 also could acquire the better accuracy of imputation for rare variant than the other two strategies. The one reason is more reference information used for both strategy 2 and strategy 3. Another reason is that strategy 2 avoided the noise brought about because of repeated imputation (strategy 3), and retain more reliable genotype information from high accuracy reference. Conclusion: We illustrate and compare different imputation strategies to integrate multiple references. Our study is helpful in guiding application of imputation methods in next generation association analyses.

3535F

Improved Filtering and Annotation of Next-Generation Sequencing Variants in Clinical Gene Panels. *H.K. Chong, S. Mexal, A.M. Elliott, H.M. Lu, H. Lu, X. Li.* Ambry Genetics, Aliso Viejo, CA.

Gene panels offer a cost-effective and targeted approach to investigating complex inherited disorders, allowing clinicians to test multiple genes in parallel by target capture followed by Next-Generation sequencing. Comprehensive gene panels can often offer a definitive molecular diagnosis and aid in making a differential diagnosis. A major limitation to gene panel testing is that data interpretation becomes more challenging as more genes are tested and hence an increasing number of variants detected. To reduce the number of variant calls for clinical interpretation, we evaluated the ability of a new in-house custom pipeline in filtering and annotating variant calls from two comprehensive diagnostic gene panels offered on our test menu; the Pan Cardio Panel, an inherited cardiovascular disease panel testing 79 genes (1573 number of amplicons), and the CancerNext panel, a 22 gene panel (477 number of amplicons) test spanning genes implicated in hereditary breast, ovarian, and colon cancers. The pipeline filters out variants based upon population frequency from several databases and annotates alterations using disease mutation databases and evolutionary conservation. The pipeline ultimately classifies variants into several different categories, including common polymorphisms, mutations, and unclassified variants that require further investigation. In a pilot study of eight wild-type and five previously characterized dilated cardiomyopathy (DCM) samples, the Pan Cardio Panel generated an average of 200 variants prior to filtering and an average of 28 variants after the alterations were run through the pipeline. Across wild-type and previously characterized cancer samples (8 total), the CancerNext Panel generated an average of 57 variants prior to filtering and an average of 16 variants after the alterations were generated through the pipeline. Among the variants left in both the Pan Cardio and CancerNext samples, the pipeline successfully identified all the previously characterized mutations in the DCM and cancer patients as either deleterious based upon HGMD entries or evolutionary conservation. These data illustrate that our custom Next-Gen panel pipeline can successfully filter out non-causative variations resulting in a decrease in the number of variants that need to be verified by side assays, saving significant time and money.

3536W

Discovery of a *cis*-regulatory SNP at the *PPARG* diabetes risk locus. *M. Claussnitzer[1,2,3,4], H. Grallert[3,5], S.N. Dankel[6,7], B. Klocke[11], H. Lee[1,2,3], S. Hauck[8], V. Glunk[1,2,3], C. Hoffmann[2,10], M. Seifert[11], G. Mellgren[6,7], T. Illig[5,9], H. Hauner[1,2,3], H. Laumen[1,2,3].* 1) EKFZ, Chair for Nutritional Medicine, Technical University Muenchen, Freising-Weihenstephan, Germany; 2) ZIEL, Research Center for Nutrition and Food Sciences, Technical University Muenchen, Freising-Weihenstephan, Germany; 3) Clinical Cooperation Group Nutrigenomics and Type 2 Diabetes, Technical University Muenchen, Freising-Weihenstephan, and Helmholtz Zentrum Muenchen, Muenchen-Neuherberg, Germany; 4) Hebrew SeniorLife Institute for Aging Research and Harvard Medical School, Boston, MA, USA; 5) Research Unit of Molecular Epidemiology, Helmholtz Zentrum Muenchen, Muenchen-Neuherberg, Germany; 6) Institute of Medicine, University of Bergen, N-5021 Bergen, Norway; 7) Hormone Laboratory, Haukeland University Hospital, N-5021 Bergen, Norway; 8) Research Unit Protein Science, Helmholtz Zentrum München, Muenchen-Neuherberg, Germany; 9) Hanover Unified Biobank, Hanover Medical School, Hannover, Germany; 10) EKFZ, Chair for Molecular Nutritional Medicine, Technical University Muenchen, Freising-Weihenstephan, Germany; 11) Genomatix Software GmbH, Muenchen, Germany.

Genome-wide association studies (GWASs) have revealed numerous risk loci associated with a diverse range of diseases. However, identifying the disease-causing variants remains a major challenge in medical genetics. We used phylogenetic module complexity analysis (PMCA), a novel bioinformatic approach for delineating cis-regulatory variants hidden in GWAS loci. PMCA combines sequence conservation with functional conservation by considering the organisation of transcription factor binding sites in complex cis-regulatory modules. We have cross-validated our predictions for type 2 diabetes (T2D) susceptibility loci. For the PPARG risk locus we found genotype-dependent regulation of endogenous PPARγ2 mRNA expression. We identified the homeodomain transcription factor PRRX1, binding to the PMCA identified variant rs4684847, as an isoform-specific transcriptional repressor of PPARγ2 mRNA expression in adipose patient samples of risk allele carriers. Though the PPARγ2 Pro12Ala variant of the T2D susceptibility locus was reproducibly associated with improved insulin sensitivity, the Ala12 allele results in reduced PPARγ2 protein transcriptional activity. This contradictory finding might be explained by the identified cis-regulatory variant rs4684847. Thus, PMCA may contribute to translate genetic association signals to the mechanism underlying the biology of complex diseases.

3537F

dbGENO: A web based complex disease curation pipeline. *D. Cook, J. Dagle, K. Ryckman, J. Murray.* University of Iowa, Iowa City, IA.

Despite improving genetic technologies, there remain significant challenges for investigators trying to understand the genetic etiology of complex diseases. Disease/trait specific databases aim to further our understanding by indexing publications by genetic loci found to be associated (forming a comprehensive list of known genetic associations). Investigators and clinicians can browse and retrieve data for incorporation in the design of a study, for replication purposes, to further understand a complex disease, or for other data related purposes. While such databases have wide applicability, they are time consuming and challenging to assemble and deploy. We have developed a web-based pipeline of tools for the curation of disease/trait specific databases that aims to improve collaboration and efficiency of the development process. Additionally, our tool works to reduce error and increase the thoroughness of literature searches. Finally, our tool works to automate the process of finding and identifying new publications which may be candidates for inclusion. The pipeline begins with a utility for importing publications, by performing PubMed searches directly or via a custom list of publications relevant to a disease or trait of interest. From these imports, a variety of features are extracted such as SNPs, genes, diseases, and mutations. These extracted features are used as a basis for indexing and later searching and sub setting data for retrieval. Following feature extraction, users examine these features to ensure they are correct and appropriate and to add additional publication details if necessary. Finally, a web-based interface allows anyone to browse and export information from the database. The pipeline is written in PHP and uses MySQL for data storage. NCBI's EFetch tools are used heavily for data retrieval of publication, SNP, and gene records. A combination of regular expressions and rule-based algorithms are used to extract and filter features for indexing. We are performing diagnostics and continually improving feature extraction algorithms by testing against manual curation. Initially, we plan to assemble a database on the basis of the complications of prematurity. As our understanding of the human genome continues to grow, the need for database tools accessible to clinicians and researchers alike grows concomitantly. dbGENO aims to fulfill this need.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3538W

EVA: Exome Variation Analyzer, a tool for filtering strategies in medical genomics. *S. Coutant[1,2], C. Cabot[2], A. Lefebvre[2], M. Léonard[2], E. Prieur-Gaston[2], D. Campion[1], T. Lecroq[2], T. Frebourg[1,3], H. Dauchel[2].* 1) Inserm U1079, Institute for Research and Innovation in Biomedicine, (IRIB), University of Rouen, Normandy, France; 2) LITIS EA 4108 Computer science, Information processing and systems laboratory, IRIB, University of Rouen, Normandy, France; 3) Department of Genetics, University Hospital, Rouen, Normandy, France.

Exome-scale sequencing is a revolution in medical genetics, impacting both fundamental research and diagnostic. However, the challenge remains in efficient filtering strategies to find causing disease variant(s) among ~20,000 per individual exome. For this purpose, analytical procedures with various filtering strategies are required. With this aim, in parternship with and for medical geneticists, we developed EVA (Exome Variation Analyzer) a user-friendly web interfaced sofware dedicated to filtering strategies for medical projects. EVA allows managing data through different modules. Among them, the first is an offline Variation Integration module. It takes as input standard Variant Calling Format files and annotates the variations thanks to ANNOVAR and the Variation Effect Predictor Ensembl API. Annotated variations are then stored in a database. On the web interface, the Variation Statistics module displays basic count of SNV and indel for all the variations corresponding to a project or a selection of individuals, chromosomes, regions or genes. The filtering module proposes to combine multiple filters to drastically narrow down the variations number. It compares data to international catalogues of variations (dbSNP, 1000 genomes, ESP Cohort) and sorts out variations depending on their: a) functional categories (synonymous, missense, nonsense, frameshift...) b) genic region (UTR, exons, introns, splice sites) and c) quality. Finally, the major strength of EVA is the implementation of inheritance filters that considers intersection or differential exome strategies: a) recurrence strategy for dominant or recessive non-familial cases; b) intra-familial strategy with filters for homozygous or heterozygous composite; c) differential exomes strategy for de novo sporadic cases (trio). EVA offers export files and cross-links to external relevant databases and softwares for further inspection of the small subset of sorted candidate variations and genes. EVA is developed to be a user-friendly, versatile, and efficient-filtering assisting software. Thereby, it provides a response to new needs at the expanding era of medical genomics for both fundamental research and clinical diagnostics. EVA will be soon available for free downloads.

3539F

Random Design Tool for Additional MIDs for Roche Sequencing Platforms. *S.H. Eck, B. Grumbt, I. Nieberle, C. Marschall, K. Hirv, H-G. Klein.* Center for Human Genetics and Laboratory Medicine, Dr. Klein, Dr. Rost and Colleagues, Martinsried, Germany.

In order to achieve a higher grade of multiplexing we implemented a Perl tool to design additional MIDs for the Roche sequencing platforms. MIDs are designed following a simple two step procedure. In the first step randomized sequences of the desired length are generated. In the second step the sequences are checked against a set of rules to ensure compatibility with the Roche sequencing systems in general and existing barcodes in particular. These rules include a minimum number of mismatches to existing MIDs (default 4), defined number of flows to sequence (existing barcodes range from 4–6 flows, default 5) and limited homopolymer length within the barcode (default 1 = no subsequent identical nucleotides). Existing MIDs are loaded for comparison, if a new designed MID passes the rule check it is automatically added to the list of known MIDs to enable comparison with further designed MIDs. We validated designed primers by sequencing 96 barcoded samples (8 amplicons of mean length 570) on a quarter of a Roche FLX slide. 48 MIDs were taken from the Roche extended MID set, 48 MIDs were random designed by the tool. In total we generated 133,780 passed filter reads on this part of the slide. Mean number of reads per MID were 1237 (max 1826, min 784) and 1122 (max 1644, min 15) for the random designed and Roche extended MID set, respectively. Standard deviation was 299 and 336 for the two sets. We also included a negative control which counted 12 reads. In total only 1492 reads (~1.1%) could not be assigned to any MID. The tool is flexible and can be used to design arbitrary numbers of additional MIDs for the Roche platforms. MID parameters like length, minimum differences to existing MIDs and target flow count can be defined by the user depending on the particular experimental reads. The tool is designed to generate Roche compatible MIDs but can be easily extended to generate barcodes for every sequencing platform.

3540W

A functional gene-based test for association in Tourette's syndrome and obsessive-compulsive disorder. *P. Evans, TSA Consortium for Genetics, The International OCD Foundation Genetics Collaborative.* Dept of Med., Univ Chicago, Chicago, IL.

Tourette's syndrome (TS) is a developmental disorder characterized by motor and phonic tics that is present in 0.03–0.08 of the population, though more prevalent in males than females. Family members of probands with TS are significantly more likely than members of the general population to be diagnosed with obsessive compulsive disorder (OCD), and vice versa. Twin and family studies support a high heritability for both disorders, yet very little is known about the genetics of either disorder. In order to better understand the underlying genetics of TS and OCD we used a gene-based method to look for genes that may be playing a role in the genetic epidemiology of these disorders. One of the disadvantages to traditional genome-wide association tests is the large penalty paid for multiple testing correction for the one million plus SNPs that are commonly interrogated in these studies. This makes detection of the functional units contributing to disease difficult, as effect sizes are expected to be small for most loci contributing to complex. One alternative to this type of test is to use genes rather than SNPs as the unit of analysis, greatly reducing the number of tests performed. We have chosen to use this method by combining SNPs for a given gene based on function. This is accomplished by annotating each gene with nonsynonymous and nonsense mutations within a given gene, as well as any known cis or trans eQTLs for a given gene. The p-values from a traditional genome-wide association analysis are combined for all SNPs annotated to a given gene using an application of Fisher's method. Gene-based p-values are then obtained by permuting the data and seeing how often a similar score is achieved using the same number of minor allele frequency matched SNPs. We performed a function-based genome-wide association study on results from two studies on TS and OCD using eQTL data from lymphoblastoid cell lines. For TS, the gene CCDC102B was genome-wide significant given the total number of genes examined (p=0.00004). For OCD, the genes FZD5 and KCTD17 were genome-wide significant with p-values of 0.00009 and 0.00002 respectively. These results may provide some insight into the underlying mechanisms contributing to these disorders as no SNPs achieved genome-wide significance in the traditional SNP-based analysis. We are now conducting similar analysis using eQTLs identified in transcriptome studies on brain, which will be presented at the meeting.

3541F

Statistical considerations for decisions on applying globin RNA reduction in blood samples for gene expression experiments. *S. Feng[1], J. Dungan[2], S. Gregory[2], C. Haynes[2], W. Kraus[2].* 1) Biostatistics & Bioinformatics, Duke Univ, Durham, NC; 2) Center for Human Genetics, Duke Uniersity, Durham, NC.

The technique of globin RNA reduction (GR) aims to remove unwanted globin mRNA from whole blood total RNA. Some studies have suggested that GR can improve the quality of gene expression data. However, the evidences supporting those claims have not been absolutely convincing. In this study, we wished to address two questions: (1) the significant benefit (increment in data quality) gained by applying GR; and more important, (2) if a benefit does exist, how useful it is, i.e., how much additional statistical power will be gained for the complicated multi-step microarray data analysis? We conducted a pilot experiment to measure the differences and similarities between the pre- and post- GR gene expression profiles, as well as the standard summary statistics. A straightforward statistical simulation study was also designed to investigate the impact of GR on the statistical powers under various scenarios. The results confirmed others' observations, i.e., if the technical procedure is successfully performed, the overall sample variations are likely to be reduced. However, the overall benefit is clouded by significant sample and RNA loss (~16%) due to the additional experimental steps; the inconsistent results among samples and the new and untraceable variations generated by the procedure. The simulation study also suggested that improvements in data quality do improve statistical power, but is only relevant under certain conditions. Based on those observations, we conclude that the overall benefits of applying GR are much less certain than expected. Thus the decision of whether to apply GR should be specific to different experiment designs and conditions, as revealed by the simulation study.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3542W

Biofilter 2.0 for Advanced Predictive Model Development, Testing, and Hypothesis Generation using Expert Domain Knowledge Resources. *A. Frase[1], J. Wallace[1], C. Moore[1,2], N. Katiyar[1], S.A. Pendergrass[1], M.D. Ritchie[1].* 1) Center for Systems Genomics, Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, PA; 2) Center for Human Genetics Research, Vanderbilt University.

Leveraging the incredible diversity and wealth of biological information collected from the genome, transcriptome, proteome, and other -omic sources, is key for advanced and directed predictive model for the dissection of complex disease. We developed Biofilter for using biological information from public databases to direct complex risk model generation. The first application of Biofilter (Bush, et al., *Pacific Symposium on Biocomputing*, 2009) used 7 sources of domain knowledge: the National Center for Biotechnology (NCBI) dbSNP and gene Entrez database information, Kyoto Encyclopedia of Genes and Genomes (KEGG), Reactome, Gene Ontology (GO), Protein families database (Pfam), and NetPath - signal transduction pathways. Originally Biofilter was for use with gene or single-nucleotide polymorphism (SNP) based data, such as filtering genome wide association (GWA) genotypic data for biologically directed model generation. For example, through directing the search space using Biofilter with GWA data, pairwise SNP-SNP interaction models were developed, replicating across datasets (Turner et al.*PLoS One*, 2010). We expanded and updated Biofilter to version 2.0 in several important ways. First, we migrated the domain knowledge used by Biofilter into a separate database called the Library of Knowledge Integration (LOKI) and incorporated 4 additional data sources: Molecular INTeraction database (MINT), Biological General Repository for Interaction Datasets (BioGrid), Pharmacogenomics Knowledge Base (PharmGKB), Open Regulatory Annotation (ORegAnno), and the database of Evolutionary Conserved Regions (ECRBase). The independence of LOKI from Biofilter 2.0 facilitates easy update of the data sources contained within LOKI, reflecting the update of public LOKI data sources. Biofilter 2.0 now accepts a greater range of data types, including SNP, rare variant, copy-number variation (CNV), and evolutionary conserved region (ECR) data. Future directions include development of Biofilter for use with RNA-Seq based data. Biofilter 2.0 allows more user customization of filtering and annotation, and easier use of custom data sources, providing a flexible framework for the integration and user-driven combination of multiple data types, providing a way to use the ever-expanding expert biological knowledge that exists to direct complex predictive models for elucidating the etiology of complex phenotypic outcomes. The software is available for use at http://ritchielab.psu.edu/.

3543F

Iterated Correction of a Diploid Parental Reference Sequence and its use during Exome Alignments for Disease Gene Discovery. *T.R. Gall[1], D.R. Adams[1,2,3], C.F. Boerkoel[1], K. Fuentes-Fajardo[1], W.A. Gahl[1,2,3], P. Cherukuri[1], M. Sincan[1], C. Toro[1], C.J. Tifft[1,3], T.C. Markello[1,2,3].* 1) NIH Undiagnosed Diseases Program, NIH, Bethesda, MD; 2) Medical Genetics Branch, NHGRI, NIH; 3) Office of the Clinical Director, NHGRI, NIH.

A major remaining problem in using next-generation sequencing is the substantial number of mismapped and misaligned sequences that lead to false-positive and false-negative errors when using alignment to a reference genome sequence as the method to reconstruct a patient's genome from short read sequences. This is due to the fact that the degree of difference between any one individual and the reference genome is no less than the difference between any one gene and an evolutionary homologue resulting from duplication and divergence that occurred within a family of genes. This is equivalent to low complexity since these closely related sequences require a longer read length to uniquely identify different locations, when compared to what is expected by a random sequence data. One way to address this is to use information, obtained independently and not from the short-read sequencing data, to improve the reference genome. The purpose of this modification is to more closely match the reference to the exome sequence being reconstructed. SNP chip data, haplotype data from the inheritance state of a nuclear family, and parental sequence data are all readily available sources of independent information that can be applied to resolve low-complexity region mapping and alignment ambiguities. We have established and tested a bioinformatic pipeline that takes SNP and parental sequence data from trios and finds the most likely alleles for the 4 independent haplotypes, generates a set of alignments, and then iterates these results for 3 cycles to improve the reference sequence used to generate the final alignment output for the patient. In addition simultaneous multiple sequence genotype calling, together with inheritance state-aware genotype calling, is used to produce the final variant list. This pipeline has produced measurable improvements from a simulated data set and for two trio data sets from a quartet sequenced in the Undiagnosed Diseases Program.

3544W

Kinship index variations among populations and thresholds for familial searching. *J. Ge, B. Budowle.* Dept Forensics and Investigative Genetics, Univ of North Texas Health Science Center, Fort Worth, TX.

Current familial searching strategies are developed primarily based on autosomal STR loci, since most of the offender profiles in the forensic DNA databases do not contain Y-STR or mitochondrial DNA data. There are generally two familial searching methods, Identity-by-State (IBS) based methods or kinship index (KI) based methods. The KI based method is an analytically superior method because the allele frequency information is considered as opposed to solely allele counting. However, multiple KIs should be calculated if the unknown forensic profile may be attributed to multiple possible relevant populations. An important practical issue is the KI threshold to select for limiting the list of candidates from a search. There are generally three strategies of setting the KI threshold for familial searching: (1) SWGDAM recommendation 6; (2) minimum KI $\geq$ KI threshold; and (3) maximum KI $\leq$ KI threshold. These strategies were evaluated and compared by using both simulation data and empirical data. The minimum KI will tend to be closer to the KI appropriate for the population of which the forensic profile belongs. The minimum KI $\geq$ KI threshold performs better than the maximum KI $\leq$ KI threshold. The SWGDAM strategy may be too stringent for familial searching with large databases (e.g., 1 million or more profiles), because its KI thresholds depend on the database size and the KI thresholds of large databases have a higher probability to exclude true relatives than smaller databases. Minimum KI $\geq$ KI threshold strategy is a better option, as it provides the flexibility to adjust the KI threshold according to a predetermined number of candidates or false positive/negative rates. Joint use of both IBS and KI does not significantly reduce the chance of including true relatives in a candidate list, but does provide a higher efficiency of familial searching.

3545F

Transparent, accessible, and robust functional analysis of SNPs. *B. Giardine, R. Burhans, C. Riemer, A. Ratan, R. Harris, G. Von Kuster, R.C. Hardison, Y. Zhang, W. Miller, Galaxy Development Team.* Center for Comparative Genomics and Bioinformatics, Pennsylvania State Univ, University Park, PA.

Human genetic variants, both common polymorphisms and rare mutations, are being discovered at an unprecedented rate. Genetic studies find associations between some of these variants and certain phenotypes, both simple and complex. However, interpretation of each variant, and especially finding a direct connection to a function in the phenotype of interest, is still difficult. A large number of computational tools have been developed to aid in SNP interpretation. Many of these are deployed within a powerful computational framework, called Galaxy, that facilitates analysis by any user and provides a transparent setting for sharing both results and the methods used to obtain them. We also provide resources, derived from ENCODE data, that are helpful in predicting functions of variants in non-coding regions.

Galaxy is a free computational service that allows users to run a wide variety of tools on their data. We have added tools to Galaxy for working with SNP data. These tools can be used for tasks such as finding phenotype-associated SNPs in a single genome, and comparing SNPs from case-control studies. To illustrate the power of these tools for SNP interpretation, suppose we begin with a well-known example of a SNP responsible (at least in part) for light skin color in humans of European ancestry (SLC24A5 A111T). This was originally mapped by genetic association in a case-control study {PMID 16357253}. We want to determine how many other variants, including rare and non-coding variants, have similar allele and haplotype distributions to this one. How close can we come to finding a strongly phenotype-associated allele without doing a full genome-wide association study? Using the genotype results from low-coverage 1000 Genomes data, we found approximately 4,000 variants that are fixed in Caucasians of European ancestry and for which the other allele is at very high frequency in Yorubans from West Africa. A set of over 2,500 showed clustering, consistent with regions of diminished variation indicative of a selective sweep. Of these, ~25 were in coding regions, including SLC24A5 A111T. Most of the clustered SNPs were non-coding, and many fell in DNA segments associated with gene regulation via ENCODE data (DNase hypersensitive sites and transcription-factor occupied sites). These are additional candidates for regions that could be involved in skin color. URL: usegalaxy.org.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3546W

ParseCNV Integrative CNV Association Software with Quality Tracking. *J. Glessner[1,2], H. Hakonarson[1,2].* 1) The Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, PA; 2) Department of Pediatrics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA.

A number of copy number variation (CNV) calling algorithms exist but comprehensive software tools for CNV association studies are lacking. We describe ParseCNV, software which takes CNV calls generated by any algorithm based on SNP array, array CGH, or sequencing and creates probe based statistics for CNV occurrence in cases and controls or family based studies with de novo and inheritance events. Case enriched significant probes are then separated from control enriched significant probes. Probe-based Fisher's exact test statistic output is then merged into CNVRs based on proximity and similar significance of neighboring probes. CNVRs are defined in a dynamic manner to allow complex CNV overlap and narrow association. The entire genome-wide dataset can be run at once without a static CNVR definition or repetitive SNP based statistics. Using this approach, we avoid failure to converge and non-monotonic curve fitting weaknesses of CNVtools (likelihood ratio trend test) and CNVassoc (latent class model). Plink (permutation-based test) only provides combined CNV state probe based statistics, not state specific CNVRs which are captured by ParseCNV statistics to increase power. In addition to p-value and odds ratio for each CNVR for all combined CNV state definitions, contributing samples, CN states, genomic features, and average number of probes of CNV calls are provided for confidence and biological interpretation. Contributing calls to each association, UCSC custom tracks, genotype and intensity values, and Plink .ped files are created for additional association review and characterization. In addition to up-front quality filtering, confidence scoring of significantly associated loci is done automatically based on genomic context, significance range, CNV overlap profile, number of probes underlying CNVs, external controls and inflated samples. 785 autism cases and 1,110 controls were assessed with 561,308 probes and resulted in 7 deletion and 12 duplication significantly associated CNVRs to demonstrate utility of the software. ParseCNV use has resulted in PCR validation rate of >80%, an often problematic stage due to inadequate significant association review. ParseCNV is freely available at http://parsecnv.sourceforge.net.

3547F

Web-based Interactive Visual Analysis of NGS Data using Galaxy. *J. Goecks[1], A. Nekrutenko[2], J. Taylor[1], The Galaxy Team.* 1) Biology and Math & Computer Science, Emory University, Atlanta, GA; 2) Center for Comparative Genomics and Bioinformatics, Penn State University.

Visualization is an essential scientific tool, making it possible to view large amounts of data simultaneously, identify patterns and outliers within data, and communicate findings to others. Using the Galaxy platform (http://galaxyproject.org), we have developed a Web-based visualization framework that addresses two challenges in visualizing next-generation sequencing (NGS) data: usability and integrated visual analysis. The Galaxy Visualization Framework (GVF) includes three visualizations: (1) Trackster, a genome browser; (2) Circster, a Circos-style genome-wide visualization; and (3) ParamaMonster, a visualization for exploring different parameter settings of an analysis tool. All visualization functionality is available using only a Web browser; no downloads/plugins are required. Composing visualizations is done using a graphical user interface, and even complex visualizations with many datasets can be created without any programming experience. Once created, visualizations can be shared with colleagues or published on the Web and included in publications' online supplementary materials. All shared and published visualizations are fully functional and can be accessed via a simple URL. By making it easy to create and share Web-based genomic visualizations, GVF makes genomic visualization more accessible and useful to investigators. Together, Trackster and ParamaMonster form a visual analysis environment that couples visualization and NGS analysis tools to enable interactive, user-friendly experimentation on NGS data. Trackster provides a fully-featured, Web-based genome browser that can be used to visualize most major genomic dataset formats. Any compatible Galaxy tool can be used in Trackster to analyze and visualize data in real time. Using Trackster, a tool's parameters can be set, the tool run, and the tool output visualized in a single, integrated interface. ParamaMonster enables visual exploration of a tool's complete parameter space, making it simple to visually compare outputs from many different combinations of parameter settings. We have used GVF to investigate interactions between structural variation and gene expression in pancreatic cancer. Trackster and ParamaMonster have been used for interactive RNA-seq analysis to find good parameter settings for assembling transcripts and to choose filters that remove assembly artifacts. Circster visualizations reveal regions where structural variation impacts gene and isoform expression.

3548W

Computational and informational challenges in providing clinically-relevant genome interpretation from high-throughput sequencing data. *R.K. Hart, Locus Development, Inc.* Locus Development, Inc., San Francisco, CA.

High-throughput genome and exome sequencing herald new opportunities to understand the genetic etiology of clinically-relevant conditions on an unprecedented scale. As the acquisition of sequencing data becomes cheaper and more routine, there is an increasing awareness that significant gaps remain in our ability to turn these data into meaningful clinical interpretations.

Locus Development provides sequencing and interpretation services to physicians from patient blood samples. Our value is based on three key assets: a high-quality database of associations of variants and conditions, custom sequencing assays, and computational analysis. Our panel is comprised of over 150 conditions in 250 genes and 10000 variants, curated by in-house geneticists and subjected to careful selection criteria. We use custom sequencing assays to target only these regions of interest and minimize costs. The entire process is substantially automated; we are currently targeting two weeks from the arrival of blood to the delivery of a clinical report.

In this presentation, we will describe the computational tools and pipeline that support our curation and interpretation efforts. We have integrated a complex assortment of external and in-house tools to provide a best-of-breed system that automates the analysis of reads, variant and haplotype calling, VUS analysis, clinical condition interpretation, and report generation. We will also discuss the suitability of publicly available data and tools.

3549F

Capture-Recapture Models for Evaluation of Algorithms Estimating Functionality of Missense Mutations. *S. Hicks[1], S.E. Plon[2], M. Kimmel[1].* 1) Department of Statistics, Rice University, Houston, TX; 2) Departments of Pediatrics and Human Molecular Genetics, Baylor College of Medicine, Houston, TX.

Whole genome and whole exome sequencing projects yield thousands of missense mutations with unknown effects on protein structure and function. Direct estimation of the sensitivity and specificity of bioinformatic algorithms such as SIFT or PolyPhen-2 which predict the impact of missense mutations on protein function requires a “gold standard” or a set of mutations with known functionality. In the absence of a “gold standard”, indirect statistical methods are needed to estimate the accuracy of these algorithms. Informative predictions depend on the algorithm and sequence alignment and in general the algorithms disagree as to which mutations are predicted to be deleterious or neutral. To investigate the level of agreement, we define disjoint categories of mutations depending on which algorithms predict which mutations to be deleterious or neutral. We develop two statistical models called Bernoulli Mixture (BM) and Augmented Bernoulli Mixture (ABM) based on the capture-recapture technique which use these disjoint categories. Application of these models allows to jointly estimate the sensitivities and specificities of each algorithm considered without the use of a gold standard and to estimate the proportion of deleterious mutations in a given set of mutations. When considering n algorithms, there are 2^n disjoint categories employed by the ABM model which includes $2n + 3$ parameters and the BM model is a special case of the ABM model which includes $2n + 1$ parameters. To estimate the parameters of the models, we use the Expectation-Maximization (EM) algorithm. We test the models using simulated and data-based predictions when a gold standard is available (Hicks et al. 2011). We perform simulation studies to evaluate convergence of the EM algorithm and to estimate bias and mean squared error (MSE) estimates which are reported as a function of sample size. Confidence intervals and confidence regions of the sensitivity and specificity estimates for each algorithm are reported. As an application, we estimate the sensitivity and specificity of the algorithms using mutations obtained from whole exome sequencing. We used 4 exome datasets sequenced at the Human Genome Sequencing Center at Baylor College of Medicine to identify cancer susceptibility genes for acute lymphocytic leukemia and lymphoma in children. Supported by CPRIT grant R83940, NCI grant CA155767 and NCI T32 training grant CA096520.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3550W

The Unified Sample Identifier - A universal sample coding system to manage large numbers of biological samples. *P. Hoffmann[1,2], J.E. Bolz[2], A. Reinscheidt[2], T.F. Wienker[4,5], S. Cichon[1,2,3], S. Herms[2].* 1) Institute of Human Genetics, University of Bonn, Bonn, Germany; 2) Department of Genomics, Life & Brain Center, University of Bonn, Bonn, Germany; 3) Institute of Neuroscience and Medicine (INM-1), Structural and Functional Organization of the Brain, Genomic Imaging, Research Center Juelich, Juelich, Germany; 4) Institute of Medical Biometry, Informatics and Epidemiology, University of Bonn, Bonn, Germany; 5) Gene Mapping Center, Max-Delbrueck-Center, Berlin, Germany.

Since the introduction of high-throughput DNA genotyping and sequencing technologies, the number of DNA samples analyzed for genetically complex phenotypes has tremendously increased. Furthermore it has become very important to combine samples from different sites to further increase statistical power to detect small genetic effects. This has led to serious challenges concerning the individual sample coding. Often encountered problems which hamper the unambiguous identification of individual samples include samples having the same identifier by chance as well as spelling errors in the sample identifier. To address these problems we have developed a universal sample coding system the Unified Sample Identifier (USI) over the last years. The USI is based on the ISO certified IBAN system which is used in the international banking system since more than 30 years. The USI includes a checksum making it very resistant to spelling errors and data about project, study, sample type and aliquot, but does not include any information on the affection status, diagnosis, gender and ethnicity. Here we present data on the implementation of the USI in our database system, the use of automatically generated barcodes and the evaluation of the checksum system. We especially analyzed the probable risk of collisions - different lab codes sharing the same check sum - in the normal lab context. We checked for random one letter changes first and focused on common human letter recognition issues in a second analysis. Based on these results we are further optimizing the algorithms for automatic correction of coding errors.

3551F

Systematic Evaluations of Sequencing Errors in Next-Generation Sequencing Data. *Y. Hu[1], Y. Liu[2], J. Ferguson[3], I. Silverman[4], B. Gregory[4], M. Reilly[3], M. Li[2].* 1) Department of Mathematics, University of Pennsylvania, Philadelphia, PA; 2) Department of Biostatistics and Epidemiology, University of Pennsylvania, Philadelphia, PA; 3) Cardiovascular Institute, University of Pennsylvania, Philadelphia, PA; 4) Department of Biology, University of Pennsylvania, Philadelphia, PA.

The emergence of next-generation sequencing technologies has revolutionized biological research. Next-generation sequencing involves the parallel sequencing of millions of DNA fragments simultaneously. Generally, these fragments are sequenced one base at a time in which the base is determined through fluorescent detection. Incorrect base calls can have a deleterious impact in sequence alignment, variant calling, and RNA editing detection. Although several studies have shown that base-call errors are not randomly distributed in sequenced reads, the error distributions and the impact of errors have not been thoroughly evaluated. By utilizing paired-end RNA sequencing data on three individuals sequenced using Illumina's HiSeq 2000 (4 RNA-Seq samples per individual), we performed a detailed characterization of sequencing errors. We generated 28M to 69M 101 bp paired-end data with 57% to 74% mapped uniquely to the reference genome. Our analysis was based on the observation that sequencing errors will leave a footprint in the base calls at overlapping sites of a mate pair. Specifically, a sequencing error was declared if one of the calls at a site agreed with the reference genome, but the other in the mate-pair read did not. Since the same individuals were sequenced multiple times, we can also identify systematic errors, which manifest as individual base-call errors from sequencing reads generated from independent RNA-Seq experiment occurring at the same genomic position. On average, each RNA-Seq sample had 29,287,719 sites covered by at least one base-call pair, among which 457,933 had errors, yielding an error rate of 0.43%. The most frequent error type is A->C (20%), followed by T->G (14%) and T->C (11%). We noticed that many of the error bases had low quality. If restricting analysis to bases with quality score ≥30, then only 10,698 error sites remained, yielding an error rate of 0.011%. Since variant detection and RNA editing analysis typically consider high-quality bases, we focused our systematic error detection on those bases only. On average, we identified 450 sites showing systematic errors with A->C (41%) being the most common, followed by T->G (13%) and A->G (12%). The error sites did not overlap with RNA editing sites identified using the same data (unpublished). Our analysis suggests that systematic sequencing errors are unlikely to confound with RNA editing sites if the editing sites were identified using a carefully designed pipeline.

3552W

EGprofile: a tool for rapid profiling analysis of epigenetic marks with ChIP-seq data. *W. Huang, L. li.* NIEHS, NIH, Research Triangle Park, NC.

Epigenetic environment or epigenome plays critical roles in chromatin remodeling and gene transcription regulation. Changes in epigenome (e.g., histone acetylation and methylation) can be detected by changes in enrichment level of a variety of epigenetic marks of histone modifications. Genome-wide survey of changes of epigenetic marks with high-throughput next-generation sequencing has now been widely used for gene regulation control studies, e.g., the ENCODE project. To facilitate such studies, we developed an efficient tool EGprofile that can be used to perform rapid profiling analysis of epigenetic marks using ChIP-seq data. Our tool can generate the empirical distribution of an epigenetic mark for a group of genomic regions/sites with a common genomic feature (e.g., TSS, TFBS), or a read coverage plot of any single genomic region. It can also automatically normalize data to enable cross-comparison of epigenetic mark profiles of different genomic regions/features. In addition, our tool can determine optimized region size for easy visualization when epigenetic marks having a single-mode distribution. Furthermore, EGprofile supports all popular read align file formats including SAM, BAM, MAQ, and UCSC BED files, and it is efficient with the typical running time less than 10 minutes for a single-lane Illumina ChIP-seq data. The core part of EGprofile was implemented with C++, and it will be available as an independent R package.

3553F

SOAPfuse detects gene fusions from paired-end RNA-Seq data with single base resolution. *W. Jia[1,2], K. Qiu[1], M. He[1], P. Song[1], Q. Zhou[1,2,3], F. Zhou[1,2,4], X. Hu[1], Y. Li[1], G. Guo[1,2].* 1) BGI-Shenzhen, Shenzhen, China; 2) BGI-Americas, Cambridge, USA; 3) School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu, China; 4) School of Bioscience and Bioengineering, South China University of Technology, Guangzhou, China.

We developed an open-source software tool, SOAPfuse, which detects gene fusions with single-base resolution from paired-end RNA sequencing (RNA-Seq) data. SOAPfuse applies several pioneering algorithms to overcome deficiencies of other widely-used fusion-detecting software tools. The most important component in our method for detecting fusions is SOAPfuse's local exhaustion algorithm, which is used for the highly efficient construction of a putative junction library. Moreover, by combining an iteratively trimming-realigning procedure with several strong filtrations, SOAPfuse achieves good detectability and high accuracy. We compared SOAPfuse with other tools, including chimerascan, deFuse, FusionHunter and TopHat-Fusion, based on both actual and simulated datasets. Actual data are from two published studies, which have several validated fusions. Compared with other tools, SOAPfuse consumed the least amount of computing resources (CPU-time and memory) to redetect the most validated fusions. We simulated 150 fusions, and generated RNA-Seq data at nine depth levels (5–200X) using MAQ. SOAPfuse showed the lowest false negative and false positive rates at all depths. We also applied SOAPfuse to RNA-Seq data from two bladder cancer cell lines, and identified 16 fusions, in which 15 (93%) are validated successfully. We reanalyzed this dataset using deFuse, which detected only nine out of the 15 validated fusions. We also found six pairs of recurrent fusions, some of which show strong signals, therefore inferring their structure variation sources. HADHB-RBKS is formed by genes from different strands, indicating an inversion. CIRH1A-TMCO7, PSMD8-SIPA1L3 and TIAM1-ATP5O are formed by intrachromosomal same-strand genes in their reversed genomic orientations, showing potential translocations. Interestingly, these recurrent fusions are not reported by previous studies on bladder cancer. To our knowledge, SOAPfuse is the only tool that uses a local exhaustion algorithm. SOAPfuse runs quickly and saves computing resources, making it suitable for analyzing vast samples in parallel. SOAPfuse not only identifies genuine fusions effectively at single-base resolution, but also provides a practical means to explore the potential structural variations that transcript to fusions. Our work offers a novel and efficient fusion detection tool for research on credible fusions. SOAPfuse is freely available from http://soap.genomics.org.cn/soapfuse.html.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3554W

Personalized genomics: personalized pathway mapping. *P. Jia[1], Z. Zhao[1,2].* 1) Biomedical Informatics, Vanderbilt University, Nashville, TN; 2) Department of Cancer Biology, Vanderbilt University, Nashville, TN.

The rapid application of next generation sequencing (NGS) in biomedical research revolutionized personalized medicine, but numerous challenges exist. Currently, the application of traditional pathway enrichment methods to next-generation sequencing data is prone to several potential biases including (1) genomic/genetic factors (e.g., the particular disease and gene length) and (2) environmental factors (e.g., personal life-style and frequency and dosage of exposure to mutagens). Therefore, new methods are urgently needed for these new data types, especially for individual-specific genome data. We proposed a novel method for the pathway analysis of NGS mutation data by explicitly taking into account the gene-wise mutation rate. We estimated the gene-wise mutation rate based on the individual-specific background mutation and the gene sequences. Taking the mutation rate as a weight for each gene, our method implements the weighted resampling strategy and builds the null distribution for each pathway while matching the gene-wise mutation patterns. The empirical P value obtained then provides an adjusted statistical evaluation. We demonstrated our method in a lung adenocarcinomas dataset and a glioblastoma dataset, and compared it to other widely applied methods. Our method could effectively reject many marginally significant pathways detected by standard methods, including several long gene based, cancer unrelated pathways. We further demonstrated that by reducing such biases, pathway crosstalk for each individual and pathway co-mutation map across multiple individuals can be objectively explored and evaluated. This method performs pathway analysis in a sample-specific fashion, and provides an alternative way for accurate analysis of personalized cancer genomes. We further extended this approach to network based analysis in personalized genomics and found promising results as well.

3555F

Correcting for expression heterogeneity while identifying regulatory hotspots. *J. JOO[1], J. Sul[2], B. Han[2], E. Eskin[2].* 1) Bioinformatics IDP, UCLA, Los Angeles, CA; 2) Computer Science, UCLA, Los Angeles, CA.

Expression quantitative trait loci (eQTL) mapping is a powerful tool that can systematically identify genetic variations affecting gene expression. From eQTL mapping studies, it has been known that thousands of genes are trans-regulated by eQTLs referred as ``regulatory hotspots". However, recent studies have shown that various confounding effects such as a batch effect induce heterogeneity in expression data and may cause spurious associations including spurious regulatory hotposts. Several computational methods have been developed to correct for the expression heterogeneity. However, the challenge in previous methods is that in an attempt to eliminate confounding effects, the genetic effects are also eliminated. Here, we introduce a novel method, that attempts to eliminate spurious regulatory hotspots while retaining hotspots caused by genetic effects on a new statistical framework. Our method is based on a rationale that confounding factors affect most of the genes while genetic factors affect fewer genes. Our method leverages a recently developed statistic to detect genes that do not have genetic effects but only confoundinf effects. Using those genes, we are able to remove confounding effets but preserve genetic effects. Applying to both simulation and real data, we showed that our method successfully corrects for expression heterogeneity while identifying regulatory hotspots.

3556W

GeneTalk: an expert exchange platform for assessing rare sequence variants. *T. Kamphans[1], P.M. Krawitz[2].* 1) Smart Algos, Berlin, Germany; 2) Charité, Berlin, Germany.

Next-generation sequencing (NGS) has become a powerful tool in personalized healthcare and is now starting to transform our society. The 1000 $ genome is within range and the really challenging question is: What is going to happen if Joe Consumer starts analyzing his own genome? GeneTalk is the first professional social network website for genome analysis and is currently growing faster than facebook. Right now, it is a networking tool used by about 10% of Europe's geneticists for expert based interpretation of family specific human sequence variants. This professional user base contributed mainly scientifically proven annotations about rare genetic disorders to GeneTalk's wiki-like knowledge base. Starting this year the first laymen joined the portal and showed a completely different user behavior. A first evaluation of the crowdsourced annotations from users searching and commenting their personal genomes illustrates that sequence variants associated with food intolerance, obesity and metabolism of recreational drugs were frequently searched. The genetic counseling of the years to come will certainly have to address questions that arise by private genome analysis.

3557F

UCSC Genome Browser 2012: distributed data, enhanced interactivity, variant annotation integrator. *W.J. Kent, A. Hinrichs, D. Karolchick, A. Zweig, B. Raney, H. Clawson, M. Cline, L. Guruvadoo, K. Learned, R. Kuhn, B. Rhead, T. Dreszer, L. Meyers, C. Li, M. Diekhans, G. Roe, P. Fujita, D. Haussler.* Dept Biomolecular Engineering, Univ California, Santa Cruz, CA.

The UCSC Genome Browser and database has improved in several ways over the past year. Here we will describe three new features - track hubs for distributed annotations, Javascript-based improvements to make the browser more interactive, and the development of a new tool aimed at human geneticists - the Varian Annotation Integrator. Our new Track Data Hub systems allows users to develop large scale annotation sets that can be viewed in context of other annotations on the browser. These hubs consist of files in standard distributed file formats such as BAM, BigWig, BigBed, and VCF placed in a directory on the users own web site, along with simple text formatted files that say how to label, color, and arrange the tracks in the hub. The data from hubs is fetched as needed in parallel, and cached at UCSC, with the result that the track hubs perform nearly as well as built in tracks, and allow for more permanence and flexibility than the existing custom track mechanism. We've used the JQuery library and local JavaScript development to increase the interactivity of the browser in several ways. It's now possible to rearrange the order of tracks, and of subtracks within tracks, by drag and drop. Drag scrolling is also now possible, and drag-zoom has been enhanced. Right click (control click on the Mac) now brings up a menu with context sensitive options depending on where you click. We've also worked to enhance the speed of the site. We are developing a new tool, the Variant Annotation Integrator. This tool lets you intersect variants in VCF and other formats with the gene and other annotation tracks at UCSC. This allows the user to see which sequence variants disrupt genes, regulatory features, or are just SNPs seen commonly across the human population.

3558W

Curating genomic epidemiology data in The PAGE Study. *G. Kumaraguruparan[1], G. Mehta[1], A.Q. Nato[2], J.L. Ambite[1], S. Buyske[3], R. Mayani[1], C. Cai[1], J.S. Vockler[1], E. Deelman[1], T.C. Matise[2].* 1) Information Sciences Institute, University of Southern California, Marina del Rey, CA 90292; 2) Department of Genetics, Rutgers University, Piscataway, NJ 08854; 3) Department of Statistics and Biostatistics, Rutgers University, Piscataway, NJ 08854.

The “Population Architecture using Genomics and Epidemiology” (PAGE) study investigates the epidemiologic architecture of genetic variants associated with complex diseases in large, ethnically diverse population-based studies. PAGE is genotyping 200–200,000 genetic variants/year in tens of thousands of European American (EA), African American, American Indian, Hispanic/Mexican American, and Asian/Pacific Islander participants. PAGE is addressing generalization of associations and variability of effects in diverse populations; identifying genetic and environmental modifiers; and investigating associations with novel phenotypes. The PAGE Coordinating Center (CC) warehouses and disseminates the results of the PAGE Consortium. Using predefined templates, studies submit SNP, phenotype, and association results files to the CC for data quality control and submission to dbGAP. A challenge of working with multiple institutions is the differences in which similar data are recorded, such as gender being either `Male' or `M,' or reporting proportions vs. percentages. To automatically review the PAGE results data, we provide a web-based Quality Control (QC) and Standardization system that checks whether the submitted data adhere to the prescribed templates and data formats, standardizes the various fields to the prescribed formats, and uploads them to a database backend for advanced quality checks and submission to dbGaP. The backend models the QC steps as a workflow deployed by the Pegasus Workflow Management System (WMS) and supported by a relational database to store and process data. Over 25 QC steps rigorously check the data to ensure that it is complete and semantically correct. In addition, QC ensures that data are reported according to a common strand orientation, which facilitates meta-analyses. Pegasus WMS can manage and execute PAGE workflows on local and distributed resources such as campus clusters or clouds. The system can run several QC workflows simultaneously allowing several studies to submit large amounts of data to be curated in a short time. Pegasus publishes the results of QC to a web GUI that displays the status of the QC steps that are running and reports errors at various stages with appropriate logs and messages so that the submitting scientists can correct and re-upload their data as needed. The PAGE QC workflow provides an efficient approach to data cleaning for a multicenter study using common data templates.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3559F

Predicting causal variants in exome sequencing of Mendelian disorders. *J. Kwan[1,2], M.X. Li[2,3,4], P.C. Sham[2,3,4].* 1) Department of Medicine, University of Hong Kong, Hong Kong, China; 2) Department of Psychiatry, University of Hong Kong, Hong Kong, China; 3) Centre for Reproduction, Development and Growth, University of Hong Kong, Hong Kong, China; 4) Genome Research Centre, University of Hong Kong, Hong Kong, China.

A major challenge of exome sequencing of Mendelian disorders is to pull out the causal variant(s) from a list of candidate mutations. Because the most common cause of Mendelian disease is a non-synonymous single-nucleotide variant (nsSNV) that results in a single amino acid change in the protein encoded by that gene, current strategy mainly relies on removing variants conflicting with the disease inheritance mode and/or catalogued in SNP databases, and then predicting the functional consequences of the nsSNVs left on disease. Here, we examine the actual performance of pathogenicity prediction in the analysis of exome sequencing data and propose a combined model tailored for predicting Mendelian disease-causing nsSNVs in exome sequencing based on the scores from multiple prediction algorithms. When used to distinguish Mendelian disease-causing mutations from other rare variations, our full combined model (based on the scores from five different algorithms, including SIFT and Polyphen version 2) outperforms the five individual algorithms and another well-known combined model, Condel (which is being used in Ensembl's Variant Effect Predictor). In addition, our reduced model (based on a subset of the scores from the five algorithms) has an even better predictive performance than our full model. However, no method is able to distinguish autosomal dominant disease mutations from autosomal recessive disease mutations. Since we estimate that human genomes typically contain ~100–200 genuine Mendelian disease-causing nsSNVs in heterozygous state, a method to classify Mendelian disease-causing mutations as autosomal dominant or autosomal recessive will be useful for exome sequencing studies.

3560W

Analysis of Ewing Sarcoma NGS transcriptome data highlights mechanisms of cancer progression. *M. Laurance, J. Billaud, C. Bullitt.* Ingenuity Systems, Redwood City, CA.

The Ewing Sarcoma family of tumors is a category of cancers that mostly affects teenagers of age 10 to 20. To understand the molecular mechanisms of tumorigenesis and metastatic processes in Ewing Sarcoma (ES), gene expression profiling of samples from Ewing Sarcomas patients was performed using Next-Generation-Sequencing (NGS) technology. In this study we present the results of combined bioinformatics and biological analyses that demonstrate the differences between samples from primary and metastatic tissues from one representative patient with ES. Initial statistical analysis of the RNA-Seq data was performed with SeqSolve software from Integromics. Ingenuity iReport was then used to analyze and understand the functional implications of the RNA-Seq data and identify molecular mechanisms participating in the development of ES. Using iReport we identified differentially expressed isoforms and highlighted their potential roles in tumor progression and/or metastatic processes. We also characterized and identified signalling and metabolic pathways (such as glucose and fatty acid metabolism) and biological processes (such as resistance to anoikis) that may be involved in the progression of ES. Subsequent transcription factor analysis of the RNA-Seq data using Ingenuity's IPA program enabled prediction of the activity state of several transcription factors (such as SREBF1, PPARGC1B, IRF3 and IRF7) that may contribute to the pathological context observed in this patient. IPA transcription factor networks were used to visualize potentially induced transcriptional programs in the metastatic process. Using complementary analysis capabilities in iReport and IPA we were able to highlight important biological processes and underestimated contributors to tumorigenesis (such as metabolic pathways) and identify key isoforms and molecular players in ES tumor progression (enzymes, transcription factors). The data presented here is from one patient, therefore similar studies are needed to understand if these insights are supported by data from other Ewing Sarcoma's patients and also to extend these insights by identifying miRNAs with roles in progression and mechanisms of metastasis in ES.

3561F

Low concordance of variant calling algorithms in exome sequencing. *G. Lyon[1,7], T. Jiang[2], G. Sun[2], W. Wang[3], J. Hu[2], P. Bodily[4], L. Tian[5], B. Moore[6], H. Hakonarson[5], J. Wang[2], M. Yandell[6], E. Johnson[4], Z. Wei[3], K. Wang[8].* 1) Stanley Institute for Cognitive Genomics, Cold Spring Harbor Laboratory, NY; 2) BGI-Shenzhen, China; 3) Department of Computer Science, New Jersey Institute of Technology, Newark, NJ; 4) Department of Medicine, Boston University School of Medicine, Boston MA; 5)) Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, PA; 6) Eccles Institute of Human Genetics, University of Utah, Salt Lake City, UT; 7) Utah Foundation for Biomedical Research, Salt Lake City, UT; 8) Zilkha Neurogenetic Institute, Department of Psychiatry and Preventive Medicine, University of Southern California, Los Angeles, CA.

There is a growing gap between the generation of massively parallel sequencing output and the ability to process, analyze and interpret the resulting data. New users to the sequencing era are left to navigate a bewildering maze of base calling, alignment, assembly and analysis tools. For the implementation of genomic analysis in the clinic, it will be critically important to optimize and standardize one overall pipeline with high sensitivity and specificity for variant calling. One must first have sufficient depth of sequencing coverage as well as high quality of sequencing data to tackle this problem. We optimized our depth of coverage in a pilot set of exomes to a depth of at least 20 reads per base pair in greater than 80% of the target region, captured using Agilent SureSelect v.2 and subsequently sequenced on the Illumina platform. We demonstrate that there is poor concordance (~60–70%) even for single nucleotide variants (SNVs) among the five variant calling algorithms used (SOAP, GATK, SNVer, GNUMAP, and SAMtools) under default parameters; concordance for indels is even worse, typically between ~25–50% for SOAPindel, SAMtools and GATK. Furthermore, concordance for novel variants is significant lower than common variants. While some researchers have suggested sequencing the same sample on two separate sequencing platforms (e.g. Illumina and Complete Genomics) and using only shared variants as determined by one variant calling platform (GATK), this may be impractical due to storage and economic considerations. As an alternative, we recommend using variants called by two or more variant calling approaches on one set of sequencing data. Additionally, we discuss the different expectations and tolerance to false positives in a research setting versus in a clinical setting, and the need to apply different filtering strategies in different settings. Finally, on a biological level, these 15 exomes are carefully selected from four large families with many symptoms of Tourette Syndrome, attention deficit hyperactivity disorder, and obsessive compulsive disorder. We are using various tools such as ANNOVAR and VAAST to prioritize variants from the families for follow-up in case-control and biological studies.

3562W

DIVERGENOMEnrich: Expanding information on a Genetic Variation database through automatic online data retrieval. *W. Magalhaes[1], G. Souza[1], G. Kingman[2], E. Tarazona[1], M. Rodrigues[1], Epigen-Brazil.* 1) Department of General Biology, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil; 2) Departamento de Ciência da Computação, Pontifícia Universidade Católica de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil.

Large-scale genomics initiatives such as the HapMap project and the 1000-genomes rely on powerful bioinformatics support to assist data production and analysis. Contrastingly, few bioinformatics platforms oriented to smaller research groups exist to store, handle, share, and integrate data from different sources, enrich genetic diversity analyses, as well as to assist these scientists to perform their analyses efficiently. We developed such a bioinformatics platform, DIVERGENOME, to assist population genetics and genetic epidemiology studies performed by small- to medium-sized research groups. The platform is composed of two integrated components, a relational database (DIVERGENOMEdb), and a set of tools to convert data formats as required by popular software in population genetics and genetic epidemiology (DIVERGENOMEtools). In DIVERGENOMEdb, information on genotypes, polymorphism, laboratory protocols, individuals, populations, and phenotypes is organized in projects. Here we present a new component, DIVERGENOMEnrich, which allows users to achieve relevant genetic and biomedical knowledge, enriching variants with other types of biological data, such as biochemical pathways, pharmacogenetics, and genome wide association studies even though these data are fragmented in different databases.. To implement this feature we propose an Agent-based enrichment system that integrates different sources of biologically relevant data in order to improve the genetic diversity analysis of native and Latin-American populations in particular. In our system, agents are autonomous programs that perform a specific task and communicate with each other in order to combine information from different sources. We believe that an agent-based approach is suited to implement such enrichment system given the heterogeneity of different types of genetic and biological data and their distribution over distinct sources over the Web.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3563F

Fast and accurate local ancestry inference on whole genome-scale data using conditional random fields. *B. Maples, F. Zakharia, S. Gravel, E. Kenny, C. Bustamante.* Genetics, Stanford University, Stanford, CA.

Humans show assortative mating based on a diversity of geographic, linguistic, cultural, and biological traits. As a result, most humans trace their origins to two or more statistically distinct groups. Using genome-wide data, it is often possible to infer not only the populations from which an individual derives, but also the population from which individual loci along her genome originate. From a medical genomics perspective, such local ancestry information enables phenotype-to-genotype associations through admixture mapping for traits that correlate with ancestry. In population genetics, local ancestry tracts can shed light on the natural history of both the ancestral populations and the admixture process. The increasing availability of whole genome data provides unbiased ancestry informative markers and haplotypes to enable the inference of local ancestry at a sub-continental scale. However, most local ancestry inference methods are highly computationally intensive, making them difficult to apply to the number of markers and reference individuals needed to perform accurate genome-scale inference. In light of these limitations, we propose a fast, accurate statistical method using conditional random fields parameterized by random forests to classify genomic diversity and perform local ancestry inference. We compared an implementation of this method, called RFMix, to local ancestry inference results performed by different groups using recent data from the 1000 Genomes project and a diversity of widely-used inference methods, including HapMix, LAMP-LD and MULTIMIX. Accuracy of the method was 99.4% when inferring ancestry tracts on simulated African American samples, and 96.6% when applied to simulated Mexican individuals. Additionally, RFMix is similarly accurate but 10 times faster than the next fastest method, LAMP-LD (with S=10) . This new approach can be readily parallelized for multi-processor machines and enables fast and accurate inference for large-scale studies.

3564W

Exome Sequencing of Mayo Clinic Biobank samples for evaluation of quality and quantity of data. *S. Middha, SK. McDonnell, ZC. Fogarty, MS. DeRycke, KJ. Johnson, NM. Lindor, DJ. Schaid, JE. Olson, JR. Cerhan, SN. Thibodeau.* Mayo Clinic, Rochester, MN.

BACKGROUND: Whole exome sequencing (WES) is a powerful next generation sequencing technology that is mainly used in the research setting, but is rapidly moving towards the clinic. In order to better understand quality and reporting metrics, we conducted a study of 39 exomes from deceased participants in the Mayo Clinic Biobank. METHODS: The Mayo Clinic Biobank is a research resource that has enrolled over 20,000 Mayo Clinic patient volunteers since 2009. Participants are enrolled regardless of their health history, and provide biologic specimens, complete a health history, and allow access to their medical record. For this study, we selected participants who had died after enrollment, did not have blood cancer, and had a detailed medical history at Mayo. We used the Agilent 50Mb capture kit and sequenced on the Illumina HiSEQ 2000 using 100bp paired-end reads on one lane per sample. An in house analysis and annotation workflow (TREAT) was used to generate variant reports. Variants were called Tier I if they were nonsense, loss of stop or splice site variants; and Tier II if they were missense variants. Samples were also genotyped using the Illumina 5M SNP chip in order to verify data quality. RESULTS: A total of 202,064 SNVs (21,477 genes) and 19,832 indels were identified using SNVMix and GATK respectively. This includes 48,528 novel SNVs in 15,205 genes not already reported in dbSNP or 1000 Genome Project and 743 unique nonsense variants in 674 genes. In aggregate, there were 1,260 Tier I SNVs, 1,774 Tier I indels, and 43,529 Tier II SNVs. On a per subject basis, there were a mean of 271 Tier I SNVs, 388 Tier I indels, and 10,792 Tier II total variants (including known and novel variants). All samples had excellent concordance (more than 99 percent) between WES and the Illumina 5M SNP chip. The average depth of sequencing coverage was excellent across the entire exome (more than 80 percent of targeted region at 20x), but was quite variable at the gene level. CONCLUSIONS: Whole exome sequencing results conducted on blood of 39 deceased participants of the Mayo Biobank are described in detail. WES shows very high concordance with SNP genotyping. Establishing the frequency and types of variants likely to be found in our target clinical population facilitates preparation for more wide-spread clinical use of this technology. Correlation of molecular findings with clinical information may help interpret some of the vast number of variants detected.

3565F

Genome wide association study on the world fastest supercomputer, k computer. *K. Misawa[1], A. Hasegawa[1], T. Tsunoda[1,2].* 1) Research Program for Computational Science, Riken; 2) Center for Genomic Medicine, Riken.

More than a million single-nucleotide polymorphisms (SNPs) are analyzed in any given genome-wide association study (GWAS), therefore performing multiple comparisons can be problematic. One attempt to cope with multiple-comparison problems in GWAS was the development of haplotype-based algorithms to correct for multiple comparisons at multiple SNP loci in linkage disequilibrium (Misawa et al. 2008, J. Hum. Genet 53: 789-). In addition, permutation tests were found to also control problems inherent in multiple testing (Kimmel and Shamir 2006, Am. J. Hum. Genet. 79: 481-). However, both the calculation of exact probability and the execution of permutation tests are time-consuming and faster methods are required. To this end, we developed a suite of computer programs, ParaHaplo 4.0. ParaHaplo 4.0 can conduct parallel computation of accurate P-values in haplotype-based GWAS through use both Misawa et al.'s (2008) algorithm and Kimmel and Shamir (2006) algorithm. Our program is intended for workstation clusters and supercomputers using the Intel Message Passing Interface (MPI). GWAS requires a huge amount of data, such that file input/output can become a bottleneck in analysis. Thus, we implemented distributed file input/output in ParaHaplo 4.0. We compared the performance of ParaHaplo 4.0 on the HapMap dataset for Japanese in Tokyo, Japan and Han Chinese in Beijing, and Chinese. Computation was conducted on k computer, which is the world fastest computer as of November 2011. The parallel version of ParaHaplo 4.0 conducts genotype imputation much faster than the non-parallel version of ParaHaplo. ParaHaplo executable binaries and source code are available at http://en.sourceforge.jp/projects/parallelgwas/releases/.

3566W

A bioinformatics approach for the identification of developmental QTL candidate genes. *A.Q. Nato[1], B. Li[2], F. Chen[1,3], J.H. Millonig[1,2,4], T.C. Matise[1].* 1) Department of Genetics, Rutgers University, Piscataway, NJ; 2) Center for Advanced Biotechnology and Medicine, UMDNJ-Robert Wood Johnson Medical School, Piscataway, NJ; 3) MedStar Health Research Institute, Washington, DC; 4) Department of Neuroscience and Cell Biology, UMDNJ-Robert Wood Johnson Medical School, Piscataway, NJ.

Neural tube defects (NTDs) occur in ~1/1,000 live births in the United States and are considered one of the most common birth defects. Approximately 20.5 million Americans over age 40 have cataract in either eye. We have formerly established a spontaneous mouse mutant, *vacuolated lens* (*vl*), as a multigenic model for cataract and NTDs. *Vl* arose on the C3H/HeSnJ background and exhibits cataract and NTDs with 100% penetrance. A mutation in an orphan G protein coupled receptor, *Gpr161*, which is expressed in the developing neural tube and lens, causes these disorders. However, the severity and incidence of the *vl* phenotypes are modified when the mutation is crossed to C57BL6/J, CAST/Ei, or MOLF/Ei backgrounds indicating that unlinked loci can bypass the effect of the *Gpr161* mutation. Quantitative trait locus (QTL) analysis mapped the position of five modifier loci (Modifiers of vacuolated lens: *Modvl1–5*). Several resources are available for the identification of candidate genes for adult QTL but not for developmental QTL. For this study, we developed a bioinformatics pipeline specifically designed for the identification of developmental QTL candidate genes. In-house computer scripts were created to mine the NCBI Entrez Gene, expressed sequenced tag (EST), and Gene Expression Omnibus (GEO) databases, and to identify genes within the *Modvl* 95% confidence intervals expressed at particular developmental stages. We utilized the dataBase for Gene Expression Evolution (Bgee) to determine which of these developmentally expressed genes are localized to the developing CNS and eye. dbSNP data, resequencing, and bioinformatics analysis were then employed to identify coding polymorphisms in these CNS- and eye-restricted transcripts and determine whether they are predicted to be functional. Finally, we identified pathways that are most significantly associated with these developmentally restricted transcripts and may therefore be relevant to neural tube closure and lens development. We have successfully identified candidate genes for each *Modvl* locus through our new bioinformatics pipeline that allows EST, microarray data, spatial/temporal expression, and dbSNP data to be searched successively. This pipeline (http://compgen.rutgers.edu/devQTL) may be utilized by other investigators to identify candidate genes for developmental QTL.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3567F

Multivariate profiling approach to inference on differential expression in RNA-Seq data with small sample size. *S. Oh[1], M. Kim[2], S. Song[3].* 1) Division of Human Genetics, Department of Pediatrics, Cincinnati Children's Hospital Medical Center Research Foundation, Cincinnati, OH; 2) Division of Epidemiology and Biostatistics, Department of Pediatrics, Cincinnati Children's Hospital Medical Center Research Foundation, Cincinnati, OH; 3) Department of Mathematical Sciences, McMicken college of Arts and Sciences, University of Cincinnati, Cincinnati, OH, 45221.

Identifying differential expression is one of the most widely used downstream statistical analyses of genomic data. We propose a multivariate profiling approach to the identification of differential expression in RNA-Seq data with small sample size. RNA-Seq is a powerful approach to transcriptome profiling that uses deep-sequencing technologies and is fast replacing microarray with superior sensitivity and precision, dynamic resolution, and isoform level quantification. Contrary to its gaining popularity, cost to sequencing is still expensive and often only a small number of samples for biological replications are included in individual experiments. The proposed fully bayesian multivariate approach will overcome limitations rising from the small sample size issue by pooling information and correlation structure across genes and samples. This idea of multivariate profiling is well justified by well-known facts that genes work collaboratively in a biological system as a network module and the functional connections are manifested as correlated structure of transcriptional regulation in gene expression among genes and among relevant tissue samples. Many multivariate approaches have successfully proposed for the analysis of microarray data. We will adopt them by addressing the discrete nature of the RNA-Seq data. The proposed approach will allow both identification and ordering of differential expressions. We will compare the proposed multivariate approach with gene-by-gene univariate approaches both in real data analysis and Monte Carlo simulation study. Keywords: RNA-Seq, differential expression, univariate/multivariate profiling approach, and Monte Carlo simulation study.

3568W

Challenges and approaches to computational candidate gene prioritization using gene networks. *P. Pavlidis[1], E. Mercier[1], J. Gillis[1,2].* 1) Department of Psychiatry and Centre for High-Throughput Biology, University of British Columbia, Vancouver, Canada; 2) Cold Spring Harbor Laboratory, NY.

It has long been hoped that computational methods will assist human genetics studies with sophisticated variants of “guilt by association” (GBA), based on the exploitation of information from gene networks. In principle such methods have wide application. Genome wide analyses of human genetic traits and disorders are often faced with the challenge of prioritization of findings for follow-up. For example, analysis of copy number variation results would be simplified by being able to choose which of the genes covered by the variant are most important to the phenotype. In addition, researchers planning focused studies could use computational methods to prioritize targets. In previous work, we found limitations of GBA methods to give specific results (Qiao Y et al., Hum Genet. 2010 128:179) and showed that such methods are affected by severe biases induced by genes which are highly “multifunctional” (Gillis J and Pavlidis P, PLoS Comput. Biol. 2012 8(3):e1002444 and PLoS One. 2011 6(2):e17258). We predict methods that do not account for multifunctionality will tend to prioritize genes which are either already highly annotated, or which have specific associations with highly-annotated genes. The result is that the same genes will tend to be prioritized for numerous phenotypes, simply due to biases in the underlying gene networks and annotations. Here we show that the non-specificity reported by Qiao et al. (2010) can indeed be accounted for by multifunctionality effects. We also present a novel approach for prioritization called Favourite Gene-Favourite Function analysis (FGFF) that considers not just the association of a gene with the phenotype or function of interest, but also the degree to which that phenotype is the one most specifically predicted to be associated with the gene. Thus each gene “votes” for its “favourite function”, and only if that function is the one of interest to the investigator is that gene considered a strong candidate. We describe several beneficial features of the approach based on analysis of large human protein interaction networks. FGFF dramatically reduces the degree to which the same gene is predicted for multiple functions, and it reduces the impact of highly annotated “hub” genes. Extensions of the approach to tissue-specific gene coexpression networks and other data types are under way and will be discussed. Our methods provide a way to increase the utility of computational gene prioritization for genetics research.

3569F

BlueSNP: An R Package for High-Scalability Genome-Wide Association Studies on Compute Clusters. *R.J. Prill[1], H. Huang[2,3], S. Tata[1].* 1) Healthcare Informatics, IBM Almaden Research Center, San Jose, CA; 2) Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD; 3) Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge MA.

Depending on the study design and statistical methodology, a genome-wide association study (GWAS) can sometimes require substantial computational resources. An example of a particularly demanding task is estimating empirical p-values via data permutation when a test statistic does not follow its parametric distribution or when a novel test statistic simply has no parametric distribution. Another example is the analysis of expression quantitative trait loci (eQTL) study designs with tens of thousands of gene expression phenotypes.

We introduce the BlueSNP R package, which distributes GWAS computations over a compute cluster configured with the MapReduce framework, making processor and memory intensive analyses feasible for large datasets. MapReduce is a parallel programming methodology for splitting a large problem into sub-parts (the map step), computing partial solutions on subparts independently, then assembling the partial solutions into the overall solution (the reduce step). BlueSNP removes the complexity of interacting with a compute cluster enabling the researcher to focus on advanced analytics using the R programming environment, primarily through the high-level gwas() function. Researchers can utilize the association tests provided with BlueSNP (e.g., case/control, regression, etc.) or supply a novel association test as a user-defined R function.

To assess performance, we simulated datasets with up to 4M SNPs, 100K individuals, and 100 traits using PLINK. We simulated 10 causal SNPs each contributing 1% of variance explained to a quantitative phenotype. We introduced 1% missing values to simulate genotyping failures. In most cases the time to complete an analysis was halved by doubling the number of compute nodes indicating near-optimal scaling. On our 40-node cluster configured with 200 map cores and 80 reduce cores it took 34 minutes to test 1M SNPs against 100 phenotypes in 100K individuals. Estimation of empirical p-values to a precision of $p \leq 10\text{-}8$ on the same dataset took 7.6 hours for one phenotype. The running-time of the adaptive permutation tests is largely determined by the number of highly significant SNPs (10 in this case).

3570W

Fast and Accurate Identification of Novel Sequences in *de novo* Human Genome Assemblies using NSIT. *B. Pupacdi[1], A. Javed[2], M.J. Zaki[3], M. Ruchirawat[1].* 1) Chulabhorn Research Institute, 54 Kamphaeng Phet 6, Laksi, Bangkok Thailand 10210; 2) Genome Institute of Singapore, A*STAR 60 Biopolis Street, Genome, #02-01 Singapore 138672; 3) Department of Computer Science Rensselaer Polytechnic Institute Troy, New York USA 12180-3590.

Novel sequences comprise a relatively new class of DNA sequence variations, and they are defined as non-repeating sequences present in a person's genome that are absent in the reference genome. A recent study showed that these sequences are consistent with the known migration paths and are potentially functionally important. Novel sequences can be detected both via sequence reads and via a *de novo* genome assembly. 2–5 Mbps of novel sequences have been reported per individual via both schools of methods. Our contribution is in the latter. The rapidly declining cost of sequencing contributes to the rising availability of *de novo* genome assemblies; hence there is a need for fast, accurate, and memory-efficient software to locate novel sequences. Using our software, NSIT (Novel Sequence Identification Tool), we identified 2.6Mbps, 2.5Mbps, and 2.3Mbps of novel sequences in the YH (Asian), NA18507 (African), and NA12878 (European) *de novo* genome sequences, respectively. NSIT was run on a 32-bit commodity desktop in 1.5 hours, on average, using less than 2GB of memory. Blastn and RepeatMasker were applied as post-processing steps, which took 30 minutes on average. The YH and NA18507 genomes were earlier reported to contain 5.1Mbps and 4.8Mbps of novel sequences, via aligning their *de novo* assemblies to the reference sequence using BLAT, LASTZ, Blastn, and RepeatMasker. The size discrepancies above are largely because the earlier work aligned to Build 36 only, while we also included Build 37. Moreover, the RepBase database in RepeatMasker used in our study was obviously more up-to-date. We adjusted their results accordingly and reduced the number of novel bases to 2.3Mbps and 2.2Mbps, respectively, and compared our results to them. Our identified novel sequences yielded 99% sensitivity and 85% precision. In addition, to the best of our knowledge, the novel sequences of NA12878 have never been reported elsewhere. Preliminary investigation revealed that about 8–15% of the novel sequences found by NSIT can be aligned to the haplotype chromosomes and unplaced contigs of the human genome. In addition, 70–75% of them can be aligned to the HuRef sequence. We also found that each pair of the three novel sequence sets overlap by 80–90%. Lastly, we compared NSIT with other aligners that have been used for this task, i.e. BLAT, LASTZ, and QueryLookupTable, and found that NSIT is several times faster.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3571F

NGS for the Masses: Empowering Biologists to Improve Bioinformatic Productivity. *K. Qaadri.* Biomatters, Inc., Newark, NJ.

Numerous enterprise-scale organizations spanning academic institutions, government facilities and commercial companies rely on a mixture of custom, open-source and proprietary commercial software for their NGS bioinformatics workflows. Dedicated bioinformaticians have been crucial to developing the tools, infrastructure and automated workflows available today, but are increasingly swamped by having to perform fairly routine bioinformatic tasks that could be done by biologists. To successfully engage biologists in mainstream NGS analysis requires a simplification of the software landscape with a data delivery platform and an interface that empowers biologists to access powerful hardware and software. This talk will focus on use cases that demonstrate how to build a flexible infrastructure for sequence-based research that combines commercial, open source and custom tools to improve ROI and accelerate discovery.

3572W

Evaluation and visualization of functional perturbations of pathways based on whole genome variations. *H. QIN, Y.Y. Shugart.* National Institute of Mental Health (NIMH)/NIH, Bethesda, MD.

While whole genome sequencing has opened a new era for personalized medicine for human complex diseases such as cancer, the complexity of genomic data presents challenge to the bioinformaticians. We have long anticipated that the involvement of multiple genes in pathways, and genetic heterogeneity of cancer cells determine drug response to treatment. Therefore, evaluating the contributions for the genetic variations to the functional change of pathways in cancer patients could be of high significance for personalized medicine. In this study, we aim to develop a new tool focusing on illustration the functional perturbations introduced by genetic variations in a whole genome scale. This tool is featured by: 1).Supporting multiple data sources for analysis, including data generated by many commercial services, such as Illumina and Complete Genomics. 2) Systematically integrating multiple pathway databases, including NCI-Nature Curated pathway, BioGrid, KEGG and Reactome; 3) Modeling functional changes of pathways by integrated functional evaluation algorithms and statistical methods. 4) Using new algorithms to compute the similarity of pathway models across samples and model-specific visualization. The features of the software include not only the robustness of integration of data by system biology methods, but also the ones used to evaluate the functional perturbations of pathways by comparing the similarity of matrixes. We applied this software to normal- cancer pair whole genome sequence data from the Complete Genomics in the initial stage and identified a tumor suppressor gene regulation pathway from 147 NCI-Nature curated pathways (including 29,626 nodes and 33,976 edges) as one of the most important pathways contributing to the cancer development. Currently, it takes 125.8 minites to finish all the processes. More features will be added gradually. We will also employ machine learning strategies for pathway ranking in simulation data and real data with larger sample size. This software will lead to the improved efficiency of the identification of the “shared” pathways and the pathways which may be “private” to the patients that are affected with complex diseases.

3573F

Use of an ad-hoc information system to support data search at a genomics center. *H. Qiu, F. Mentch, E. Frackelton, C. Kim, L. Hermannsson, H. Hakonarson.* Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, PA.

In the setting of a large genomics research center, data integration and presentation to end users greatly impact efficiency in utilization of available phenotype and genotype information residing in disparate data sources which includes both open source and commercial relational databases. We develop a two-component information system to support up-to-date integrated data view and complex analytical queries. First, we develop a distributed real-time query system that, through web interface, allows users to search all production databases at the center: the phenotype database which includes electronic medical records, survey conducted at enrollment of a study and extra information from collaborators, patient families and disease related registries; biological sample laboratory information management system (LIMS); and genotyping and sequencing production LIMS. We support real time query functionality in keyword search, patient/project status lookup and information validation, data query up to a combination of a limited number of modalities, search limited within an uploaded list of predefined patient IDs. The real-time query system does not scale well in case of really complex queries which require joining data from multiple modalities as they consume significant computing resources. Next, we deploy an i2b2 (Informatics for Integrating Biology and the Bedside) based framework for user friendly complex query composition and ontology-driven cohort selection. From the first step, based on user activity logs and interactive discussions with end users, we are able to continuously identify most frequently used data fields and types of popular queries. The subset of data fields are extracted and transformed into the i2b2 data model and loaded into a separate i2b2 data warehouse. The i2b2 based system is highly optimized for query speed and response time for complex queries, which makes it effective in support of research cohort selection, hypothesis generation and advanced analytics. In conclusion, the use of the ad-hoc information system with two complementary components (real time distributed query tool and data warehouse based i2b2 tool) provides an efficient and user friendly way to facilitate phenotype and genotype information search and knowledge discovery at our center.

3574W

Using biosignatures to stratify clinical response: Evaluation of classification methods using simulated gene expression signatures and real patient cohorts. *P. Ravindran, C. Ooi, L. Kai, H. Zhong, A. Belousov, H. Bitter.* Hoffmann-La Roche Inc, Nutley, NJ.

The need for personalized health care is evident in the heterogeneity of patient treatment outcomes. Identifying biomarkers predictive of patient responses will improve clinical efficacy and reduce the sample size and cost, and improve the safety of clinical trials. While single biomarkers have been used to stratify patient treatment responses, panels of biomarkers or biosignatures, e.g. gene expression and protein analytes, have not been used widely in the clinic for several reasons. For example, some progress has been made in identifying gene expression biosignatures predictive of treatment response, but no systematic study for translating the identified biosignature to clinical usage has been conducted.

In this study, we performed simulations to investigate the relationship between clinical response, gene expression biosignatures, and classification methods with the goal of understanding which methods are preferred in different contexts. One important characteristic of gene expression datasets is the correlation structure among transcripts. To preserve this correlation structure, we compared the methods using biosignatures derived from real gene expression datasets from oncology and immunology. We simulated different levels of correlation between clinical response and biosignature to mimic various scenarios ranging from weak to strong prediction power from a biosignature. We then performed a systematic and comprehensive evaluation of several major classification and regression methods to classify patients into their response subgroups. The methods studied include classical machine learning algorithms and gene expression specific algorithms like Gene Set Variation Analysis, BinReg, and Connectivity Score. We defined several metrics, including sensitivity, specificity, and association estimate to quantify the strength of a biosignature, and used them to evaluate the performance of the classification methods. We discovered that most algorithms performed similarly in the presence of strong correlation between biosignatures and clinical responses. However, the algorithms' performance varied substantially when correlation is weak to moderate, and the variation depends on the correlation structure of the genes in signature panel. This study provides guidelines on how these classification methods can be optimally used to translate a gene expression biosignature predictive of treatment response to use in clinical trials.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3575F

Mitre; A tool for bulk updates of SNP identifiers. *N.W. Rayner[1, 2], N. Robertson[1, 2], M.I. McCarthy[1, 2, 3].* 1) WTCHG, Univ Oxford, Oxford, United Kingdom; 2) OCDEM, University of Oxford, Oxford, United Kingdom; 3) Oxford NIHR Biomedical Research Centre, Churchill Hospital, Oxford, UK.

With the increasing use of sequencing technologies there is a need to ensure the novel SNPs discovered can easily be updated to rs numbers with each new release of dbSNP. This coupled with the large quantity of GWAS data generated on previous genome builds means there is a requirement to easily assign rs identifiers to various SNP listings which may comprise multiple identifier types. To solve this we have developed Mitre a simple to use web based Java application with an H2 embedded database deployed in a Jetty HTTP Server that allows users to upload a list of SNPs with their current chromosome and position from any genome build and receive a set of summary files based on any other genome build. The output comprises several files, including a list of SNPs with their newly assigned rs number, SNPs without a current rs number and those where the identifier is different from that uploaded. The system is designed to cope with multiple SNP identifiers, including Affymetrix and Illumina SNP chip ids. To test the system we considered three likely use cases. 1) Sequence data comprising SNPs with chromosome and position (chr:pos) identifiers, 2) conversion of existing GWAS SNP ids from previous genome builds and 3) assigning up to date rs identifiers to common SNP genotyping chips. To test the bulk SNP assignment for sequencing we took 17,571,460 chr:pos identifiers from a sequencing experiment and compared them to dbSNP builds 132, 134 and 135 which comprise 30.4, 41.3 and 52.2 million SNPs respectively, from this we obtained 53%, 63% and 70% of SNPs with rs identifiers respectively. For the GWAS update we selected a data set imputed on NCBI genome build 36 using HapMap2 comprising 2,614,446 SNPs with various identifiers, this we remapped to NCBI genome build 37 dbSNP 135. Using rs identifiers (2,610,944 SNPs) 2,606,702 (99.84%) do not change, 4212 (0.16 %) are merged with another rs identifier and 30 (0.001%) lose their rs identifier. Using the positions 2,613,635 (99.97%) can be assigned an rs number, 811 (0.03%) are unassigned. Many genotyping chips contain non rs identifiers, the Human Exome chip from Illumina comprises non rs identifiers for all 247,867 SNPs on the chip, using Mitre and dbSNP 135 we were able to assign rs numbers to 187,968 (75.8%) SNPs. Mitre is an easy and convenient way to update many varied SNP based data sets and with dbSNP gathering new SNPs quickly Mitre is capable of being run repeatedly with each new dbSNP release.

3576W

Genomic Analysis from Sequencer to Bedside: An Integrated Analysis Pipeline for Research and Clinical Sequencing. *J.G. Reid[1], M.N. Bainbridge[1], F. Yu[1], M. Dahdouli[1], D. Challis[1], P. Pham[1], D. Sexton[1], E. Boerwinkle[1,2], R.A. Gibbs[1].* 1) Human Genome Sequencing Ctr, Baylor College Med, Houston, TX; 2) University of Texas Health Science Center, School of Public Health, Houston, TX.

As sequencing technology has grown ubiquitous it has naturally matured to the point where many groups are now offering whole exomes and genomes to clinicians. Bringing sequencing to the clinic requires a significant effort to monitor processes, perform validations, provide security, and ensure repeatability, all on top of the analysis challenges already inherent in sequencing. As bioinformatic genomic analyses are integral to sequence production and the generation of clinically useful results, we have developed an integrated, CAP/CLIA-approved analysis pipeline to meet the needs of clinical sequencing at the BCM Whole Genome Laboratory. A flexible and clinically validated genomic analysis software pipeline is somewhat of a paradox, as the flexibility required for best-in-class genome sequencing can run counter to the strict regulatory requirements for change control. However, by maintaining a set of strict validations along with the code (workflows with test data), we are committed to maintaining a current, validated clinical analysis pipeline.

Our analysis pipeline (“Mercury”) consists primarily of a Master Control Program (MCP) for handling data flow and communication. The MCP starts by initiating vendor primary analysis software for generation of raw reads using CASAVA, and then moves the resulting reads through a BWA-based mapping pipeline. After mapping, the resulting BAM files go through a BAM finishing process including GATK recalibration and realignment, and into Atlas-SNP and Atlas-indel for variant calling, producing a VCF file after applying minimal filtering to called variants. At this point the SNPs and indels are annotated using a variety of data sources, and a final annotated report is delivered. Throughout the pipeline QC metrics are generated, and input/output hooks are available to bring the pipeline together with LIMS and other metadata tracking resources. Currently this analysis framework is in place in the Human Genome Sequencing Center's main production pipeline and handles roughly 20 terabases of sequence data each month, including a variety interesting Mendelian disease pedigrees. A clinically validated version is in place at the BCM Whole Genome Laboratory where it has been used to process over a hundred clinical cases, and 11 validation samples resequenced and reprocessed multiple times.

3577F

Human genome pattern mining framework for complex diseases. *M. Riemenschneider, M. Stoll.* Leibniz Institute for Arteriosclerosis Research, Muenster, Germany.

Recent advances in genomic research have significantly improved our understanding of the genetic architecture underlying common complex diseases. However, a large proportion of the genetic variance leading to common diseases remains elusive. Apart from genetic variants, epigenetic regulation is thought to play an important role in disease pathogenesis. Recently, we reported regions of conserved gene clusters in the human genome enriched for disease variants, which are subject to selection pressure. We demonstrated the evolutionary spread of disease variants and the conservation of disease associated genes in the vertebrate lineage. Interestingly, these regions also exhibited patterns of co-expression on the transcriptome level, leading to the assumption that epigenetic mechanisms may contribute another layer of genetic information to the complex genomic architecture underlying common diseases. However, due to the partial knowledge on the interaction of epigenetic factors and other genomic elements further investigations are required. We are developing a comprehensive bioinformatics approach, which combines data on genetic variants (SNPs, SNVs, structural variants), methylation patterns and transcriptome data involving chronic inflammation. As a backbone, publicly available data from HAPMAP, the 1000 Genomes Project, the UCSC Genome Project and the Encyclopedia of DNA elements (ENCODE) are used to establish the data framework for our analyses. Upon completion of the basic framework, data on two different complex diseases comprising whole genome SNP data, methylation patterns as well as the corresponding transcriptome data generated on our Illumina HiScanSQ platform will be used to annotate our previous in silico observation with actual data, targeting the genetic architecture underlying inflammatory disease phenotypes on multiple levels. We will present our computational approach as well as the current status of our work, which likely contributes to the understanding of disease pathogenesis.

3578W

A Tool for selecting endogenous control for qRT-PCR using high throughput expression data. *C.S. Rocha, C.V. Maurer-Morelli, I. Lopes-Cendes, F.M. Artiguenave.* Department of Medical Genetics, Unicamp - FCM, Campinas, São Paulo, Brazil.

Objective: Quantitative real-time PCR has become a standard for quantitative gene expression, and for validation of genes detected as differentially expressed in high throughput experiments, but the reliability of gene expression studies by mRNA quantification is highly dependent upon several experimental procedures, especially the choice of reference genes used for data normalization. qRT-PCR gene expression results are generally normalized using endogenous control genes, these reference genes should be expressed at a constant level across all sample groups in a study, and should not be influenced by study treatments or conditions. The normalization usually is made with one or more endogenous control genes, which must have its expression stable within the samples to be compared, regardless of experimental conditions, treatments or tissue differences. With that in mind we developed a tool to select genes that have the most stable behavior throughout the experiment. **Methods:** This tool was written using R software environment for the statistical calculation and data filtering, and receives as input a file with the expression intensity values. Soon a web interface will be deve loped and the tool will be freely available. **Discussion:** In statistics, the relative standard deviation (RSD or %RSD) is the absolute value of the coefficient of variation. It is useful for comparing the uncertainty between different measurements. So, using high throughput expression experiments data, once we have all intensities values, the tool calculates the RSD for each gene or probeset for the whole experiment, and selects the data with the lowers values (a default cut of 5% is set, but the user can change it), creating a list with the better candidates to endogenous control. This list is divided in three parts: Genes with high expression, medium expression and low expression.**Conclusion:** Genes used as endogenous controls in qRT-PCR experiments are often chosen arbitrarily from a pool of commonly used endogenous control genes such as GAPDH, and actinβ without verifying their variation over the experiment. Selecting an endogenous control gene to normalize gene expression data is one of the most important phases in the experimental design. That is why we believe that this tool will be very useful for selecting the best endogenous control for qRT-PCR.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3579F

Golden Helix GenomeBrowse: Cloud-enabled visual analytics of DNA and RNA-Seq NGS data. *G. Rudy, S. Gardner, M. Thiesen.* Golden Helix, Bozeman, MT.

High-throughput sequencing (HTS) has recently provided price competitive alternatives to microarrays for both RNA expression profiling with the RNA-seq protocol and DNA genotyping with whole genome and whole exome sequencing. Although the bioinformatics tools have matured for secondary analysis of sequence data, including alignment, variant calling and gene and transcript level quantification, the outputs of these tools often require inspecting the "raw read alignments" for putative variants and genes with interesting expression profiles. Because the BAM files containing these reads are both very large and computationally expensive to create, it is ideal to host them in a cloud environment and only access them "on demand".

We introduce Golden Helix GenomeBrowse, a free visualization tool for DNA and RNA sequence alignment along with annotations tracks from a rich catalog of pre-curated public data. GenomeBrowse is built from the ground up with the guiding principles of (1) working seamlessly with cloud-hosted data as fluidly and quickly as local files, (2) being intuitive to use for non-bioinformaticans to utilize in their research, and (3) having a multi-threaded architecture to make big-data visualization and analysis accessible to anybody capable of navigating Google Earth. By including integration with a rich repository of public data, users have no barriers to the process of interpreting their sequencing results. In particular, we demonstrate the ability of GenomeBrowse to stream RNA-seq samples sequenced to the depth of 25 million reads from the Amazon Cloud at the whole genome down to the gene level with annotation tracks ranging from 1000 genomes, dbSNP, genes and miRNAs. With a set of public breast cancer cell lines, we show how GenomeBrowse can highlight Single Nucleotide Variants and small Insertion/Deletions that may explain differentially expressed genes between cell types.

3580W

Informatics challenges on utilizing next-generation sequencing to advance disease understanding. *S. Saisanit, J. Hakenberg, Y. Li.* Pharma Research & Early Development Informatics, Roche Pharmaceuticals, Nutley, NJ.

Next generation sequencing provides new avenue to deepen our understanding of disease processes. The increasing availability of data from large international consortiums such as TCGA/ICGC, 1000 Genomes as well as from our own sequencing projects poses a full spectrum of challenges in informatics including data management, data retention, data analysis, and integration. This presentation describes our approach to manage these challenges. Together with many other informatics colleagues across our organization, we contribute to building a foundation platform named Understanding Disease Informatics System (UDIS). The aim is to integrate NGS data with other kinds of omics and next gen data with comprehensive sample annotations. The ultimate goal is to support better target and biomarker selection, validation and profiling for Roche Pharma R&D.

3581F

Integrated genome-phenome analysis. *M.M. Segal[1], M.S. Williams[2], J.G. Gleeson[3].* 1) SimulConsult, Chestnut Hill, MA; 2) Geisinger Health System, Danville, PA; 3) UCSD, La Jolla, CA.

The cost of sequencing the human genome has fallen low enough that the main concern about using whole genome information is now the cost of interpretation. We have addressed this concern by integrating genome sequencing with automated comparison to known phenotypes of diseases using a diagnostic decision support tool, SimulConsult, already in clinical use. The integrated genome-phenome analysis is done by importing the variant table consisting of HGNC gene names, variant severity scores and zygosity information into the diagnostic decision support tool to match it up with the patient's findings. Variant severity scores are derived from available protein structure and mutation data, with high severity scores deemed highly unlikely to occur incidentally, and thus have a powerful effect on diagnosis. Zygosity information is used to match variants with known data about whether a gene abnormality in each known disease is autosomal dominant, X-linked, etc. Compound heterozygotes are analyzed using a worst-case scenario in which the two most severe variants are considered to be on opposite chromosomes, and the abnormality is flagged to allow the clinician to evaluate the clinical plausibility of the disease and decide whether to do traditional sequencing to test for compound heterozygosity. The diagnostic tool uses its detailed database of clinical and laboratory abnormality in thousands of diseases to present the genome information in the context of the clinical presentation (pertinent positive and negative findings, including clinical signs and symptoms, laboratory test results and imaging information). The output includes a display of likely diagnoses ranked by probability, including not only genetic diseases detectable on next generation sequencing, but also genetic diseases not well detected (e.g., trinucleotide repeats and large copy number variants), and nongenetic diseases. For each disease, the tool takes into account the incidence, treatability, and family history, and for each finding in each disease it takes into account the frequency, age of onset and age of disappearance. The output also includes a display of findings, including genes, ranked by pertinence (measured by the ability of a finding to modify the differential diagnosis). The output is not just a one-time report. As the clinical query changes, laboratory studies become available, or the phenome database is updated, an updated analysis and report can be generated.

3582W

Variation data services at NCBI: archives, tools, and curation for research and medicine. *S. Sherry, K. Addess, V. Ananiev, C. Chen, D. Church, M. Feolo, J. Garner, T. Heffron, D. Hoffman, M. Kholodov, A. Kitts, J. Lee, J. Lopez, D. Maglott, R. Maiti, L. Phan, G. Riley, W. Rubinstein, D. Rudnev, Y. Shao, E. Shekhtman, K. Sirotkin, D. Slotta, R. Tully, R. Villamarin-Salomon, Q. Wang, M.H. Ward, H. Zhang.* National Center for Biotechnology Information, National Institutes of Health, Bethesda, MD.

NCBI operates several archives for sequence variation data that are increasingly relevant for medical research. These archives contain general variation data in the public domain and high density surveys of genetic diversity in medical study populations. Small scale variations are accessioned and distributed through dbSNP at http://www.ncbi.nlm.nih.gov/snp/ and larger structural variations (>50bp) are accessioned and distributed through dbVar at http://www.ncbi.nlm.nih.gov/dbvar/. Research participant genotypes, participant phenotypes and analysis results are distributed through the dbGaP controlled access system http://www.ncbi.nlm.nih.gov/dbgap/. Assertions of clinical significance for variants and alleles are accessioned and distributed through ClinVar at http://www.ncbi.nlm.nih.gov/clinvar/.

Tools for exploring and visualizing variation data include the 1000 genomes browser http://www.ncbi.nlm.nih.gov/variation/tools/1000genomes/, the 1000 genomes data slicer http://trace.ncbi.nlm.nih.gov/Traces/1kg_slicer/ and the phenotype-genotype data integrator http://www.ncbi.nlm.nih.gov/gap/PheGenI. The clinical remap tool at http://www.ncbi.nlm.nih.gov/genome/tools/remap#tab=rsg will provide sequence coordinates for variations on a clinical RefSeqGene record, and the variation reporter service at http://www.ncbi.nlm.nih.gov/variation/tools/reporter will provide a list of known variants and the functional consequences for a region of interest in BED format or set of variants in HGVS or GVS format.

The presentation will also report the results of NCBI's participation with the GET-RM consortium to establish standardized sequence data for next generation sequencing in clinical laboratories and to identify lists of variants with clinical phenotypes, and lists of common variants not known to be medically important. The latter are frequently used to filter normal variation from next generation sequencing results.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3583F

snpActs: A versatile webinterface for annotating and prioritizing SNV data sets. *B. Stade, D. Ellinghaus, B. Petersen, M. Forster, A. Franke.* Institute of Clinical Molecular Biology, University Hospital Schleswig-Holstein, Kiel, Germany.

The rapidly decreasing prices for sequencing entire human exomes and genomes resulted in large amounts of variation data. To cope with this variation data we developed snpActs, a database-driven toolset that allows scientists to annotate single nucleotide variants (SNVs) and to categorize them comprehensively. snpActs scans different gene annotations and identifies SNVs in functional elements. Additionally, snpActs utilizes the results of several established mutation effect prediction algorithms, such as SNAP, SIFT, and Polyphen2, to distinguish between deleterious and functionally neutral amino acid changes caused by SNVs. For this, it also checks the Human Gene Mutation Database (HGMD). In comparison to other annotation-programs snpActs can filter SNVs lists can be filtered using special rules (e.g. coding SNVs), special masks (e.g. cancer regions) or based upon other SNV lists (e.g. presence in relatives). snpActs further implements a classical and precise linkage analysis to examine regions that are identity-by-descent in data sets from complex pedigrees. Via these functionalities it is possible to identify potential disease causing genes in examined individuals.

3584W

Games for gene annotation and phenotype classification. *A.I. Su, S. Loguercio, C. Wu, B.M. Good.* Molecular and Experimental Medicine, The Scripps Research Institute, La Jolla, CA.

The Empire State Building was built with 7 million hours of human effort. The Panama Canal took 20 million hours to complete. By comparison, it is estimated that up to 150 billion hours are spent playing games every year (9 billion on Solitaire alone). Obviously people play games because they are enjoyable and fun. But aside from that enjoyment, games largely result in no tangible benefit, neither to the individual nor to society at large.

Recently, several groups have built “games with a purpose”, a class of games that focuses on collaboratively harnessing gamers for productive ends. In biology, games have been built to fold proteins and RNAs, and to perform multiple sequence alignment. Here, we present our efforts to apply games to two critical challenges in genetics.

First, we have built games focused on organizing and structuring gene annotations. With the increasing popularity of genome-scale science, many analysis strategies (including gene set enrichment, pathway analysis, and cross-species comparisons) depend on comprehensive and accurate gene annotations. These structured annotations are mostly the result of centralized manual curation efforts, but these initiatives do not scale well with the explosive growth of the biomedical literature. We describe several games that target working biologists to extract their expert domain knowledge in computable form. We present actual preliminary game play data, and demonstrate how mined annotations systematically improve gene set enrichment analysis.

Second, we describe a game for predicting human phenotypes from molecular descriptors. Researchers can now relatively easily characterize any biological sample according to a number of features, including genotype, gene expression, and epigenetics. A key challenge in the field is identifying exactly which of those molecular features can be used to predict a clinical phenotype like disease susceptibility or adverse drug events. While statistical classifiers have been applied to this challenge, they typically do not incorporate human expert knowledge, and they often fail to replicate in external test populations. Here, we describe a new classification strategy called “Human Guided Forests”. Based on the random forest algorithm, this hybrid classification approach incorporates crowdsourced human input through an online strategy game.

Prototypes of these and other games for genetics research are available at http://genegames.org.

3585F

HMM-Fisher: a hidden Markov Model based method for identifying differential methylation. *S. Sun[1, 2], X. Yu[2].* 1) Case Cancer Center, Case Western Reserve Univ, Cleveland, OH; 2) Epidemiology & Biostatistics, Case Western Reserve Univ, Cleveland, OH.

DNA methylation (the addition of a methyl group to a cytosine) is one of the most common molecular changes in cancer cells and plays a key regulatory role. Therefore, it is important to study cancer DNA methylation, especially differential methylation patterns between two groups of samples. With next generation sequencing (NGS) technologies, it is now possible to identify differential methylation patterns by considering methylation at all CG sites in an entire genome. However, it is challenging to analyze large and complex NGS data. In order to address this difficult question, we have developed a new statistical method using a hidden Markov model and the Fisher Exact test to identify differentially methylated regions. In particular, we first use a hidden Markov chain to model the methylation signals to infer the methylation state as Not methylated (N), Partly methylated (P), and Fully methylated (F) for each individual sample. We then use the Fisher Exact test to identify differentially methylated CG sites. The advantage of our method is that it can incorporate neighboring CG site methylation information and reduce the impact of sequencing errors. In this presentation, we show our HMM-Fisher method and compare it with a linear model based statistical method and a different HMM-based method (HMM-DM) that has been developed by our group using a publicly available data set.

3586W

Automated HGVS-recommended sequence variant description. *P. Taschner, J.F. Laros, M. Vermaat, J.T. den Dunnen.* Dept Human Genetics, Leiden Univ Medical Ctr, Leiden, Netherlands.

Errors in and consequent confusion from sequence variant reporting in relation to genetic disease is highly undesired. Although the recommendations for variant description of the Human Genome Variation Society (HGVS)[1] have been been largely accepted world-wide as the standard nomenclature, errors in reporting remain. To promote error-free reporting we have developed the Mutalyzer sequence variation nomenclature checker[2]. Construction of variant descriptions accepted by Mutalyzer requires basic knowledge of the standard nomenclature and comparison of the reference sequence and the variant sequence.

The chance of finding a complex variant increases with the advert of sophisticated variant callers (e.g., Pindel) and the rise of long read sequencers (e.g., PacBio). We have developed a new tool called Description Extractor[3] to help users describe complex variants. The algorithm used closely follows the human sequence comparison approach to describe a variant. It will first find the “area of change”. Then it finds the largest overlap between the original area and the area in the observed sequence. This process is repeated until the smallest description is found. The tool ensures that the same description will be generated every time researchers observe the variant. A significant advantage of the tool is that knowledge of the HGVS nomenclature is not required to generate correct descriptions. This not only helps clinicians to generate correct descriptions, but its implementation also allows automation of the description process.

Funded in part by the European Community's Seventh Framework Programme (FP7/2007–2013) under grant agreement No. 200754 - the GEN2PHEN project.

1) http://www.hgvs.org/mutnomen/
2) https://mutalyzer.nl/
3) https://mutalyzer.nl/descriptionExtract.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3587F

RNAseq Analysis using the Pipeline Graphical Workflow Environment in Neuropsychiatric disorders. *F. Torri[1,2], I.D. Dinov[2,3], A. Zamanyan[3], S. Hobel[3], P. Petrosyan[3], Z. Liu[3], P. Eggert[3,4], I. Guella[1], J. Pierce[3], A.P. Clark[5], J.A. Knowles[5], J. Ames[2], C. Kesselman[2], A.W. Toga[2,3], S. Potkin[1], M.P. Vawter[1], F. Macciardi[1,2].* 1) Department of Psychiatry and Human Behavior, University of California, Irvine, Irvine, CA 92617; 2) Biomedical Informatics Research Network (BIRN), Information Sciences Institute, University of Southern California, Los Angeles, CA 90292; 3) Laboratory of Neuro Imaging (LONI), University of California, Los Angeles, Los Angeles, CA 90095; 4) Department of Computer Science, University of California, Los Angeles, Los Angeles, CA 90095; 5) Zilkha Neurogenetic Institute, USC Keck School of Medicine, Los Angeles, CA 90033.

RNAseq is increasingly becoming the method of choice for investigating the transcriptional landscape of cells. This includes not only transcript abundance (i.e. differential expression), but also transcript diversity due to alternative splicing events, different promoter or transcription start sites or differential base pair editing of mRNA. An increasingly large number of workflows are available today to manage high-throughput genomics sequencing data, from basic data processing to high-quality visualization of results. Graphical workflow environments are emerging as useful tools for constructing, modifying, interconnecting and executing computational genomics protocols using data processing workflows, such as pipelines. We have developed the Graphical Pipeline for Computational Genomics and Visual Informatics for RNAseq (GPCGR) with a flexible graphical infrastructure for efficient RNA-seq analysis computing and distributed informatics research within the framework of a joint collaboration between LONI (Laboratory of Neuro Imaging) at UCLA and BIRN (Biomedical Informatics Research Network) at UCI, ISI and USC. We use this pipeline to analyze RNAseq data from 10 controls, 9 subjects with Bipolar Disorder (BD), and 8 Schizophrenics (SZ). Using our developed set of pipelines we performed, reference genome alignment and assembly of the reads into transcripts, estimation relative transcript abundance, identification of differentially expressed and spliced genes. We will report the possible diagnosis dependent differential expression and splicing of genes in pathways relevant to BD and SZ compared to controls. This study demonstrates the utility of collaboratively developed graphical pipelines to improve the efficiency of high-throughput RNAseq data analysis in investigations of the underlying pathophysiology of neuropsychiatric disorders.

3588W

ENViz: a Cytoscape plugin for integrative statistical analysis and visualization of multiple sample matched data sets. *A. Tsalenko[1], A. Kuchinsky[1], R. Navon[2], M.L. Creech[3], I. Steinfeld[4], Z. Yakhini[2].* 1) Agilent Laboratories, Santa Clara, CA; 2) Agilent Laboratories, Petah Tikva, Israel; 3) Blue Oak Software, Los Altos, CA; 4) Dept of Computer Science, Technion, Israel.

Modern genomic, metabolomics and proteomic assays produce high quality multiplexed measurements, often obtained from the same set of samples. These measurements characterize the molecular composition and activity of biological samples from complimentary points of view. Integrative analysis of such measurements still remains a challenge to researchers and practitioners in the field. Here we present ENViz - a plugin to the Cytoscape network biology software platform, that implements an enrichment networks approach to the joint analysis of two types of sample matched datasets. ENViz facilitates the extraction of biological insights from multiple types of highly multiplexed measurements via integrated statistical analysis of these measurements and by using available systematic annotations. ENViz analyses a primary data set (e.g. gene expression) with respect to a pivot data set (e.g. miRNA expression, metabolomics or proteomics measurements) and primary data annotation (e.g. pathway or gene ontology) in the following way. For each pivot entry, we rank elements of the primary data based on the correlation (or anti-correlation, or absolute value of correlation) to the pivot data across all samples, and compute the statistical enrichment of annotation elements (gene sets) in the top of this ranked list based on an mHG model (minimum hypergeometric statistics - Eden et al, PLoS CB 2007). Significant results are represented in Cytoscape as an enrichment network - a bipartite graph with nodes corresponding to pivot and annotation elements, and edges corresponding to pivot-annotation entry pairs with enrichments scores better than the user defined threshold. In addition, for analysis of enrichment in biological pathways, correlations of primary data and corresponding pivot data are visually overlaid on biological pathways for each significant pivot-annotation pair, using the WikiPathways (Kelder et al, NAR 2012) resource. Enrichments of GO categories are overlaid on top of the Gene Ontology DAG. Edges of the enrichment network, representing significant associations, may point to functionally relevant mechanisms. In Enerly et al. 2010 an association between miR-19a and the cell-cycle module was substantiated as an association to proliferation, further validated using high-throughput transfection assays where transfection of miR-19a to MCF7 cell lines resulted in increased proliferation.

3589F

Determining probability of rare variants: Design implications for family-based sequencing studies. *W. Wang, G. Peng.* Bioinformatics & Comp Biology, Univ Texas MD Anderson Cancer Ctr, Houston, TX.

It is still challenging to find rare variants in order to better understand the genetic basis of human disease. Family-based sequencing studies have been performed, in search for functionally important genes. A high false discovery rate in calling rare variants precludes comprehensive downstream functional analyses. To improve accuracy, recent studies have used linkage information among relatives. Ongoing sequencing studies include data on extended family members. We developed FamSeqPro, which evaluates the probability of variants in family-based sequencing data, given raw measurements from all family members, at a single base level. Using FamSeqPro, we performed simulation studies to identify cost-effective designs for finding inherited genetic mutations in families with diseases of interest. We illustrate the utility of FamSeqPro in real data: whole-genome sequencing of extended family pedigrees.

3590W

Harvest: A Web-Based Biomedical Data Discovery and Reporting Application Development Platform. *P.S. White[1,2,3], B. Ruth[1], M.J. Italia[1], J. Miller[1], J.W. Pennington[1].* 1) The Children's Hospital of Philadelphia, Philadelphia, PA; 2) Division of Oncology, The Children's Hospital of Philadelphia, Philadelphia, PA, 19104, USA; 3) Department of Pediatrics, University of Pennsylvania School of Medicine, Philadelphia, PA 19104, USA.

Biomedical researchers share a common challenge of making complex data understandable and accessible. This need is increasingly acute as investigators seek opportunities for discovery amidst an exponential growth in the volume and complexity of sequence data derived both from advances in experimental technology and the availability of large patient-oriented data stores relevant to genomic research. We sought to address this need by constructing novel tools to more effectively enable data discovery workflows that effectively integrate subject observational data with molecular data derived from patient samples. We developed Harvest, an open-source framework that provides a set of modular components to aid the rapid development and deployment of custom data discovery software applications. Harvest incorporates visual representations of multidimensional data types in an intuitive, web-based interface that promotes a rapid, iterative approach to exploring complex clinical and experimental data. These visualizations present data in contexts and formats immediately familiar to researchers without exposing them to underlying physical data models, which facilitates data exploration for hypothesis generation and pre-defined queries. The Harvest architecture capitalizes on standards-based, open-source technologies (e.g., HTML5, django, python, jQuery) to address multiple functional needs critical to a research and development environment, including domain-specific data modeling, abstraction of complex data models, and a modular, customizable web client. Harvest provides a large library of data visualization modalities, the ability to instantly identify subject cohorts based on user selection of both phenotypic and molecular attributes, an intuitive ETL tool, REST APIs to facilitate data exchange, and the ability to save, recall, and export user-defined data sets. Interfaces for novel disease investigations with well-structured data can be generated in days or weeks. We have demonstrated the utility of Harvest by developing and deploying multiple applications across several research domains, including hearing impairment, cardiology, and mitochondrial disorders. Examples include the NIDCD Audiological and Genetic Database (audgendb.chop.edu), the NHLBI Pediatric Cardiac Genetics Consortium (pcgc.research.chop.edu), and the NHGRI CSER Program, each of which support national-scale collaborative research. Harvest is available at http://harvest.research.chop.edu.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3591F

Pinpoint: A new algorithm to detect gene-gene interactions in large-scale genome-wide association studies. *D. Wong[1], S. Szymczak[2], J.E. Bailey-Wilson[2].* 1) Pinpoint Analytics, San Francisco, CA; 2) Statistical Genetics Section, Inherited Disease Research Branch, National Human Genome Research Institute, National Institutes of Health, Baltimore, MD.

Genome-wide association studies (GWAS) have identified more than one thousand single-nucleotide polymorphisms (SNPs) that are reproducibly associated with hundreds of phenotypes. However, these variants explain only a small proportion of the overall heritability. The standard approach in GWAS is to analyze each SNP separately and is therefore not designed to identify genetic variants having a joint effect on the phenotype. Different statistical methods exist to identify gene-gene interactions, e.g. multifactor dimensionality reduction (MDR) or logic regression. However, most of these approaches do not scale well for testing all interactions in GWAS data, making these analyses computationally challenging. We evaluated a new computationally efficient algorithm, Pinpoint, that extracts multiple combinations of SNPs out of the data without restricting the search to a specific degree of interactions. We compared Pinpoint with logistic regression and random forests, a machine learning approach that jointly analyzes all SNPs. Genotypes for 100K independent SNPs in 10,000 samples were simulated with twenty-one SNPs having an effect on disease status by two-, three-, four-, five- and six-way interactions. In this model, the genetic main effects and interactions completely accounted for the risk of disease in the population. In this simple simulated model, all methods identified only causal SNPs, indicating good control of false positive findings. However differences were found in the ability to extract causal SNPs, the ability to interpret the extracted model, and the amount of computing resources required. In terms of the faintest signal successfully extracted by the algorithm, random forest was about two times more powerful than logistic regression, and Pinpoint was about five times more powerful. In terms of interpretability of the extracted model, Pinpoint provides explicit modeling of gene-gene interactions, whereas logistic regression and random forest only identify individual SNPs. In terms of required computing resources, Pinpoint requires several orders of magnitude less runtime than both logistic regression and random forest, and can be run on a PC. We plan to extend our simulation by simulating SNPs with realistic local linkage disequilibrium patterns and by simulating more realistic disease models with smaller effect sizes on risk of the disease. We will also present comparisons with other algorithms for analyzing large-scale datasets such as MDR.

3592W

A Unified Analysis Framework for Detecting Genetic Variations from Next-Generation Sequencing Data. *C. Xiao, S. Sherry.* NIH/NLM/NCBI, bethesda, md.

The per-base cost of whole genome sequencing has dropped significantly due to recent advancement of next generation sequencing (NGS) technologies, and a number of large-scale re-sequencing projects, e.g. 1000 Genomes, ICGC, TCGA, GO-ESP, and CGCI etc., have been initiated to extend our knowledge of single nucleotide polymorphisms (SNPs), short insertions/deletions (INDELs) and structural variations (SVs), and relate these variants to human diseases. However, data generation and analysis in a timely fashion present numerous challenges to researchers. In order to facilitate NGS data analysis in biomedical research, NCBI develops an integrated analysis framework (VPIPE) to profile genetic mutations from next-generation sequencing data in a uniform manner. This pipeline manages parallel-computing resources, aligns short reads to the reference genome sequences, refines the mapping of placed reads, calls SNPs, INDELs, and SVs, and performs de novo assembly and functional annotation according to data availability and project-specific policies. A centrally implemented pipeline streamlines the data processing workflow for the data generated by next-generation sequencing technologies.

3593F

HMM-DM: identifying differential methylation patterns using a Hidden Markov Model. *X. Yu[1], S. Sun[1,2].* 1) Epidemiology & Biostatistics, Case Western Reserve University, Cleveland, OH; 2) Case Cancer Center, Case Western Reserve University, Cleveland, OH.

DNA methylation is essential in both normal organism development and cellular differentiation. It is important to study the differential methylation patterns between two groups of biological conditions, especially for cancer diagnosis, prognosis, and therapeutics. With the rapid development of next-generation sequencing (NGS) technologies, it becomes possible to detect and analyze methylation patterns in single-base resolution for a large genome. However, it is challenging to process the enormous amount of complex methylation data generated by NGS technologies. Therefore, an accurate and effective NGS-based statistical method is in great need. To fulfill this demand, we have developed a Hidden Markov Model based method (called HMM-DM) to identify differential methylation between two groups. In particular, a Hidden Markov chain is used to infer the methylation of each CG site as Hypermethylated, Equally methylated, or Hypomethylated between two groups. Differentially methylated regions are then summarized based on the inferred hidden states, number of sites, and the distance between neighboring CG sites. We've compared our HMM-DM method to a CpG island-based linear model using a public available data set and shown that our method has the advantage of identifying local methylation patterns, and it can easily identify any differentially methylated regions in a genome. In addition, we have also compared this HMM-DM method with another HMM-based method (HMM-Fisher) that has been developed by our group.

3594W

Genotype calling for next-generation sequencing data from multiple populations. *K. Zhang, D. Zhi.* Dept Biostatistics, Univ Alabama at Birmingham, Birmingham, AL.

Genotype calling for next-generation sequencing (NGS) data from larger samples typically achieves higher accuracy than that from smaller samples. This is because current LD-based genotype calling methods, the best performing ones, leverage sharing of chromosomal segments among individuals, and there is a higher chance of chromosomal sharing in larger samples. As NGS data for multiple populations are becoming available, it is desired to combine the samples from multiple populations to achieve a large sample size. However, current LD-based methods typically assume a flat population structure, and thus may not work well with non-homogeneous samples. We developed a LD-based method that incorporates population structure into the commonly used chromosome segment sharing Hidden Markov models. Using simulated and real data, we show that our new models improve the accuracy of genotype calling and haplotype inference for samples from multiple related populations.

3595F

Detecting structural variants in cancer from whole genome and exome sequencing. *J. Zhang[1], Y. Shi[1], W. Foulkes[1,2], J. Majewski[1].* 1) Department of Human Genetics, McGill University and Genome Quebec Innovation Centre, Montreal, Quebec, Canada; 2) Program in Cancer Genetics, Departments of Oncology and Human Genetics, McGill University, Montreal, Quebec, Canada.

Next generation high-throughput sequencing technologies have been applied to investigate genetic disorders and gene expression changes on a large scale. Structural variants represent important types of genetic alternations that can lead to abnormal cell growth and proliferation, and therefore contribute to the development and progression of cancer. Various computational algorithms have been developed to identify and characterize structural variants from genome sequencing data. These approaches can be divided into four categories: read-pair, read-depth, split-read and assembly. However, most of these computational methods are not integrated, and their abilities to detect structural variants from cancer samples are not clear. Here we describe and evaluate three software: BreakDancer, CNAseg and Pindel, which are based on read-pair, read-depth and split-read algorithms, respectively. We performed the analysis using two cancer genomes from familial Wilms' tumor patients, and compared the three software based on their sensitivity and accuracy. Approaches for structural variant detection in exome sequencing data are more limited because of the discrete nature of the data as well as the presence of systematic biases that result in non-uniform distribution of read-depth. As a consequence, most current methods for copy number variation detection rely on pair-wise comparison of normalized read-depth; split-read and paired-end methods are inapplicable. Presently, we evaluate two software (Seqgene and CONTRA) and introduce our own modified implementation (fishing CNV), all of which are based on read-depth. Preliminary analysis of exomes from Neonatal opsoclonus and ovarian cancer patients using these software show highly comparable results, and suggest that read-depth approaches are capable of detecting copy number alterations at the resolution of the exon, as well as identifying large scale genomic alterations in cancer.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3596W

Causal inference of gene regulation based on sub-network assembly. *W. Hsieh[1], C. Peng[2], A. Dai[1], S. Peng[4], T. Yen[3,4].* 1) Inst Statistics, National Tsing Hua Univ, Hsin-Chu, Taiwan; 2) Resource Center for Clinical Research, Chang Gung Memorial Hospital, Taoyuan, Taiwan; 3) Head and Neck Oncology Group, Chang Gung Memorial Hospital and Chang Gung University, Taoyuan, Taiwan; 4) Department of Nuclear Medicine and Molecular Imaging Center, Chang Gung Memorial Hospital and Chang Gung University, Taoyuan, Taiwan.

Identifying disease-causing genes requires studying the nature of genes in the context of regulatory systems. We develop a four-step strategy to integrate both gene expression and SNP variation to make causal inference on a gene regulatory network. The first step towards our goal is to select gene modules on an interaction map that includes the protein-protein interaction, gene-gene interaction and well recognized pathways. Those modules are overlapping to each other and they individually are highly connected according to the gene expression profiles. The modules are also selected according to their discriminative power of distinguishing patients' disease status. The second step is to select the sequence variation underlying each of the modules. A multivariate modeling strategy is adopted to select SNPs that are strongly associated with the expression. It is close to the expression QTLs while the response we concern now is the joint variation of the whole module. The third step is to build directed sub-networks with each module and its associated SNPs. The sub-networks are then assembled with a ranking strategy that transfers the local information into global information. The proposed procedure was applied on a cohort of oral squamous cell carcinoma patients that were profiled with both exon expression and SNP variation. The modules were selected according to their association with lymph node metastasis. The causal network of our inference shows a change of regulation direction among a few key genes in the PPAR signaling pathway.

3597F

Identification of candidate variants in population based whole exome sequencing data through integration of functional prediction with bayesian penalized regression. *J. Lu[1,2], A. Sabo[1], J. Reid[1], D. Muzny[1], E. Boerwinkle[1,4], R. Gibbs[1,3], ARRA Autism Sequencing Consortium.* 1) Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX; 2) Department of Structural and Computational Biology and Molecular Biophysics, Baylor College of Medicine, Houston, TX; 3) Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas, USA; 4) Institute of Molecular Medicine, University of Texas Health Science Center, Houston, Texas, USA.

Despite their success in identifying variants associated with complex disease, current SNP by SNP methods used in genome-wide association studies (GWAS) are inadequate for use with population case -control with whole exome sequencing data. Rare variants inherent in exome data (minor allele frequency < 5%) challenge traditional frequentist approaches. Further, joint modeling of rare and common SNPs and INDELs while computationally challenging, provides more accurate analysis of genetic structure. For example, any statistical model for exome data should accommodate models where multiple rare variants throughout the population may affect genes and pathways at different nucleotide positions but still produce the same disease phenotypes.

We are implementing two concurrent solutions to identify the significant variants in case-control whole exome sequencing data. We employ both a Bayesian Probit and Logistic penalized (LASSO) regression. Through the appropriate use of priors, these models are able to engage in variable selection. Parameters in these models are estimated using Markov Chain Monte Carlo or with variational methods, respectively.

We model our data in several ways: we first employ both additive and dominant genetic models. We then include in step wise fashion functional, conservation, and gene based annotation data from ANNOVAR and dbNSFP. Lastly we aggregate the variant level data into our hierarchical regression as gene and pathway level effects. To test our methods we apply them to the ARRA Autism Sequencing project, a multi-institutional collaboration focused on identifying heritable variation in Autism Spectrum Disorders via whole exome sequencing of 1000 cases and 1000 matched controls. These methods are generating candidate variants and genes for further functional testing.

3598W

Joint association of multiple correlated phenotypes using matrix-variate linear mixed-models. *N. Furlotte, E. Eskin.* University of California, Los Angeles 3532-J Boelter Hall Los Angeles, CA 90095-1596.

Genome-wide association study (GWAS) methods have become a favorite tool among those available to researchers attempting to dissect complex traits. However, most GWAS methods are designed to interrogate individual phenotypes, while ignoring the availability of multiple correlated phenotypes. Multi-trait mapping approaches, exploiting the utility of correlated phenotypes in the mapping procedure, have been proposed in the context of linkage analysis and a small number of works have proposed similarly motivated methods for GWAS. Unfortunately, the complex dependency structures present in many current GWAS datasets, often caused by population structure or cryptic relatedness, make it difficult to translate linkage-based approaches to the GWAS context and limit the applicability of the proposed multi-trait GWAS approaches. Linear mixed-models (LLMs) are a widely accepted approach to avoid the pitfalls of population structure in single trait association analysis and LMM-based methods for performing artificial selection in agricultural and livestock communities using multiple correlated traits have a rich history in the literature. Such LMM-based multi-trait models may be adapted for use to map multiple correlated traits in GWAS. However, such approaches become computationally intractable, when considering multiple correlated phenotypes across large GWAS cohorts. In this paper, we introduce a genome-wide association study methodology for use with multiple correlated phenotypes, which we call the matrix-variate linear mixed-model (mvLMM). Our mvLMM method has a significantly lower computational cost when compared with a standard LMM-based formulation. We first show that this method results in increased power over single trait mapping. Next, we show how when using the mvLMM framework, it is possible to isolate the fraction of the total phenotypic correlation that is attributable to genetics from that which is attributable to environment. Finally, we apply our method to map traits in the 1966 Northern Finnish Birth Cohort and evaluate the ability of the proposed method to identify known associations with increased significance and to identify novel associations that were not uncovered in single trait analysis.

3599F

The Human Gene Connectome: A Map of Short Cuts for Morbid Allele Discovery. *Y. Itan[1], S.Y. Zhang[1], G. Vogt[1], A. Abhyankar[1], L. Quintana-Murci[3], L. Abel[2], J.L. Casanova[1,2].* 1) The Rockefeller University, New York, NY; 2) INSERM, Paris, France; 3) Pasteur Institute, Paris, France.

High-throughput genomic data reveal thousands of gene variants per patient and it is often difficult to decipher which of these variants underlie disease in individuals. At the population level however, there can be some level of phenotypic homogeneity, with alterations of specific physiological pathway underlying the pathogenesis of a particular disease. Here we describe the human gene connectome (HGC) as a new approach facilitating the interpretation of abundant genetic data, guiding subsequent experimental investigations. We identify the set of degrees of separation, shortest plausible biological distances and routes between all pairs of human genes, by applying a shortest distance algorithm to the full human gene network. We demonstrate a hypothesis-driven application of the HGC in which we generate a TLR3-specific connectome, which may be useful for the genetic dissection of herpes simplex virus encephalitis of childhood. We also developed the functional genomic alignment (FGA) approach from the HGC. In FGA, the genes are clustered according to their biological proximity (rather than the traditional evolutionary genetic distance), as estimated from the HGC. We discuss HGC application for discovering the comprehensive set of disease causing alleles in full cohorts of patients and for complex diseases and traits. The HGC and FGA should facilitate the genome-wide discovery and experimental validation of disease-causing alleles.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3600W

Transcriptome sequencing of nPOD type 1 diabetes pancreatic samples for viral sequence identification. *S. Morfopoulou[1], G. Zhao[2], R. Ferreira[3], A. Pugliese[4], J. Petrosino[5], L. Thackray[2], V. Plagnol[1], nPOD-Virus group.* 1) University College London (UCL) Genetics Institute, United Kingdom; 2) Division of Immunobiology, Washington University, St Louis; 3) WT/JDRF Diabetes and inflammation laboratory, University of Cambridge, Cambridge, UK; 4) Diabetes Research Institute, Miller School of Medicine, University of Miami; 5) Department of Molecular Virology and Microbiology, Baylor College of Medicine, Houston.

A growing body of genetic and epidemiological evidence suggests a link between viral infections and type 1 diabetes (T1D). Viruses may act as a trigger for T1D onset and the identification of viral agents that are present in young onset cases may provide new clues into T1D pathogenesis. Transcriptome sequencing, or RNA-Seq, is a powerful technology to detect active viral infections in human tissue. To identify viruses relevant to T1D etiology, we undertook as part of the network of pancreatic donors virus group (nPOD-V) initiative a RNA-Seq experiment for a set of post-mortem pancreatic samples from young onset T1D cases and controls. We sequenced total RNA from n = 7 pancreatic samples using a combination of Illumina short sequencing reads and longer 454 reads. We used a metagenomic analysis to detect pathogens present in these samples. To maximize the power of this study, we developed a novel Bayesian statistical approach to detect organisms expressed at very low levels in metagenomic mixtures. Our results highlight the difficulties and opportunities provided by deep short read sequencing to generate insights into the contribution of viral infections to T1D.

3601F

VARITAS: Variant Analysis with Rapid Incorporation of Annotation Sources. *B. Powell.* Medicial Genetics, Baylor College Med, Houston, TX.

The analysis of genome-scale data for purposes of diagnosis or gene discovery begins with automated analysis but typically involves manual evaluation of variants of possible significance by comparison with a number of databases that are undergoing frequent upgrades. This manual step can be facilitated by presentation of a broad class of existing annotations to the expert annotator. VARITAS is a software package that allows for incorporation of information from internally-developed or external annotation sources such as the 1000 Genomes Project, NIEHS Environmental Genome Project, NHLBI Exome Sequencing Project, dbSNP, etc. This software was designed for a particular balance of performance with ease of installation and use. The tabix indexing scheme is used, obviating the need for maintenance of a generalized database program. Reference datasets are stored in a compressed format, minimizing disk utilization and read-time when used in network file storage. The standard VCF (variant call format) is used for both input and output, allowing VARITAS to be easily included in other data analysis pipelines. Configuration may be performed at the command-line or with a script-like configuration file; addition of new data sources does not require recompilation. Variants are processed in a streaming fashion, so runtime memory usage is less than 1GB. Processing time varies with the number of data sources incorporated, but annotation of variants from a typical exome can be performed in less than 20 minutes on a personal computer. VARITAS runs using the Java virtual machine, so it can be used on all major computing platforms. This work is supported through 5R01CA138836 and 5T32GM007526.

3602W

Clinical Diagnostic Tools for Family Genome and Exome Data. *A. Russell, F. De La Vega, J. Rule, M. Reese.* Omicia Inc, Emeryville, CA.

Continuing declines sequencing costs combined with successes in the elucidation and diagnosis of Mendelian diseases through NGS whole-genome and exome sequencing is creating enthusiasm for using these technologies in the clinic. These analyses are much more powerful when unaffected or affected family members are analyzed together with the patient, resulting in higher diagnosis rates. However, the analysis pipelines described to date are either custom workflows or tailored for expert researchers and are complex to use or install, lack secure access and project management capabilities, or use terminology unfamiliar in a clinical setting. Here we present the family analysis capabilities of the Omicia Opal system, a secure online service which enables analyses of whole-genome, exome or gene panel data provided in commonly used formats such as VCF and Complete Genomics (CGI). When supplied family pedigrees, it verifies that family relationships are correctly labeled to rule out relationship misspecification, and employs the VAAST algorithm to rank genes with regard to the likelihood that they are disease-causing. The user is presented a filterable table of candidate disease-genes and their variants, along with their genotypes in each family member,,classified following the guidelines from the American College of Medical Geneticists (ACMG). Information is presented on any protein impact the variant might cause, allele frequency information, and disease information associated with the either the genes or variants from publically-available and licensed databases, or from the user's own gene and variant databases, as well as sequencing metrics such as quality values and read depth. Users can annotate relevant variants, flag non-disease causing variants or variants of unknown significance (VUS) where appropriate, share their findings with colleagues and generate reports in PDF format suitable for consumption by clinicians. We present results for several simulated cases, constructed by spiking two trios of Puerto Rican and Yoruban origin sequenced by CGI, with disease variants for several recently elucidated Mendelian diseases. These cases are analyzed with Opal's variant mining interface and VAAST report showcasing how the tool could be used in clinical research and eventually diagnostics. The Opal system addresses deficiencies in current analysis software and provides capabilities that enable interpretation of NGS genome and exome data in a clinical setting.

3603F

ASAP: an easy to use pipeline for sequencing data processing. *E. Torstenson[1], C. Li[1,2].* 1) Center for Human Genetics Research, Vanderbilt University, Nashville, TN; 2) Department of Biostatistics, Vanderbilt University, Nashville, TN.

Modern sequencing technologies have given researchers access to deep sequencing data on a large scale and in a timely fashion. However, data of this size presents many new challenges, including lengthy processing times and increased probability of human error when processed manually. We have developed a new program, Advanced Sequence Automated Pipeline (ASAP), as a way to simplify the process of translating fastq sequence data into annotated variant calls by typing as few as three commands; all without the need for administrative rights or dedicated hardware. ASAP works both on computer clusters and on standalone machines with minimal human involvement and maintains high data integrity, while allowing complete control over the configuration of its component applications. It is highly configurable and allows users to align fastq data, realign reads around insertions and deletions, recalibrate quality scores, perform variant calls and annotation in any combination while offering flexible sample grouping schemes to suit various needs and resources. It also provides tools for dividing jobs into small pieces for maximum throughput on the local cluster and offers an easy to use interface for submitting and tracking jobs and identifying job failures. ASAP also allows basic quality checking such as transition-transversion and het:hom ratios so that researchers can immediately spot problems before committing resources toward analysis. The program allows users to incorporate configuration settings from other researchers to allow shared, “Best Practice” methodologies, assist groups in using common resource paths as well as providing a simple mechanism for reproducibility. ASAP is designed to minimize user involvement and maximize flexibility for use and future development. We used ASAP to process 5 exome and 2 whole genome paired end sequencing data with >30x coverage which completed without any further user intervention. Steps performed include alignment, local realignment, recalibration and SNP calling. Each step was broken down into a number of jobs which were run in parallel. For the 5 exomes, a total of 287 jobs were executed with average runtimes of 43.6, 29, 7 and 1.6 minutes respectively at each step. For the 2 whole genome sequence data, a total of 192 jobs were created with average runtimes of 277.5, 181.7, 109.3 and 9.8 minutes respectively at each step.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3604W

Identifying biological pathways and human diseases in high-throughput human genetic studies. *C. Xie[1], X. Mao[2], J. Huang[1], Y. Ding[1], J. Wu[3], S. Dong[1], L. Kong[1], G. Gao[1], C. Li[4], L. Wei[1].* 1) Center for Bioinformatics, State Key Laboratory of Protein and Plant Gene Research, College of Life Sciences, Peking University, Beijing, China; 2) Computational Systems Biology Lab, Department of Biochemistry and Molecular Biology, Institute of Bioinformatics, University of Georgia, USA; 3) Cancer Research Program, Garvan Institute of Medical Research, 384 Victoria Street, Darlinghurst, NSW, Australia; 4) Laboratory of Bioinformatics and Genomic Medicine, Institute of Molecular Medicine, Peking University, Beijing, China.

Genome-wide high-throughput technologies give the opportunity to study polygenic effects existed in complex diseases. These high-throughput human genetic studies often identify dozens to hundreds of relevant genes. From these genes, geneticists want to understand which biological pathways may be involved in and which other human diseases may be related with. Here, we report a method to deal with this problem. We retrieved and integrated data from five pathway databases (KEGG PATHWAY, PID, BioCyc, Reactome and Panther) and five human disease databases (OMIM, KEGG DISEASE, FunDO, GAD and NHGRI GWAS Catalog) as our background knowledge base. For an input set of genes, our method annotates them with putative pathways and disease relationships based on mapping to genes with known annotations. Then it performs hypothesis tests and FDR corrections to identify statistically significantly enriched pathways and diseases. Our method can be accessed at http://kobas.cbi.pku.edu.cn. We further optimized our statistical method, and incorporated knowledge from Gene Ontology, which improved the results in our researches of complex diseases.

3605F

GenAMap: An Visual Analytics Software Platform for eQTL and GWAS Analysis. *E. Xing[1,3], R. Curtis[2,3], S. Lee[1], S. Shringarpure[1], J. Yin[3].* 1) School of Computer Science, Carnegie Mellon University, Pittsburgh, PA; 2) Joint Carnegie Mellon-University of Pittsburgh PhD Program in Computational Biology, Carnegie Mellon University, Pittsburgh, PA; 3) Lane Center for Computational Biology, Carnegie Mellon University, Pittsburgh, PA.

Structured association mapping has been shown to be a powerful approach to find disease associated genetic polymorphisms. However, these algorithms are often difficult to use as it requires expertise and effort to customize. Because of the difficulty required to use these state-of-the-art techniques, simpler and less powerful methods are more popular. To make algorithms easily accessible to users, we developed a software platform called GenAMap. It integrates computational methods for association mapping and population genetics with novel visualizations and graphical user interface. GenAMap uses an automatic processing system called Auto-SAM, which enables biologists to automatically perform structured association mapping using various algorithms. Auto-SAM includes algorithms to discover gene-networks and find population structure. It can also run popular association mapping algorithms. For efficient execution of these algorithms, Auto-SAM makes use of parallelization. Auto-SAM is available through GenAMap. GenAMap and AutoSAM are implemented in JAVA; GenAMap is available through http://sailing.cs.cmu.edu/genamap.

3606W

PHV: A high accuracy SNP and INDEL variant caller based on the Profile Hidden Markov Model. *M. Zhao, W. Lee, G. Marth.* Biology Dept., Boston College, Chestnut Hill, MA., USA.

Accurately calling single nucleotide polymorphisms (SNPs), and insertion and deletion (INDELs) polymorphisms is crucial to medical and population genetics. A variety of algorithms have been developed to detect such variations using next-generation sequencing (NGS) data and have been successfully applied to large-scale re-sequencing projects (for example the 1000 genomes project), but INDEL detection still presents significant problems. Moreover, most of the existing algorithms are designed for high-quality short-read data with a low INDEL sequencing error rate whereas upcoming technologies such as Oxford Nanopore single molecule sequencing technique produce long reads at the cost of an increased sequencing error rate. This leads to more challenging variation detection, and is beyond the scope of present variant detection algorithms. We have developed a tool based on Profile Hidden Markov Model (PHMM) for accurately calling SNPs and INDELs from data with varying quality. In particular, it focuses on jointly modeling sequencing and alignment uncertainties, which are overlooked or not properly modeled by most existing variant calling algorithms. While this tool is developed for use with long read technologies, it also demonstrates an improvement in variation detection when deployed on currently available short read technology data. A single sample (NA12878 from the 1000 genomes project) with deep sequenced whole genome data (Illumina and LS454) has been used to test the algorithm. Comparisons with variants called using the popular Samtools and Freebayes software, PHV (the PHMM supported Variation detector) showed higher sensitivity and specificity for INDEL detection and higher specificity for SNP detection, especially in the genomic regions with poorly aligned reads. We plan to use banded PHMM in a sliding window along the whole genome to achieve comparable processing speeds to other variation callers and test the performance of this tool on PacBio and Ion Torrent sequenced E.coli data at the next step.

3607F

Developing copy number variation based case-control association analysis tool and its application for disease analysis. *Y. Chung[1], J.H. Kim[1], H.J. Hu[1], S.H. Yim[1], J.S. Bae[2], S.Y. Kim[3].* 1) Integrated Res Ctr Genome Polymorphism, Catholic Univ Korea,Sch Medicine, Seoul, South Korea; 2) Laboratory of Genomic Diversity, Sogang University, Seoul, Korea; 3) Medical Genomics Research Center, KRIBB, Daejeon, Korea.

Copy number variation (CNV) is one of the major components of human genetic variations and it is thought to contribute to inter-individual differences in diverse phenotypes. Although there are several tools for CNV association studies, most of them do not provide appropriate definitions of CNV regions (CNVRs), which are essential for CNV-association studies. We previously develop a user-friendly program called CNVRuler for CNV-association studies (Bioinformatics 2012 doi: 10.1093/bioinformatics/bts239). Outputs from the ten most common CNV defining algorithms can be directly used as input files for determining the three different definitions of CNVRs. CNVRuler supports chi-squared and Fisher's exact tests in addition to logistic and linear regression analyses using defined CNVRs and clinical information. Clinical information can be easily coded into a simple tab-delimited text file format. Both the false discovery rate and Bonferroni correction can be used for multiple testing with this software. CNVRuler supports the likelihood ratio test (LRT), which can be used to assess the goodness-of-fit of logistic regression models. For population stratification, CNVRuler uses principal component analysis. It calculates eigenvectors and uses 0 to 3 principal components as covariates of regression. In order to validate the performance of CNVRuler, we applied the data of 4,574 CNVs identified by Bae et al. in 500 cases of subarachnoid aneu-rysmal hemorrhage (J Hum Genet. 55, 726). The three different CNVR-defining algorithms identified different numbers of CNVRs: 1,843 CNVRs, 2,211 ROs and 2,797 fragments. We compared the lists of our CNVRs with their significant CNVs. Two significant CNVs identified by Bae et al. were detected by all three algorithms of CNVRuler. In association analyses, the two CNVRs were consistently significant in univariate models regardless of the CNVR-defining algorithm. After validation, we adopt CNVRuler to analyze the genome-wide lupus-CNV association and got two significant results. Here we present our CNV-disease association result using CNVRuler.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3608W

DWAC-seq - Dynamic Window Approach for CNV detection using next-generation targeted and full genome sequencing tag density. *V. Koval[1], J. van Rooij[1], K. Estrada[1], P. Arp[1], M. Jhamai[1], R. Kraaij[1], A. Uitterlinden[1], E. Cuppen[2,3], V. Guryev[2].* 1) Genetic Laboratory, Department of Internal Medicine, Erasmus MC, Rotterdam, Netherlands; 2) Hubrecht Institute KNAW, Utrecht, Netherlands; 3) Department of Medical Genetics, UMCU, Utrecht, Netherlands.

We present a novel method for the detection of relative copy number variants (CNVs) between a set of genomes based on next-generation sequencing (NGS) tag density. It is suitable for both targeted (e.g., exome) and full genome sequencing. The distinctive dynamic window approach allows for efficient CNV detection independent of sequenceability and mappability of genomic regions. The accurate breakpoints estimation is rigid to the sequencing noise and the presence of off-target reads. We demonstrate the sensitivity and high dynamic range of DWAC-seq by analyzing a family trio with full exome NGS data, full genome NGS data of the Genome of The Netherlands (GoNL) project, and SNP array (Illumina 550K, 610K, 2.5M, 5M) results of the Rotterdam Study (RS), a deeply phenotyped population-based cohort study of ~15,000 Dutch Caucasian elderly subjects. The tool opens wide possibilities for GWAS studies based on sequencing CNV data. DWAC-seq will be applied to full exome NGS data of 3,000 RS subjects in the CNV analysis of diseases and phenotypic traits.

3609F

Fast detection of de-novo copy number variants from case-parent SNP arrays identifies a deletion on chromosome 7p14.1 associated with non-syndromic isolated cleft lip/palate. *I. Ruczinski[1], R.B. Scharpf[1], S. Younkin[1], H. Schwender[2], M. Marazita[3], A.F. Scott[1], T.H. Beaty[1].* 1) Johns Hopkins Univ, Baltimore, MD; 2) University of Dortmund, Germany; 3) University of Pittsburgh, Pittsburgh, PA.

We define a statistic called “minimum distance” to capture differences in copy numbers between offspring and parents using genome wide SNP array data. Following segmentation of the minimum distance by circular binary segmentation, final inference regarding de novo copy number events is based on a posterior calling step. We apply both the MinimumDistance and PennCNV algorithms to simulated data, and to a study of oral clefts, validate several de novo regions by qRT-PCR, and assess the overall concordance of these two algorithms. We conclude that MinimumDistance is an effective approach for reducing technical and experimental sources of noise which can generate false positives in experimental datasets. Coupled with smoothing by segmentation and posterior probabilities to incorporate information regarding allelic frequencies, this MinimumDistance algorithm compares favorably with PennCNV, and it is much faster. Using trios not selected for phenotype as controls, we identify a region on chromosome 7p14.1 with a (genome-wide) significantly higher number of de-novo deletions in the case-parent trios ascertained through an oral cleft.

3610W

VarioML Framework for Comprehensive Variation Data Representation and Exchange. *M. Byrne[1], I.A.C. Fokkema[2], O. Lancaster[3], T. Adamusiak[10], A. Ahonen-Bishopp[9], D. Atlan[7], C. Beroud[11], M. Cornell[8], R. Dalgeish[3], A. Devereau[8], G.P. Patrinos[6], M.A. Swertz[5], P.E.M. Taschner[2], G.A. Thorisson[3], M. Vihinen[4], A.J. Brookes[3], J. Muilu[1].* 1) Institute for Molecular Medicine Finland (FIMM), Helsinki, Finland, Helsinki, Uusima, Finland; 2) Department of Human Genetics, Leiden University Medical Center, Netherlands; 3) Department of Genetics, University of Leicester, Leicester, UK; 4) Department of Experimental Medical Science, Lund University, Lund, Sweden; 5) Genomics Coordination Center, Department of Genetics, University Medical Center Groningen and Groningen Bioinformatics Center, University of Groningen, Netherlands; 6) University of Patras, School of Health Sciences, Department of Pharmacy, Patras, Greece; 7) Phenosystems Inc., Belgium; 8) National Genetics Reference Laboratory, Manchester, UK; 9) Biocomputing Platforms, Ltd., Espoo, Finland; 10) European Molecular Biology Institute, Hinxton, UK; 11) Inserm, Paris, France.

Sharing of data about variation and the associated phenotypes is a critical need, yet variant information can be arbitrarily complex, making a single standard vocabulary elusive and re-formatting difficult. Complex standards have proven too time-consuming to implement.

The GEN2PHEN project addressed these difficulties by developing a comprehensive data model for capturing biomedical observations, Observ-OM, and building the VarioML format around it. VarioML pairs a simplified open specification for describing variants, with a toolkit for adapting the specification into one's own research workflow. Straightforward variant data can be captured, federated, and exchanged with no overhead; more complex data can be described, without loss of compatibility. The open specification enables push-button submission to gene variant databases (LSDB's) e.g., the Leiden Open Variation Database, using the Cafe Variome data publishing service, while VarioML bidirectionally transforms data between XML and web-application code formats, opening up new possibilities for open source web applications building on shared data. A JAVA implementation toolkit makes VarioML easily integrated into biomedical applications. VarioML is designed primarily for LSDB data submission and transfer scenarios, but can also be used as a standard variation data format for JSON and XML document databases and user interface components.

VarioML is a set of tools and practices improving the availability, quality, and comprehensibility of human variation information. It enables researchers, diagnostic laboratories, and clinics to share that information with ease, clarity, and without ambiguity.

3611F

Gene-centered viewing, storing and sharing of exome/genome variant and phenotype data. *J.T. Den Dunnen, I.F. Fokkema, I.C. Lugtenburg, J. Hoogenboom, Z. Tatum, G.C.P. Schaafsma, M. Vermaat, J.F.J. Laros, P.E.M. Taschner.* Dept Human Genetics, Leiden Univ Medical Ctr, Leiden, Netherlands.

The favourite view of sequence variant data in DNA diagnostic centers is gene-centered. We have developed a new version of the LOVD platform (Leiden Open-source Variation Database, http://www.LOVD.nl) facilitating the analysis of exome and genome sequence data. During installation, web services retrieve gene and transcript information on the fly. Imported variant data are stored using chromosomal nucleotide positions as a reference. Data can be stored and displayed in several ways: variant-by-variant or all connected to one individual. Using the existing LOVD functionality, users have the option to perform query per gene or individual, to link to other resources of interest, to get genome browser views of the data and to using web services to access variants stored in other gene variant databases. LOVD 3 has the unique option to independently store both the phenotypes screened and the variants detected. This gives submitters the chance to share inconclusive results, allowing collaborators with matching data to join the gene identification project and crack the case together. In addition, LOVD3 has a new access level, designated “collaborator”, allowing submitters to share otherwise non-public data with other submitters, e.g., to share detailed phenotype information with other diagnostic labs only.

Funded by the European Community's Seventh Framework Programme (FP7/2007–2013) under grant agreement n° 200754 - the GEN2PHEN project.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3612W

Leveraging metadata for experimental discovery at the ENCODE Portal. *E.L. Hong[1], C. Sloan[2], V. Malladi[2], K. Rosenbloom[2], G. Barber[2], G. Binkley[1], E.T. Chan[1], R. Fang[2], B.C. Hitz[1], D. Karolchik[2], V. Kirkup[2], K. Learned[2], J. Long[2], M. Maddren[2], M. Wong[2], A. Zweig[2], D. Haussler[2,3], J. Kent[2], J.M. Cherry[1].* 1) Dept of Genetics, Stanford University, Stanford, CA, 94305; 2) Center for Biomolecular Science and Engineering, University of California, Santa Cruz, CA, 95064; 3) Howard Hughes Medical Institute, University of California, Santa Cruz, CA, 95064.

The Encyclopedia of DNA Elements (ENCODE) Project is a collaborative project to create a comprehensive catalog of functional elements in the human and mouse genomes. To date, the ENCODE project has generated 3000+ experiments, using over 30 different experimental techniques and 300+ cell lines and tissue types, in order to investigate the binding sites of over 200 DNA-binding proteins, refine the annotation of protein-coding genes and non-coding RNAs, and examine the chromatin structure. As the ENCODE project enters its 6th production year, the data volume is expected to increase, covering more DNA-binding proteins, chromatin structure, and transcription in more cell lines and tissue types. All data generated by the ENCODE project are submitted to the Data Coordination Center (DCC) for validation, tracking, storage, and distribution to community resources and the scientific community. The number of data sets and the complexity of the methods used makes identification of experiments that match the interests of a researcher challenging. To accommodate the complex and diverse needs of researchers a new facility is being created. The ability to identify and download the appropriate data is being enhanced with the expansion of metadata that will be used to describe an experiment, the development of tools that will facilitate searching based on these metadata, and summaries of the experimental data. Data from the ENCODE project can be accessed via the ENCODE portal (http://www.encodeproject.org) and the UCSC Genome Browser (http://genome.ucsc.edu/cgi-bin/hgGateway).

3613F

A Bayesian hierarchical generalized linear model for identifying multiple interacting genes in family-based case-control studies. *J. Li, N. Yi.* Biostatistics, University of Alabama at Birmingham, Birmingham, AL.

Genetic interactions and rare variants are hypothesized to be two important putative explanations for the missing heritability of complex human diseases. Jointly analyzing interactions and rare variants could remarkably improve our understanding of the genetic architecture of most complex human diseases. We recently proposed a hierarchical generalized linear model (GLM) approach that can simultaneously analyze numerous common and rare variants as well as gene-gene and gene-environment interactions for population-based genetic association studies. In this article, we extend our previous method to family-based case-control genetic association studies. We develop a hierarchical Poisson GLM that is statistically equivalent to the standard conditional logistic models for family-based case-control genetic association studies but is more computationally feasible. We employ the Multinomial-Poisson transformation technique to simplify maximization of multinomial likelihood in model fit. To facilitate the process of model fit in presence of high dimensionality and rare variants, we create a fast and stable algorithm by incorporating an expectation-maximization (EM) algorithm into the usual iteratively weighted least squares as implemented in the R package glm. Extensive simulations based on real data show that the proposed method outperforms the existing approaches in terms of statistical power and parameter estimates, especially for interactions and rare variants. The new method has been implemented in our R package BhGLM which is available free to practitioners (http://www.ssg.uab.edu/bhglm/).

3614W

A national platform for clinical genetic analysis of high-throughput sequencing data in Norway. *M.C. Eike[1], H. Lærum[2], T. Hughes[1], S. Bremer[3], S. Bergan[4], G. Thomassen[5], M. Aanestad[6], T. GRünfeld[1], D.E. Undlien[1].* 1) Department of Medical Genetics, Oslo University Hospital, University of Oslo, Oslo, Norway; 2) IT Department, Oslo University Hospital, Oslo, Norway; 3) Department of Medical Biochemistry, Oslo University Hospital, Oslo, Norway; 4) Department of Pharmacology, Oslo University Hospital, Oslo, Norway; 5) University Center for Information Technology (UCIT), University of Oslo, Oslo, Norway; 6) Department of Informatics, University of Oslo, Oslo, Norway.

The introduction of high-throughput sequencing (HTS) into clinical practice poses great challenges in terms of analytic resources, integration and data security. The project *Norwegian clinical genetic Analysis Platform* (genAP) was initiated as a response to these challenges, aiming to establish a centralised infrastructure for secure storage and analysis of human sequencing data that allows for disseminated clinical use. GenAP is a collaborative project between Oslo University Hospital and the University of Oslo, with implementation of the system in an established high-performance computing environment. As part of this project, we are in the process of establishing pilot HTS pipelines for a set of clinical packages, each with a defined set of targeted genes. The pilots cover diagnostic, prognostic and pharmacogenetic areas, including cardiomyopathies, breast cancer and tacrolimus dosage, respectively. Initially, the pipelines involve targeted capture and resequencing, but the system will be scalable to exome and genome data for a large number of patients. To reduce the workload associated with manual analysis, we have sought to achieve a high degree of automation, including variant annotation and quality control, filtering based on public resources, identification of previously classified variants, and standardization of report information structure for integration with existing patient journal systems. The pilot phase includes comparing the performance of HTS to conventional methods, building a database of genetic variants adapted to our patient populations and real-world testing using collaborating clinicians. The experiences gained will be used to expand to other clinical packages, and the ultimate goal is to introduce the pipelines for widespread use in the clinic. Complementing the genetic, bioinformatic and clinical issues, the project also addresses legal, ethical and organizational issues encountered when HTS is deployed in large scale clinical decision making.

This presentation will demonstrate the overall architecture of our system and present initial experiences with the pilot systems.

3615F

Haplotyping human genomes using whole-genome sequence data. *V. Bansal.* Scripps Translational Science Institute, La Jolla, CA.

Rapid advances in high-throughput sequencing technologies have enabled the sequencing of the genomes for hundreds of individuals. However, virtually all individual genomes sequenced using next-generation sequencing technologies are incomplete with respect to haplotype information. Humans are diploid and haplotype information is essential for the complete understanding of genomic variation both at the individual and population level. Sequence reads that span multiple heterozygous variants in a diploid genome are informative about the phase between neighboring variant sites. Computational tools for assembling these reads into longer haplotype segments and integrating these haplotypes with alternate sources of haplotype phase information are needed. We describe a computational framework for leveraging the haplotype information present in sequence reads generated using next generation sequencing technologies for phasing individual genomes. Although the relative short read lengths of current high-throughput sequencing technologies combined with the short inserts used to sequence individual genomes limit the length of haplotypes that can be assembled using this approach, we have found that that (a) haplotype information can be used to improve the accuracy of variant calling, (b) long haplotypes can be assembled for highly polymorphic regions of the genome, e.g. HLA and (c) sequence-based haplotypes can complement family and population based methods for resolving haplotypes.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3616W

Comparing Protein Prediction Methods Using Disease-Causing Missense Variants. *P. Duggal[1], Y. Kim[2], M.K. Tilley[2], M.M. Parker[1], A. Maroo[1], A.P. Klein[1,3].* 1) Dept Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD; 2) Inherited Disease Research Branch, NHGRI, NIH, Baltimore, MD; 3) Department of Oncology and Pathology, Johns Hopkins University, Baltimore, MD.

The success of next-generation sequencing studies to identify genetic variation associated with disease is dependent upon the ability distinguish between genetic variation that leads to a change in protein function and benign variation. Deciphering the role of these identified variants is often difficult because of low allele frequencies. Computational methods that classify the functionality of these variants are used to filter variants and to weight the relevance or importance of a given missense mutation in statistical analyses. We manually curated the Uniprot database, and the clinically associated variant list from the Human Genetic Mutation Database and identified 150 variants present in 1000 genomes data regardless of ethnicity. Using the OMIM database we identified 43 missense mutations from the 150 with published evidence of a functional role for the mutation, including differences in animal studies, clinical assays and binding affinities. We considered these 43 variants to be a “gold standard” since there were appreciable differences in function associated with each genetic change. Using 9 different prediction programs, we evaluated and compared the prediction performance. The programs used in this study were: SIFT, SNAP, PolyPhen-2, PMUT, VarioWatch, The Grantham Matrix Score, GERP++, SNPs&Go and MAPP. Depending on the algorithms and sources used in each protein prediction method, the prediction results were varied. Overall, the percent of variants predicted to be deleterious ranged from 51–77% using each program. Although for any given variant the percent of programs that predicted it to be deleterious ranged from 17–100%. These programs have been evaluated previously, but the use of a “gold standard” in which function was shown to be evident in the literature should provide the most informative comparison. The discrepancy in programs raises concerns about the use of an individual program to predict functionality in filtering or weighting analyses, since these may result in the exclusion of truly functional variants. This study suggests that multiple programs be considered when evaluating novel variants to limit the risk of false negatives, which will be especially important for complex diseases.

3617F

Considerations for the Processing and Direct-to-Consumer Return of Exome Sequences. *E.D. Harrington, C. McLean, A. Shmygelska, A. Chowdry, B. Naughton.* 23andMe Inc, Mountain View, CA.

In late 2011 23andMe announced our first publicly available sequencing product: the Exome Pilot Project. To return these data directly to consumers (DTC) we implemented a processing pipeline that would maximize the value to the consumer while maintaining data quality and security at each step. Enrollment in the pilot was limited to individuals who had been genotyped on the 23andMe platform. Samples were enriched using the Agilent SureSelect 50Mb platform and sequenced to at least 80X unaligned coverage using Illumina's HiSeq technology. Participants received their aligned raw data, variant calls, and a summary report describing relevant statistics and potential variants of interest based on a custom filtering process. We implemented a flexible and scalable pipeline for processing sequence data (exome and whole genome) using the Broad Institute's best practices. It employs a combination of standard tools (eg GATK) and custom software to automate the tracking of samples, data distribution to and collection from compute nodes, read mapping, variant calling, report generation, and validation against our existing genotype database. To quickly meet fluctuations in demand the pipeline can be deployed either locally or on a cloud platform. All variant calls were validated against existing genotype data when coverage on the 23andMe genotyping platform overlapped with the exome targets. For variant calls that passed our filters and had unambiguous stranding, we observed 99.6–99.9% concordance between chip and sequence data, consistent with the error rate of the chip. Lowering our stringency of filtering had a marked, though expected, effect on our concordance. Data security is integral to DTC data delivery. Data was encrypted with keys delivered via secure messaging on the 23andMe website. The encrypted raw data for each exome was on average 6GB, making bandwidth and data integrity another concern. We delivered data via Amazon S3 and the use of encryption made errors in transfer immediately obvious. Participant response to their exome data was varied but largely positive. Some participants had significant scientific background in genetics and interacted with others via the 23andMe community to provide guidance on their data. Some used their data to bootstrap research into rare diseases affecting family members, while others shared their results publicly via the Personal Genome Project. To date, no negative consequences of such data return have been observed.

3618W

Topological Mapping and Exploration of Genotyping Data using Iris. *J. Paquette, G. Singh, G. Carlsson, P.Y. Lum.* Ayasdi, Palo Alto, CA.

Genotyping samples from single-nucleotide polymorphism (SNP) arrays and next-generation sequencing (NGS) are large and complex. Furthermore, there is a systematic heterogeneity of samples due to sampling bias as well as natural patterns in the human population. These two issues, complexity and heterogeneity, confound the interpretation of genotyping samples via supervised detection of significant variants with respect to clinical status. A specific genomic variant might significantly confer sensitivity to a drug in one subgroup of a population whereas it may be less significant, or even have the opposite effect, in a different subgroup. Combining both subgroups in a supervised analysis of drug response will under-estimate the significance of the variant for the former subgroup and over-estimate the significance of the variant for the latter. Conventional methods for analysis and visualization of heterogeneity in a population of genotyping samples such as principal components analysis (PCA) often fail to indicate true patterns and similarities between samples. We make a case here for a novel approach to tackle the interpretation of large genotyping data sets, Topological Data Analysis (TDA), and present a software implementation, Iris, that leverages TDA to produce comprehensive network visualizations. TDA networks allow an investigator to visualize and understand complex patterns of heterogeneity in the overall population as well as local patterns of heterogeneity in select subgroups. The Iris software also contains tools for interactive supervised analysis via clinical data and exploration of significant variants in the chromosomal location and cellular pathway context. We present case studies applying TDA to the Wellcome Trust Consortium SNP array data set, the 1000 Genomes Project DNA-Seq data set, and a cancer genotyping data set. We contrast the differences between TDA and PCA, and highlight new variants discovered to significantly relate to disease incidence and drug response.

3619F

Development of 1920 Barcodes for Large-Scale Targeted Sequencing using the Access Array™ System. *X. Wang, F. Kaper, P. Chen, C. Friedlnader, G. Sun, A. May.* R&D, Fluidigm inc, South San Francisco, CA.

With the development and commercialization of high-throughput sequencing technologies, it has become possible to process hundreds or thousands of samples in parallel on a single sequencing run. In particular, targeted sequencing and genotyping of specific loci within the human genome can be enabled at high throughput with an appropriate barcoding strategy. Combining highly parallel barcoding with the Access Array™ System streamlines preparation of targeted sequencing libraries in large-scale genomic studies. We developed an algorithm to generate 1920 10-base barcodes (384×5) for use with any sequencing systems, including those from Roche, Illumina and Life technologies. 10-base barcodes were selected with different Hamming distances to enable accurate multiplexing of samples in the presence of single-base substitution errors. The barcodes have been further evaluated in the context of sequence tags required for multi-primer PCR amplification as implemented on the Access Array™ System. A number of sequence features were considered in barcode selection including potential homopolymers, hairpins or heteroduplex formation between tagging primers. Furthermore, the 1920 barcodes have been binned into different groups to incorporate different sequencing workflows. We will present data demonstrating successful amplification, sequencing and demultiplexing with 384 samples labeled with a subset of these barcodes. Sequence data are high quality, relative representation of samples, and amplicons within those samples have uniform coverage, and exhibit excellent technical reproducibility.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3620W

NGS Catalog: A database of next generation genome sequencing studies in humans. *J. Xia[1], Q. Wang[1], P. Jia[1], B. Wang[1], W. Pao[2,3], Z. Zhao[1,3,4].* 1) Department of Biomedical Informatics, Vanderbilt University School of Medicine, Nashville, TN; 2) Department of Medicine/Division of Hematology-Oncology, Vanderbilt University School of Medicine, Nashville, TN; 3) Vanderbilt-Ingram Cancer Center, Vanderbilt University Medical Center, Nashville, TN; 4) Department of Psychiatry, Vanderbilt University School of Medicine, Nashville, TN.

Next generation sequencing (NGS), also known as massively parallel sequencing, is rapidly transforming biomedical and biological research from single gene to genome scale. Within only a few years of the advent of NGS technologies, it is now possible to allow researchers to apply whole exome sequencing (Exome-Seq), whole genome sequencing (WGS), whole transcriptome sequencing (RNA-Seq), or a combination of them to investigate individual genome(s), especially those related to disease. These rapidly emerging NGS studies provide us with an exceptional opportunity to examine the potential impact of genetic variants on diseases by systematically cataloging and summarizing key characteristics of the observed mutations. To provide the research community with a comprehensive NGS resource, we have developed the database Next Generation Sequencing Catalog (NGS Catalog, http://bioinfo.mc.vanderbilt.edu/NGS/index.html), a continually updated database that collects, curates and manages available human NGS data obtained from published literature. NGS Catalog deposits publication information of NGS studies and their mutation characteristics (SNVs, small insertions/deletions, copy number variations, and structural variants), as well as mutated genes and gene fusions detected by NGS. Other functions include user data upload, NGS general analysis pipelines, and NGS software. NGS Catalog is particularly useful for investigators who are new to NGS but would like to take advantage of these powerful technologies for their own research. Finally, based on the data deposited in NGS Catalog, we summarized features and findings from germline mutations and somatic mutations studies for human diseases/traits.

3621F

Genotype Imputation via Matrix Completion. *E.C. Chi[1], H. Zhou[2], G.K. Chen[3], D. Ortega Del Vecchyo[4], K. Lange[1,5].* 1) Human Genetics, University of California, Los Angeles, Los Angeles, CA; 2) Department of Statistics North Carolina State University Raleigh, NC; 3) Department of Preventive Medicine, USC, Los Angeles, CA; 4) Interdepartmental Program in Bioinformatics University of California Los Angeles, CA; 5) Departments of Biomathematics and Statistics University of California Los Angeles, CA.

Most current genotype imputation methods are model-based and computationally intensive, taking days to impute one chromosome pair on 1000 individuals. We describe an efficient genotype imputation method based on matrix completion. Matrix completion aims to recover an entire matrix when only a small portion of its entries are actually observed. Making an appeal to Occam's razor, it seeks the simplest matrix consistent with the observed entries. Often this statement formally translates to searching for a low rank matrix with a small squared error difference over the observed entries. Genotypes can be coded as 0, 1, or 2's by counting reference alleles and entering the counts into a matrix whose rows are labeled by individuals and whose columns are labeled by SNPs. Despite matrix completion's purely empirical nature, it seemed natural to us to investigate its application to a central problem in modern genetics research. Specifically, we apply a matrix completion strategy for the problems of imputation in a traditional GWAS and of genotype calling in low coverage sequencing. We consider simulated data based on haplotypes from HapMap and 1000 Genomes. Compared to the gold standard MACH program, matrix completion as embodied in our program Mendel-Impute achieves similar or better imputation accuracy while reducing run time by orders of magnitude.

3622W

Phenome-wide association study of common PXDN variants demonstrates association with aortic aneurysms. *J. Denny[1], L. Bastarache[1], G. Bhave[1], S. McCall[1], M. Sample[1], R. Carroll[1], P. Peissig[2], A. Kho[3], C. McCarty[4], M. Brilliant[2], J. Cowan[1], R. Chisholm[3], E. Larson[5], G. Jarvik[6], C. Chute[7], I. Kullo[7], D. Roden[1], B. Hudson[1].* 1) Vanderbilt University, Nashville, TN, 37232; 2) Marshfield Clinic, Marshfield, WI; 3) Northwestern University, Chicago, IL; 4) Essentia Institute of Rural Health, Duluth, MN; 5) Group Health Research Institute, Seattle, WA; 6) University of Washington, Seattle, WA; 7) Mayo Clinic, Rochester, MN.

Peroxidasin (PXDN) is an extracellular matrix heme peroxidase that forms reinforcing sulfilimine bonds within collagen IV protein networks. The enhanced expression of PXDN within endothelial and vascular smooth muscle basement membranes along with its cross-linking function suggested a potential role in vascular integrity and disease. To test this hypothesis in an unbiased fashion, we conducted an electronic medical record (EMR)-based phenome-wide association study of 29 SNPs located in or near PXDN available within existing genome-wide SNP data performed for other phenotypes selected by the electronic MEdical Records and GEnomic (eMERGE) network. The study was performed in 13,582 participants of European ancestry. Logistic regression for each of the 1547 algorithmically-defined phenotypes was performed for each SNP using an additive genetic model adjusted for age and sex. The top associations were with rs12714332, an intronic SNP in PXDN, and abdominal aortic aneurysms (AAA) (odds ratio [OR]= 1.55, 95% confidence interval [95%CI]=1.27–1.88, $P=1.2\times10^{-5}$) and any aortic aneurysms (OR= 1.41, 95%CI=1.19–1.68, $P=8.0\times10^{-5}$). The threshold for phenome-wide significance is 3.2×10^{-5}. Post-hoc evaluation of the accuracy of the PheWAS algorithm for predicting AAA cases was 86%. We then sought to replicate the association in a cohort of ~3800 patients with genome-wide SNP data. All patients with a billing code for AAA were manually reviewed by clinicians with full access to the EMR to confirm the presence of AAA. The set identified 46 cases with AAA. Comparing to 2860 controls without any aneurysms, the p-value trended toward significance, P=0.07, OR=1.9. These data suggest that PXDN may be associated with AAA and other aneurysms in humans. The PheWAS method, applied to large populations with extant genotype data, may enable rapid discovery of potential clinical implications of genes of unknown significance.

3623F

Rare Variant Association Testing under Low-Coverage and Pooling. *E. Halperin[1], J.H. Sul[2], O. Navon[1], B. Han[2], L. Conde[3], P. Bracci[4], J. Riby[3], C. Skibola[3], E. Eskin[2].* 1) School of Computer Science, Tel-Aviv University, Tel-Aviv, Israel; 2) Computer Science Department, University of California, Los Angeles, California 90095, USA; 3) Division of Environmental Health Sciences, School of Public Health, University of California, Berkeley, California 94720, USA; 4) Department of Epidemiology and Biostatistics, University of California, San Francisco 94107, California, USA.

Deep sequencing technologies allow us to study rare variants involved in disease. To increase the statistical power to detect effects of rare variants, several methods have been developed that group variants in genes and test for association between a disease and the set of variants. However, detecting subtle associations requires studies of hundreds or thousands of individuals, which is prohibitively costly. Low coverage sequencing and DNA pooling, where a pool of DNA samples from multiple individuals is sequenced in one experiment, are strategies that can help reduce the cost associated with current sequencing technologies. However, current disease association methods cannot be applied directly to such data as they require individual genotypes, which are lost in pooling, and are susceptible to errors in low coverage sequencing. In this work, we propose two novel methods for the analysis of rare SNPs from low coverage, error-prone sequencing with or without DNA pooling. We show by simulations that our methods outperform previous methods under low and high coverage sequencing, with and without pooling. By analyzing real pooled sequencing data from a study of non-Hodgkin's lymphoma, we estimate the sequencing error rate and the accuracy of pooling by comparing sequencing data to previously obtained GWAS data on the same samples. We simulate different study designs based on the parameters from the real data, and show that for a given budget, it is desirable to include all available individuals in sequencing pools in order to maximize power.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3624W

BioBin: A Bioinformatics Tool for Biologically Inspired Collapsing of Rare Variants. *C.B. Moore[1,2], J.R. Wallace[2], A.T. Frase[2], S.A. Pendergrass[2], M.D. Ritchie[2].* 1) Center for Human Genetics Research, Vanderbilt University, Nashville, TN; 2) 1Center for Systems Genomics, Department of Biochemistry and Molecular Biology, The Huck Institutes of the Life Sciences, The Pennsylvania State University, University Park, PA.

There has been increasing interest in rare variants (RVs) and methods to detect their association to disease. We propose a flexible collapsing method inspired by biological knowledge. We have built a database, Library of Knowledge Integration (LOKI), which contains resources such as: the National Center for Biotechnology (NCBI) dbSNP and gene Entrez database information, Kyoto Encyclopedia of Genes and Genomes (KEGG), Reactome, Gene Ontology (GO), Protein families database (Pfam), NetPath - signal transduction pathways, Molecular INTeraction database (MINT), Biological General Repository for Interaction Datasets (BioGrid), Pharmacogenomics Knowledge Base (PharmGKB), the database of Evolutionary Conserved Regions (ECRBase), and the Open Regulatory Annotation Database (ORegAnno). BioBin can apply multiple levels of burden testing, including: functional regions, evolutionary conserved regions, genes, and/or pathways. BioBin has been tested using resequenced data with simulated variants, NHLBI GO Exome Sequencing Project Kabuki dataset, and low coverage data from the 1000 Genomes Project. BioBin was able to reliably detect causative variants in resequenced data with high sensitivity in a variety of simulated models and effect sizes. We successfully demonstrated the use of BioBin on NHLBI GO Sequencing data (Ng et al. Nature Genetics 2010). Lastly, we conducted a pairwise comparison of rare variant burden differences (MAF < 0.03) between ancestry groups (15 populations) of the 1000 Genomes Project and found marked differences. For example, between Yoruba (YRI) and European descent (CEU) individuals, we found that 56.8% of gene bins, 78.9% of intergenic bins, and 83.5% of pathway bins have significant differences in RV burden. Comparing two related populations, YRI and ASW (African ancestry in Southwest USA), we identified 0.39% gene bins, 0.60% of intergenic bins, and 0.98% of pathway bins had significant differences in RV burden. Ongoing efforts are examining regional characteristics using BioBin (evolutionarily conserved and/or regulatory). We are also exploring ancestry correction, the magnitude of potential stratification in RVs is large and inadequate correction could have large implications for sequence data analysis. BioBin is a useful, powerful, and flexible tool in analyzing sequence data and will be successful at uncovering novel associations with complex disease.

3625F

NCBI's ClinVar: data archive and tools for human variation of medical interest. *D. Maglott, S. Chitipiralla, D. Church, M. Feolo, J. Garner, W. Jang, J. Lee, R. Maiti, J. Ostell, L. Phan, G. Riley, W.S. Rubinstein, D. Shao, S. Sherry, K. Sirotkin, R. Tully, R. Villamarin, M. Ward.* Natl Ctr Biotech Info, NIH/NLM, Bethesda, MD.

The discovery of human sequence variation is outpacing development of resources and tools to maintain information about clinical significance. Clinical and experimental evidence on which interpretations are based have frequently been captured in local databases, as unstructured data or without standardized usage. Thus the primary use cases for maintaining these data—namely (1) identification of variation known to be disease-related, (2) identification of variation not thought to be disease-related and (3) automatable re-evaluation—are difficult. The National Center for Biotechnology Information (NCBI) established ClinVar (http://www.ncbi.nlm.nih.gov/clinvar/) to solve these problems by centralizing management of these data. ClinVar is a versioned archive of submissions reporting genotype, phenotype, clinical interpretation and supporting evidence, including observations in affected or unaffected individuals, animal models, in vitro assays, and in silico predictions. Once data are accumulated in standardized structures, they can be reviewed by experts. ClinVar provides infrastructure for such deliberations, and captures decisions attributed to the group contributing the intellectual effort. In other words, ClinVar accepts both submitted reviews of clinical interpretation by authoritative groups and primary data. To identify genes and variations about which some assessment of pathogenicity has been made, ClinVar integrates data from OMIM®, GeneReviews, the Genetic Testing Registry, locus-specific databases, as well as multiple testing laboratories. Common variation without such assessments are also reported. Users can interrogate ClinVar starting at different points—by genomic location, by phenotype, by gene, by associated citation—to name a few. The user can retrieve data interactively, or download reports after submitting large sets of variation calls, such as from high-throughput sequencing. Data are also accessible by interactive browsing, ftp, and NCBI E-utilities. These modes are intended to facilitate integration of ClinVar into clinical interpretation analysis workflows. ClinVar welcomes data sharing with resources maintaining information on human variation. For example, ClinVar is the data repository for the proposed U.S. node of the Human Variome Project (HVP).

This presentation will review data submission, data access, and current status of ClinVar.

3626W

Estimation of haplotype frequencies from pooled sequence data. *D.E. Kessner[1], J. Novembre[1,2].* 1) Bioinformatics, UCLA, Los Angeles, CA; 2) Ecology and Evolutionary Biology, UCLA, Los Angeles, CA.

DNA samples are often pooled, either by experimental design, or because the sample itself is a mixture. For example, when population allele frequencies are of primary interest, individual samples may be pooled together to lower the cost of sequencing. Alternatively, the sample itself may be a mixture of multiple species or strains (e.g. bacterial species comprising a microbiome, or pathogen strains in a blood sample). We present an expectation-maximization (EM) algorithm for estimating haplotype frequencies in a pooled sample directly from mapped sequence reads, in the case where the possible haplotypes are known. Our algorithm performs best when the SNP density of the strains under investigation is sufficiently high so that haplotype information can be extracted from individual reads. This condition is satisfied in, for example, inbred lines of Drosophila, and we anticipate that this method will be useful in other model organisms as advances in sequencing technology continue to increase sequence read lengths.

3627F

Leveraging the haplotype information in long reads for variant calling. *A.P. Singh, Y. Shen.* Center for Computational Biology and Bioinformatics, Columbia University, New York, NY.

Nanopore sequencing technologies are ushering the era of very long read sequences that can be used for alignment or variant calling. However, the high error rates of the reads are a matter of concern. Long reads (or jumping reads) are likely to harbor multiple polymorphic sites. Each read represents a haplotype. When the number of haplotypes is smaller than the number of reads, we can jointly infer the haplotypes of the reads and the genotypes of the putative polymorphic loci (PPL). Yun Li first used haplotype information to perform genotype calling in their package, Thunder. Zhi et. al extended the concept to use haplotype information from jumping reads for genotype calling from population data, in their HapSeq package. Their method is optimized for low depth coverage sequencing but can look at only two neighboring PPLs due to complexity of the HMM. We simplified their model and optimized it for the scenario where there are only two haplotypes, but can also take advantage of the haplotype information from three or more neighboring PPLs in the jumping reads. We assign the probability of a read belonging to either haplotype using a dynamic programming procedure based on sharing of known and novel candidate variants between two overlapping reads. We use an HMM to compute the genotype likelihood of the overlapping candidate variants in the reads. The emission in the HMM is based on the assumption that reads from the same haplotype are likely to produce the same variants at each loci. The candidate variants are obtained using bwasw by altering the mismatch and gap open penalties while allowing for large band widths in the alignment. We use the genotype likelihoods obtained from the shared set of known and novel snps as priors in calculating the posterior likelihoods of the genotypes. The priors for novel snps are downweighted compared to those of the known snps. We performed simulations to measure the accuracy of our method in inferring the haplotypes and determining the genotype likelihoods. Simulated reads were generated on chromosome 21 of the human reference genome at medium depth coverage (~10X) which included a set of phased snps and simulated mismatches and indels. We obtained an accuracy of more than 95% in determining the haplotype of the reads. We also obtained the new improved genotype likelihood values for snps that account for the haplotype information in this EM-like procedure.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3628W

A Model of binding on DNA microarrays: Understanding the combined effect of probe synthesis failure, cross-hybridization, DNA fragmentation and other experimental details of affymetrix arrays. *Y. A. Jakubek, D. J. Cutler.* Human Genetics, Emory University, Atlanta, GA.

DNA microarrays are used both for research and diagnostics. Affymetrix arrays are commonly used for genome wide association studies, resequencing, and gene expression analysis. Microarrays are often analyzed using statistical methods that ignore many of the details of the experimental protocol and the manufacturing process. The aim of this study was to develop a comprehensive model for hybridization on microarrays that incorporates these details and consequently predicts probe intensities for Affymetrix arrays more accurately. The model begins by assuming that hybridization kinetics on a chip are fundamentally similar to binding in solution (i.e. follow a Langmuir isotherm), but that differences between chip and solution binding largely arise from details of the experiment not usually considered. These details include target fragmentation, wash stringency, scanner settings, probe synthesis efficiency, and target concentration. All the parameters used in the model have well-established physical origins. For the 302 chips that were analyzed the mean correlation between expected and observed probe intensities was 0.701 with a range of 0.88 to 0.55. All available chips were included in the analysis regardless of the data quality. Our results show that batch effects arise from differences in probe synthesis, scanner settings, wash strength, and target fragmentation. We also show that probe synthesis efficiencies for different nucleotides are not uniform. To date this is the most complete model for binding on microarrays. This is the first model that includes both probe synthesis efficiency and hybridization kinetics/cross-hybridization. These two factors are sequence dependent and have a large impact on probe intensity. The algorithms developed in this work provide useful tools for the analysis of cross-hybridization, probe synthesis efficiency, fragmentation, wash stringency, temperature, and salt concentration on microarray intensities. The next step is to apply this model to SNP arrays in order to achieve better call rates and to call variants within CNVs.

3629F

A recursively partitioned mixture model for clustering time- course gene expression data. *D.C. Koestler[1], C.J. Marsit[1], B.C. Christensen[1], K.T. Kelsey[2], E.A. Houseman[3].* 1) Community and Family Medicine, Geisel School of Medicine at Dartmouth College, Lebanon, NH; 2) Department of Epidemiology, Brown University, Providence, R!; 3) Department of Public Health, Oregon State University, Corvallis, OR.

Longitudinally collected gene expression data provides an opportunity to investigate the dynamic behavior of gene expression and is crucial for establishing causal links between changes on a molecular level and disease development and progression. Clustering of subjects based on time-course expression data may lead to a more comprehensive understanding of temporal expression patterns that result in disease phenotypes. Although there are numerous existing methods for clustering subjects using gene expression data, most are not suitable when expression measurements are repeatedly collected over a time-course. We present a modifi ed version of the recursively partitioned mixture model (RPMM) for clustering subjects based on longitudinally collected gene expression data. In the proposed time-course RPMM (TC-RPMM), subjects are clustered on the basis of their temporal profi les of gene expression using a mixture of mixed effects models framework. This framework captures changes in gene expression over time and models the autocorrelation between repeated gene expression measurements for the same subject. We assess the performance of TC-RPMM using both simulation studies and a dataset from a multi-center research study of inflammation and response to injury (www.gluegrant.org), which consists of time-course gene expression data for 140 subjects. Methods for clustering subjects based on temporal gene expression pro files is a high priority for molecular biology and bioinformatics research. Simulation studies demonstrated the ability of TC-RPMM to correctly classify subjects on the basis of their temporal gene expression pro le. When applied to real epidemiologic data with repeated-measures, longitudinal gene expression measurements, TC-RPMM identi ed clusters that had strong biological and clinical signifi cance. Hence, TC-RPMM represents a promising new approach for analyzing time-course gene expression data.

3630W

Probing altered gene expression profiles and pathways that affect mitochondrial metabolism in cancer cells and neurodegenerative diseases through network biology approach. *A.Devi. Yadavalli[1], Naresh.B Sepuri[2].* 1) Dept of Bioinformatics, Karunya University, Coimbatore, India; 2) Department of Biochemistry, School of Life Sciences, University of Hyderabad.

Mitochondria are implicated in several metabolic pathways including cell respiratory processes, apoptosis, and free radical production. Mitochondrial dysfunction has been associated with neurological disorders, cell signaling and different kinds of cancers. Increasing evidence suggest that altered mitochondrial metabolism may play a significant role in onset or progression of these diseases. Little is known about the common genes or pathways, if any, that are connected to mitochondrial metabolism and disease status. To detect altered gene expression profiles and pathways that affect mitochondrial metabolism in cancer cells and neurodegenerative diseases, the public domain microarray data has been used. The differentially expressed genes are utilized to build the biological networks for each disease state. Computation of topological parameters for these networks showed the scale-free pattern and hierarchical organization. From the large network involving 1000's of genes, we synthesized sub-networks and annotated them with highly enriched biological processes. A careful dissection of the functional modules, important nodes, and their connections identified the common networks that are altered in neurological and cancer diseases. Further categorization of the identified genes may yield valuable mechanistic clues to understand the common metabolic pathways which may be of interest for developing therapeutic drug targets.

3631F

Drug repositioning through data integration and advanced classification tools. *D. Greco[1], Y. Zhao[2], F. Napolitano[3], R. Tagliaferri[3], M. D'Amato[1], J. Kere[1].* 1) Karolinska Institutet, Sweden; 2) University of Heksinki, Finland; 3) University of Salerno, Italy.

High-throughput technologies have allowed the production of an unprecedented amount of genomic data in the last decade. However, it has been very difficult to translate this information into clinical practice, and the impact on therapeutic intervention strategies is still negligible. De novo drug development is time consuming, expensive and laborious. Integrating several layers of information on drugs already in use (e.g. FDA approved), including data on their chemical and physical properties, their mode of action and their clinical effects, may help to highlight new therapeutic indications for translational medicine. In this study, we have re-analyzed a large collection of genome-wide gene expression data from human cell lines treated with several hundred drugs. We have found that previously published preprocessing of this data set is heavily biased, and does not allow precise estimation of the mode of action of the drugs. Conversely, our novel approach allows the identification of gene expression signatures that better reflect similarity between drugs. Further, we have assessed the relatedness of drugs from multiple layers of information concerning their chemical structures, intracellular molecular targets and known clinical effects. We show that, only when different similarity matrices are combined, highly successful prediction of therapeutic indication can be achieved. We systematically explore this possibility by building a computational classification system based on cutting edge machine learning algorithms. Finally, we use this classification tool also for guiding the selection of genes whose expression patterns best describe the activity of the drugs as well as their chemical, physical and clinical properties. Altogether, our results are valuable both in academic and industrial environments as they are useful for inferring the behavior of new drugs as well as for reliably operating drug repositioning.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3632W

Gemini: a flexible, scalable analysis framework for medical and population genomics. *U.D. Paila[1], A.R. Quinlan[1,2].* 1) Center for Public Health Genomics, University of Virginia, Charlottesville, VA; 2) Department of Public Health Sciences, University of Virginia, Charlottesville, VA.

Motivation. Genetics researchers have a wealth of affordable sequencing technologies for the discovery of genetic variation on a population scale. However, existing data analysis tools struggle to scale to the size and complexity of modern datasets and limit discovery because they are unable to integrate the wealth of informative genome annotations (e.g., ENCODE, OMIM, dbSNP).

Results. We have developed GEMINI, a novel framework for mining large, complex genomics datasets involving thousands of samples, as well as a suite of data analysis, exploration, and visualization tools. Our framework integrates and annotates population-scale genomic variation data in the VCF format with a growing collection of genome annotations from resources such as dbSNP, OMIM, KEGG, and the 1000 Genomes Project. The integrated data are stored in an efficient database framework facilitating intricate SQL queries and an intuitive interface to the data. Our database stores variation and annotations on a genome scale, and, uniquely, provides access to genotypes for thousands of individuals. We provide users with built-in analyses for quality control, standard population genetics metrics (e.g., HWE, Pi) as well as common data analysis tasks. Users can interact with our database through ad hoc SQL queries and importantly, researchers can easily integrate their own annotations into the GEMINI framework. Using this framework, we are able to quickly develop tools for medical genetics using various disease models and can easily integrate new analysis techniques such as “collapsing” methods. While most existing analysis tools are restricted to SNP variants, GEMINI represents SNPs, INDELs and structural variants, allowing researchers to explore all forms of genetic variation in the context of their disease phenotype or population analysis. To illustrate the power and flexibility of this framework, we demonstrate genome-scale analyses on genome-wide datasets from the 1000 Genomes Project. **Conclusion.** GEMINI is a scalable and easy to use analysis resource for medical and population genomics research. Its novel analysis and data integration features facilitate data exploration and discovery by enabling innovative analyses of genome variation.

3633F

NOCALLER - a tool to account for missing genotypes. *Z. Kronenberg[1], C.D. Huff[2], M. Yandell[1].* 1) Human Genetics, Eccles Institute of Human Genetics, Salt Lake City, UT; 2) Department of Epidemiology, The University of Texas M.D. Anderson Cancer Center, Houston, TX.

Next-generation sequencing data are now commonly used to identify disease-causing alleles, to infer population structure, and to identify genomic regions under selection. These analyses often do not account for missing data (no-calls). No-calls are placeholders for regions of the genome where sequence quality and/or depth of coverage preclude genotype inference. These regions are often ignored and treated as homozygous reference; this introduces systemic errors into downstream analyses, often with disastrous consequences. NOCALLER has been created to solve this problem by adding no-calls to files in Genome Variation Format (GVF). NOCALLER relies on depth-of-coverage and sequence quality data together with probabilistic modeling to identify no-call regions in exomes, whole genomes, and pooled sequencing data. NOCALLER's GVF-compliant output can be directly used in the VAAST suite of tools. Using simulations and 1000 genomes data we show that no-calling results in more accurate estimates of the minor allele frequencies and improves the accuracy of disease-gene searches.

3634W

A novel model to predict splicing consequences of intronic nucleotide substitutions in the human genome. *A. Shibata.* Nagoya University Graduate School of Medicine, Nagoya, Japan.

The next generation sequencers disclose a large amount of single nucleotide variations (SNVs) in the human genome. Tools to analyze effects of SNVs on the protein functions have been recently published PolyPhen-2 and SIFT. Effects of SNVs on pre-mRNA splicing, however, remain mostly nebulous. Here we developed a novel algorithm, the Aberrant Splicing Finder (ASF), to predict aberrant splicing events. We first extracted splicing cis-elements from the annotation database, ENSEMBL 62, that is based on the human reference sequence, GRCh37/hg19. Using the neural network modeling, we generated an algorithm to predict alternative splicing events from primary sequences of splicing cis-elements. The splicing cis-elements included the branch point sequence, the polypyrimidine tract, a nucleotide at intron -3, exonic length, and the 5' splice site. Our algorithm efficiently predicted alternative splicing events and yielded a correlation coefficient of 0.686 on the validation dataset. We also calculated ASF scores of benign SNPs in the dbSNP134 database, as well as of pathogenic mutations in HGMD, and found that benign SNPs reduce the scores by -0.00122 on average, whereas mutations reduce them by -0.0282 ($p < 0.001$). Our algorithm also disclosed that G at intron -3 has a prominent negative regulatory effect on pre-mRNA splicing.

3635F

Prioritizing disease-linked variants, genes, and pathways with an interactive whole genome analysis pipeline for the MedSeq project. *S. Kong[1], K. Lee[2], K.B. Hwang[3], J.M. Bohn[1], R.C. Green[4], I.S. Kohane[1,2].* 1) Informatics Program, Boston Children's Hospital, Boston, MA; 2) Center for Biomedical Informatics, Harvard Medical School, Boston, MA; 3) Soongsil University, Seoul, South Korea; 4) Genetics, Brigham and Women's Hospital, Boston, MA.

The increased availability and maturation of whole genome sequencing technology has shifted the researcher's burden from genetic variant identification to the interpretation of large numbers of variants per individual, and has fueled the development of a multiplicity of bioinformatics pipelines for whole-genome and exome analysis. Many of these focus on processing raw sequence data to detect highly confident genomic variants without providing any downstream annotation or filtering measures. Where such measures do exist, they are generally limited to few annotation options, static filtering, and lack of multiple-genome comparison. We have considered what combinations of annotations and statistical analyses would meet the growing needs of investigators using Whole Exome Sequencing/Whole Genome Sequencing in medical research and eventually in medical care settings such as the MedSeq project (HG006500). These considerations have led to the design and implementation of gNOME, a freely available interactive downstream analysis pipeline that combines diverse genomic annotation sources with statistical analysis methods in a scalable framework. gNOME consists of 3 core components: a knowledge database, a search engine, and an interactive web-based user interface. The pipeline's modular architecture allows further variant annotations, such as MutationTaster score, to be easily added to the database if desired. A user can interactively drill down to a small number of disease/phenotype linked variants by using dynamic reports that enable the investigation of personal genomes at three different levels simultaneously: variant-level, gene-level and gene set-level. Additionally, gNOME can apply non-parametric statistical tests to variant- and gene-level counts of filtered variants between two groups, and supports gene set enrichment analyses for biological pathways and known disease-linked genes. Sets of publicly available exomes and genomes such as the 1000 Genomes Project are prepared as comparison groups to reduce false positive incidental findings as well, and unlike most existing tools, our pipeline is easily scalable to accommodate thousands of genomes. We hope that, in combining these strengths, gNOME will be of use to the biomedical research community.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3636W

The 1000 Genomes Project, data availability and accessibilty. *L. Clarke, H. Zheng Bradley, R. Smith, I. Streeter, E. Kulesha, I. Toneva, B. Vaughan, P. Flicek, 1000 Genomes Project Consortium.* EBI, The Wellcome Trust Genome Campus, Hinxton, CB10 1SD, Cambridge, United Kingdom.

The 1000 genomes data sets continue to represent one of the largest public variation resources available to the community. Providing coherent and useful resources based on this data continues to be a key goal for the project Data Coordination Center (DCC). We present here a selection of the tools built on top of the 1000 genomes data to make it as useful as possible to the wider community. Finding data sets on the 1000 genomes FTP site can be a challenge with a 279 Tbyte data set spread over more than 260,000 files. We provide both an index in the current.tree file and a search based on this index in the form of http://www.1000genomes.org/ftpsearch to give all levels of users easy access to the complete set of files we have available. Our variant sets are all provided in VCF format, which represent a very data rich format. We have developed several tools to allow users to query theses files in a meaningful way. Many users are interested in a specific region of the genome rather that genome wide data. All our BAM and VCF files are indexed to allow retrival of slices of this data. We also provide a data slicing tool that provides users with subsections of bam and vcf files and also allow vcf files to be subset on the basis of individual or population. We have developed a variation pattern finding tool that organizes variants with functional consequences according to which individuals share particular genotypes. This allows uses to query which individuals an inheritance pattern for a given set of variation. Many users want to be able to view the linkage disequalibrium patterns for the 1000 genomes variant sets using tools like haploview. We now provide a tool that will convert our vcf files into the pedigree and locus information files that these visualization tools need. All of our BAM and VCF files are now available in the Amazon Cloud in a public s3 bucket. We also provide a simple Amazon Machine Image for interacting with this data. We also ensure users can keep up with new data releases with through several means including an email list (1000announce@1000genomes.org), rss and twitter (@1000genomes.org).

3637F

Bioinformatics optimization for the detection of low-level heteroplasmy in the mitochondrial genome in myelodysplastic syndrome. *S. Dames[1,2], E. Duncavage[1], J. Thompson[1], P. Shami[4], K. Eilbeck[2], M. Salama[1,3], R. Mao[1,3].* 1) Res & Development Sequencing, ARUP Lab, Salt Lake City, UT; 2) Department of Biomedical Informatics, University of Utah School of Medicine, Salt Lake City, UT, USA; 3) Department of Pathology, University of Utah School of Medicine, Salt Lake City, UT, USA; 4) Department of Hematology, University of Utah School of Medicine, Salt Lake City, UT, USA.

Background: Next generation sequencing (NGS) allows for the detection of low-level heteroplasmy in the mitochondrial genome (mtDNA). ARUP offers an mtDNA assay and reports heteroplasmy levels >10%. Different alignment and variant call strategies generally produce concordant results, but not always. These differences are often seen with heteroplasmy levels <10%. In an effort to accurately detect variants with heteroplasmy levels <10%, we compared different alignment and variant calling strategies using isolated stem cell DNA obtained from individuals with Myeloid Dysplasia myelodysplastic syndrome (MDS). For each patient, samples are flow-sorted using markers CD3+ (early T-cell), CD33+ (myeloid), and CD34+ (hematopoietic progenitor) to look for differences in the mtDNA sequence. **Methods**: mtDNA is enriched by long range PCR and sequenced on an Illumina HiSeq 2000. Sequences are aligned using BWA or CLCBio using different quality metrics and seed length parameters. Alignments to hg19, hg19 with chrM removed and replaced with NC_012920 (mtDNA refseq), and NC_012920 itself are compared. Refining of alignments and variant calls is performed using Samtools, Picard, GATK, and CLCBio. Different metrics for each step were compared including PCR duplicate removal and subtraction of baseline error rates, among others. **Results**: Results suggest the human genome should not be included in alignment due to areas of high homology that may skew heteroplasmy percentages and limit mtDNA coverage. PCR duplicate removal, varying seed lengths and different base and alignment quality metrics all alter low-level heteroplasmy levels, but to varying degrees. At base and alignment quality scores are >30, coverage of >200 is sufficient for the detection of >10% heteroplasmy, but should be >1,000 for <10%. When stringent quality metrics are applied, excessive coverage does not improve heteroplasmy accuracy. **Conclusions**: The initial phase of this project is complete, but further sequencing and analysis for reproducibility is underway. Results show the average background at a given base for the HiSeq is 1.4–1.6%, making it difficult to detect heteroplasmy in the 2–3% range. Furthermore, secondary verification of heteroplasmy levels <5% is difficult. We hope hypothesize that optimized alignment and variant call metrics will not only improve low-level heteroplasmy detection in mtDNA, but may also be extrapolated to potentially detect variants in mixed cell populations.

3638W

Rare Variant Discovery in Illumina Next Generation Sequencing Data: Prediction Method to Differentiate True variants from False positives. *J. Durtschi[1], R.L. Margraf[1], K.V. Voelkerding[1,2].* 1) ARUP Institute for Clinical and Experimental Pathology, Salt Lake City, UT; 2) Department of Pathology, University of Utah School of Medicine, Salt Lake City, UT.

Variant discovery for rare genetic diseases using Illumina genome or exome sequencing involves screening of up to millions of variants to find only the one or few likely causative variants. Sequencing or alignment errors create “false positive” variants. These false variants are often retained in the screening process, since they are not usually in databases used to filter out common variants. Methods to remove false variants consider parameters such as base or mapping quality or alignment biases. While somewhat effective, these methods can miss many false variants. As part of a rare variant discovery pipeline, we prioritize variants into 4 bins based on a false variant likelihood prediction. Our presented method uses the Genome Analysis Toolkit (GATK) variant calling software to calculate the variant-to-wildtype genotype likelihood ratio at each position. The Phred-scaled ratio is divided by read depth (to minimize the strong, linear effect of depth). This likelihood-by-depth (LBD) is calculated for a proband of interest as well as for approximately 40 exome and whole genome samples (proband's family or unrelated) all run on the Illumina HiSeq at our facility. At variant sites with no apparent sequencing error or alignment biases, wildtype calls cluster near -3 LBD and variant calls typically cluster above 10 LBD. Sites with systematic variant call problems (e.g., low coverage, sequencing error, or alignment error, as evident by GATK quality values and biases as well as displayed on the iGV viewer) show broadening and merging of these clusters, indicating lowered variant call accuracy. Depending on the separation of the proband's variant LBD value from the cluster of wildtype calls from other samples, the variant is put into one of the four bins of false variant likelihood with bin 1 most likely true and bin 4 most likely false. A comparison of two sets of variants from one genome sample that were either mostly true (20,000 chromosome 1 variants seen at between 10 and 20 % allele frequency in the 1000 genomes data) or mostly false (154 novel proband variants, not within the proband's family or in the 1000 genomes). The likely true variant set resulted in 85, 14, 0, and 1 % in bins 1, 2, 3, and 4, respectively. The likely false variant set resulted in 5, 27, 33, and 36 % similarly ordered. This indicates correspondence between real variants and bins 1 and 2, leading to more accurate prioritization of variants for subsequent analyses.

3639F

Memory Efficient Assembly of Mammalian Size Genome. *F. Hormozdiari, E. Eskin.* Computer Science, UCLA, Los Angeles, CA.

The ability to detect the genetic variations between two individuals is an essential component for genetic studies, where obtaining the genome sequence of both individuals is the first step towards variation detection problem. The emerge of high-throughput sequencing (HTS) technology has made the DNA sequencing practical, and is widely used by diagnostician to increase their knowledge about the causing factor for diseases. As HTS advances more data are generated every day than the amount that scientists can process. Genome assembly is one existing method to tackle the variation detection problem. The de bruijn graph formulation of the assembly problem is widely used in the field; furthermore, it is the only method which can assemble any genome in linear time. However, it requires an enormous amount of memory in order to assemble any mammalian size genome. The high demands of sequencing more individuals and the urge to assemble them are the driving force for a memory efficient assembler. In this work we propose a hybrid method which builds the de bruijn graph while consuming lower memory. Moreover, our proposed method can reduce the memory usage by 37%; compared to the existing methods. In addition we used a real data set (chromosome 17 of A/J strain) to illustrate the performance of our method.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3640W

Target capture-assisted sequence assembly for accurate genotyping of insertions and deletions. *C.J. Kennedy, N. Chennagiri, MA. Umbarger, GJ. Porreca, PC. Saunders, V. Greger.* Good Start Genetics, Cambridge, MA.

Alignment of short DNA sequences to a reference genome is the standard first step for next-generation sequencing (NGS) data analysis. While providing global genomic coordinates and variant information to individual reads, alignment often fails to detect insertions and deletions (indels) of significant size relative to the read length. De novo sequence assembly, by contrast, generates contiguous sample sequence (contigs) and generally improves indel detection; however, assembly is computationally intractable for very high volume data, depends critically on the extent of read overlap, and may disregard important read-based information such as coverage depth or base qualities. Here we describe a hybrid approach that uses target capture design to rapidly map individual reads to their constituent exons. Reads are assembled into contigs, which are then aligned to the reference with improved accuracy due to enlarged sequence context. The coordinate and variant information from this assembly is then transitively associated with the corresponding raw reads to enable coverage- and base quality- enabled genotyping. We sequenced exons of 14 genes associated with monogenic disorders from 150 samples chosen for their ethnic diversity or clinical significance. Sample-exons were sequenced to an average depth of 1500x with 85bp single-end (Illumina) reads and compared to corresponding bi-directional Sanger sequence. Assembly-derived genotypes were concordant for all 137 Sanger-identified insertions (up to 12bp) and deletions (up to 18bp) with no false positives. Alignment alone failed to detect 6 of 26 unique indels (23%) including compound heterozygous alleles within the highly polymorphic signal peptide sequence of SMPD1 (Niemann-Pick disease) as well as c.2207_2212delinsTAGATTC, which is the most common disease-causing mutation in BLM (Bloom syndrome) for individuals of Ashkenazi Jewish heritage. To test the robustness of our approach, we simulated data for rare pathogenic variants for which no sample data were available. All simulated variants, including 31 that were clinically relevant but undetectable by alignment alone, were correctly genotyped using our hybrid assembly algorithm. Clinical NGS applications apply new rigor and acceptability criteria to variant detection. We present an integrated analysis method that uses target capture-assisted assembly for accurate, robust, and scalable genotyping of indels.

3641F

CLC bio's integrated framework for identification and comparison of genomic variants in mendelian diseases. *M. Matvienko[1], A. Joecker[2], C. Boysen[1], U. Appelt[2], A. Fejes[2], L. Kahns[2], S. Mønsted[2], J. Grydholt[2], B. Knudsen[2], M. Bundgaard[2], J. Buur Sinding[2], H. Handberg[2], A.-M. Hein[2], M. Nygaard Ravn[2], A. Joecker[2], M. Værum[2], R. Forsberg[2].* 1) CLC bio, Cambridge, MA; 2) CLC bio, Aarhus, Denmark.

The use of personal genomics for predictive medicine is becoming increasingly popular in clinical settings due to its potential for identifying optimal drug treatments for particular individuals based on their genetic background. However, the bioinformatics analysis underlying the foundation of personal genomics, including the identification of “disease relevant” genomic variants, is still a bottleneck. CLC bio's Workbench and Server solution has been extended to an integrated framework for comparison and functional analysis of genomic variants with the purpose of supporting clinicians and scientists in the identification of clinically relevant variations. In this work, we show how the analysis of a publicly available data set from 11 patients with inherited hearing loss can be performed using the CLC Workbench. We will show the complete re-sequencing workflow, from mapping of the sequencing reads to the reference sequence, through to the identification of genomic variants. Furthermore, we demonstrate a process for the annotation, filtering and comparison of identified genomic variants to emphasize clinically relevant variants and interesting new findings.

3642W

Sensitivity and Resolution of Whole Chromosome Mapping for Detection and Characterization of Structural Variations in Human Genomes. *R. Moore.* OpGen, Gaithersburg, MD.

Structural changes in human genomes such as copy number variations, insertions, deletions, inversions, and translocations have been associated with a wide variety of clinical conditions including developmental delay, degenerative disorders, and cancer. Classically, chromosomal structural variations have been investigated with cytogenetic techniques such as Fluorescent In Situ Hybridization (FISH) but these techniques have limited resolution, typically > 2 Mbp. More recently, techniques such as Next Generation Sequencing (NGS) and array Comparative Genomic Hybridization (CGH) have been used to investigate genomic structural variants. However, NGS cannot detect variants larger than a few kilobase pairs without the use of costly and inefficient techniques such as fosmid or BAC libraries. Array CGH detects only unbalanced events such as copy number variation and does not provide information on the location of additional copies.

Here we investigate the performance characteristics of Whole Genome Mapping (WGM) as applied to the detection of structural variations in human chromosomes. WGM uses restriction maps from long individual DNA molecules to construct high resolution maps of human chromosomes. These maps can be compared between samples or to maps derived from sequence to detect structural variants. In this study human genomes previously characterized with FISH, array CGH, and NGS were analyzed with WGM to benchmark the performance of WGM for structural variant detection. Results demonstrate that WGM detects > 95% of homozygous structural variations larger than 2kbp. This was true for all types of structural variations including both balanced and unbalanced events.

This study suggests that WGM is an excellent compliment to NGS for complete evaluation of human genomes. NGS provides detection of single nucleotide polymorphisms and small structural variants. WGM detects structural variants larger than a few kbp and could eliminate the need for array CGH and FISH. We expect the WGM will be a powerful tool for the investigation of the role of structural variation in human disease.

3643F

Comparison of next-generation sequencing alignment programs using 215 whole genomes from the Cache County Study on Memory Health and Aging. *P.G. Ridge Jr[1,2], J.D. Durtschi[2], C.D. Corcoran[3,4], R.G. Munger[4,5], Q.O. Snell[6], M.J. Clement[6], K.V. Voelkerding[2,7], J.S.K. Kauwe[2].* 1) Department of Biology, Brigham Young University, Provo, UT; 2) ARUP Institute for Clinical and Experimental Pathology, Salt Lake City, UT; 3) Department of Mathematics and Statistics, Utah State University, Logan, UT; 4) Center for Epidemiologic Studies, Utah State University, Logan, UT; 5) Department of Nutrition, Dietetics, and Food Sciences, Utah State University, Logan, UT; 6) Computer Science Department, Brigham Young University, Provo, UT; 7) Department of Pathology, University of Utah School of Medicine, University of Utah, Salt Lake City, UT.

Introduction: For many years the primary barrier impeding the analysis of full genome sequences was the cost and time required for sequencing. With the advent of next generation sequencers the challenge has shifted to downstream processing of the millions of short reads generated by sequencers: alignment, variant detection, and variant interpretation. There are many programs available for each step, but the question remains: which of the programs, or combination of programs is most accurate? Comparisons of new algorithms against existing programs have done little to resolve the issue since reported results are inconsistent. Here we provide an unbiased comparison of four of the most frequently used alignment programs: BWA, Bowtie2, Novoalign, and SOAP. For our comparison we will use 215 complete genomes. The accuracy of each alignment is measured by comparing concordance rates of detected variants against genotyped data (arrays and Sanger sequenced variants).

Methods: Genomes were sequenced by Illumina on the HiSeq2000. Using the BAM files provided by Illumina, we used bam2fastq to regenerate fastq files. Parameters for each alignment program were selected based on the recommendations by the developers of each program for Illumina paired end, 100 bp reads. Variants were called using SamTools. All data processing was performed on Brigham Young University's supercomputer, Marylou (Marylou.byu.edu).

Results: The 215 genomes were sequenced to an average depth of 37. Based on our preliminary data, Bowtie2 was the fastest program requiring approximately 13 hours to perform an alignment, followed by SOAP (~28 hours), BWA (~52 hours), and Novoalign (~56 hours). As expected, identified variants were different for each program. Novoalign resulted in the highest number of unique variants (141854) and SOAP the fewest (23964). Collectively, all four programs detected 3329437 variants. We will validate observed common variants using genotypes obtained from Illumina 2.5M arrays. A set of rare variants will be selected for validation using Sanger sequencing to investigate the accuracy of the SNP calls obtained from each alignment method.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3644W

Continuously Evolving Informatics at a Medium-Scale Genomics Center. *L. Watkins, M. Barnhart, J. Goldstein, S. Griffith, E. Hsu, K. Roberts, D. Snyder, B. Craig, K. Hetrick, K. Doheny, Center for Inherited Disease Research.* Center for Inherited Disease Research (CIDR), Johns Hopkins University, Baltimore, MD.

The Center for Inherited Disease Research (CIDR) provides high quality next-gen sequencing (NGS), genotyping, and statistical genetics consultation to investigators working to discover genes that contribute to disease. Rapidly advancing technologies for interrogating genomes and the concomitant explosion of data create a formidable challenge for sustaining throughput of 80–100,000 samples per year on a demanding schedule while maintaining the utmost data quality. CIDR began in 1996 with human STRP 400-marker linkage panels, switched to SNP genotyping in 2004, and now runs arrays up to 5 million SNPs. With the addition of a next-gen sequencing service in 2011 and establishment of the joint Baylor-Hopkins Center for Mendelian Genomics, CIDR now releases data for about 5 studies per month varying from a few to over 35,000 samples each. Our present (May 2012) capacity for GWAS or epigenetic genotyping is 145,000 samples/year; for custom genotyping, 138,000 samples/year. Our sequencing workflow can process a 4,000 sample project targeting a custom 3Mb region, 150 whole exomes, or 15 whole genomes in 16 weeks, with multiple projects in progress simultaneously. A typical sequencing project now totals around 14TB of data per run; combined with many TB of genotyping data per week, this yields a genomic data deluge. We continuously evolve our informatics to meet these challenges by paying close attention to every step of the process and optimizing each one via improvements at every level of operation, from speeding up data movement (e.g., moving to 10 gigabit Ethernet), to writing better and smarter software (e.g., our own sample de-multiplexer is several times faster than the program it replaced), to rigorous data triage and compression. Our sequence analysis pipeline, CIDRSeqSuite, now takes 60% less time to process 5 times as much data per run as it did in 2010. We also leverage the best commercial products as well as our own software: we wrote a genotyping LIMS to maximize flexibility and productivity, adapted and extended a commercial LIMS for sequencing, and we're completely rewriting our entire sample- and problem-handling system to accommodate current and future demands. Current work emphasizes better and faster analysis tools, improved approaches to parallelism (e.g., GPU and hybrid-core compute platforms), sophisticated data staging, improved storage I/O using solid-state disks, and advanced hardware- and reference-based compression for data reduction.

3645F

Genotype Concordance Between Low-Coverage Whole Genome and High-Coverage Exome Sequencing: Results from the UK10K Study. *H. Zheng on behalf of the UK10K Consortium Cohorts Group.* Department of Medicine, Human Genetics, Epidemiology & Statistic, Lady Davis Institute, Montreal, Quebec, Canada.

Background: Given the high cost of large, next-generation sequencing studies there is considerable debate as to the optimal study design to accurately capture rare variants. While low-coverage whole genome sequence provides an accessible survey of genotypes across the entire genome, including regulatory regions, it is unclear if low-coverage studies yield sufficient genotypic accuracy to justify their utility. Aim: To compare the genotype concordance rates of low coverage sequencing and high depth sequencing in the same individuals. Methods: Low-coverage whole genome sequence (WGS) data, generated from the UK10K project (median depth per variant=6), and high-coverage whole exome sequence (WES) (median depth per variant =65.6) were compared for genotypic accuracy for 157 samples from TwinsUK. Variant calls from 22 autosomes were compared using VCFtools for variants called in both WGS and WES. Results: The 157 samples had 178,112 sites in common between WGS and WES. 190 sites had missing genotypes. The overall genotype concordance rate was 99.41%. 177,635 reference alleles had concordant genotypes, and 287 reference alleles displayed discordant genotypes. The concordance rate for reference allele homozygote, heterozygote and non-reference allele homozygote were 99.6%, 98.33% and 98.86%, respectively. The non-reference concordance rate was 96.8%. Concordance rates increased with the allele depth and allele frequency, yet remained high even for alleles of low frequency (99.1% for variants with a non-reference allele frequency between singletons and 1%). Discordance rates were also observed to be higher at specific genomic locations. Conclusion: The concordance rate between low coverage sequencing and high depth sequencing is high. These findings have important implications for sequencing study design and suggest that low coverage sequencing is able to accurately call genotypes.

3646W

Genotype calling and haplotype inference for next-generation sequencing data incorporating haplotype information in sequencing reads. *D. Zhi, K. Zhang.* Dept Biostatistics, Univ Alabama, Birmingham, Birmingham, AL.

A critical, but often overlooked advantage of sequencing over array-based genotyping is that sequencing provides a sequential readout of nucleotides along a single chromosome, and thus providing direct haplotype information. However, most existing genotype calling and haplotype inference methods only use read nucleotide and quality information at individual sites, and thus bypass such useful information. We recently developed a LD-based method that integrates haplotype information in reads into the commonly used chromosome segment sharing hidden Markov models [1]. While achieve improved accuracy to MACH/Thunder [2], our method only used read joint allele counts across two consecutive sites. In this work we developed the notion of haplotype likelihood, which is parallel to the commonly used genotype likelihood, but capturing haplotype information across multiple variant sites. Using simulation and real data, we show that our method achieve superior genotype calling accuracy than competing methods. 1. Zhi, D., et al., Genotype calling from next-generation sequencing data using haplotype information of reads. Bioinformatics, 2012. 28(7): p. 938–46. 2. Li, Y., et al., MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. Genetic epidemiology, 2010. 34(8): p. 816–34.

3647F

Exploring why whole genome sequencing methods result in different variant calls by comparing and integrating multiple datasets. *J.M. Zook, D. Samarov, M. Salit.*Biochemical Science Division, National Institute of Standards and Technology, Gaithersburg, MD.

As it is becoming possible to sequence even large whole genomes deeply, systematic errors and biases become very important. Integrating data from multiple platforms and algorithms can improve variant call accuracy and allow interrogation of the reasons behind method-specific errors and biases. To generate well-characterized human genome reference materials that help assess and improve performance of sequencing methods, we are comparing and integrating the large diversity of data from multiple versions of Illumina, Complete Genomics, SOLiD, and 454 on the CEU sample NA12878. Hundreds of thousands of differences exist between whole genome variant calls on this sample from the variety of sequencing platforms, mapping algorithms, and variant calling pipelines. Recent publications have demonstrated these differences when comparing just two different platforms, but have not explored the reasons for these differences or developed methods to increase sensitivity and specificity of variant calls by integrating multiple data sets. To arbitrate between variant calls that are discordant among datasets, we first find the distributions of various characteristics (e.g., strand bias, coverage, base quality score bias, mapping quality, and allele balance) for high-confidence variants for each dataset. Then, for discordant calls, we only use calls from each individual dataset that fall within the typical distribution of each characteristic. These integrated consensus variant calls have higher accuracy than any individual dataset, when compared to fosmids finished by Sanger sequencing or OMNI array data for NA12878. Therefore, we use our consensus calls to assess the accuracy of whole genome sequencing methods, and show that using microarrays to assess accuracy underestimates genome-wide false positive and false negative rates. In addition, we determine characteristics of false positive and negative calls to show strengths and weaknesses of platforms and algorithms and how they could be improved. By integrating the large amount of current and future data for NA12878, we will develop an increasingly sensitive and specific set of variant calls for this individual. DNA from this individual and others will be developed into NIST Reference Materials, which can be used by research, commercial, and clinical laboratories to assess and improve performance of sample preparation, sequencing, and bioinformatic methods.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3648W

Next Generation Biomarker Screening: Using Roche's 454 Sequencing Platform to Identify Somatic Biomarkers. *C. Lasyone, X. Tan, A. Pond, L. Szkotnicki, V. Venegas, K.B. Thomas, M. Berry, F. Lu.* SeqWright, Inc., Houston, TX.

It took half a century and three billion dollars to sequence the first human genome. Now, next generation sequencing technology makes it possible to screen entire genomes or specific genetic loci for mutations, producing actionable data for personalized, targeted treatment and patient screening. In addition, deep sequencing can identify biomarkers below the level of detection for most diagnostic tests. The ability to detect somatic mutations present in low percentages of the test sample holds enormous implications for the diagnosis and treatment of diseases such as cancer, and the confident validation of somatic biomarkers is critical to the development of new companion diagnostics for screening such diseases. Sanger sequencing is a widely established, validated sequencing method accepted by the Food and Drug Administration (FDA). Although accurate and reproducible, Sanger sequencing has a limit of detection of 20%, a sensitivity level insufficient to identify some clinically relevant somatic mutations. At SeqWright, we overcome these challenges by designing customized biomarker validation assays using next-generation sequencing on Roche's 454® GS-FLX platform which provides several thousand-fold depth of coverage, and therefore has the sensitivity to identify sequence variants present in much lower percentages. Using a subcloned region of human chromosome 7 with a known single nucleotide polymorphism (SNP), we created a limiting dilution series of SNP containing plasmid DNA spiked into wild type plasmid sample. Using this reference sample set, we then are able to demonstrate the accuracy, repeatability, and limit of detection of our 454 amplicon sequencing protocol. With a coverage depth ranging from 7,587 to 30,790, we demonstrate the accuracy of the platform to exceed 99%, capable of reproducibly detecting a SNP present in only 1% of the total DNA target population. Once validated, the assay can be used for the detection of the biomarker within biological samples within this range. Using Roche's 454® GS FLX sequencing platform, SeqWright provides a highly sensitive method for the detection of low percentage mutations, currently below the limit of detection of current sequencing technologies. Validated next-generation sequencing assays are critical to the development and marketability of companion diagnostics for low abundance mutations.

3649F

***De Novo* Assembly of extremely long single-molecule genome maps imaged in Irys nanoChannel arrays.** *M. Requa, M. Austin, H. Dai, P. Deshpande, O. Hampton, H. Sadowski, M. Saghbini, M. Xiao, H. Cao.* BioNano Genomics, San Diego, CA.

Despite significant advances in shotgun sequencing technology, *de novo* genome assemblies using only short read data are generally incomplete due to the complexity found in large genomes. This complexity, consisting mainly of large duplications and repetitive regions, hinders sequence assembly and subsequent comparative analysis. We present a single molecule genome analysis system based on a nanochannel array technology that resolves these sequence ambiguities. This technology provides high throughput sequence motif maps of single molecule fragments hundreds of kilobases in size. Capitalizing on the information encoded in the extremely long single molecule maps, assembly algorithms unique to this format deliver high-resolution whole-genome sequence motif maps. Parallel assembly by shotgun sequencing and sequence motif mapping offers a multi-scale pathway for whole genome *de novo* assembly solving many of the ambiguities inherent in using short read assembly alone. Here we detail the BioNano genome assemblies of select prokaryotic and eukaryotic organisms demonstrating applicability of this technology to genome sizes spanning three-orders of magnitude and to a diversity of organisms. In the cases where a high-quality genome draft exists, the BioNano assembly is compared to the reference highlighting the utility of parallel multi-scale genome assembly. With these demonstrations, *de novo* assembly of the human genome using BioNano technology is within reach.

3650W

Defining best practice guidelines for the use of NGS applications in genome diagnostics. A national collaborative study of Dutch genome diagnostic laboratories (LOD). *N. van der Stoep[1], M.M. Weiss[2], Q. Waisfisz[2], C. Ruivenkamp[1], M. Nelen[3], J.D.H. Jongbloed[4], H. BRüggenwirth[5], M. van Slegtenhorst[5], R. Lekanne dit Deprez[6], O. Mook[6], A. Van den Wijngaard[7], M. Vogel[8], B. Van der Zwaag[8], M.M.A.M. Mannens[6], Dutch National Board for DNA-diagnostics (LOD) http://www.dnadiagnostiek.nl/).* 1) Dept Clinical Genetics, Leiden University Medical Center, Leiden, Netherlands; 2) Dept. Clinical Genetics, VU University Medical Center, Amsterdam, Netherlands; 3) Dept Genetics-Genome Diagnostics, St Radboud University Medical Center, Nijmegen, Netherlands; 4) Dept. Genome Diagnostics, University Medical Center Groningen, Groningen, Netherlands; 5) Dept. Clinical Genetics, Erasmus Medical Center, Rotterdam, Netherlands; 6) Dept. of Clinical Genetics, Academic Medical Center, Amsterdam, Netherlands; 7) Dept. of Clinical Genetics, University Medical Center Maastricht, Maastricht, Netherlands; 8) Dept. of Medical Genetics, University Medical Center Utrecht, Utrecht, Netherlands.

Next-generation sequencing (NGS) has developed over the past few years into a valuable tool for research applications in genomic research. There have been remarkable advances in NGS technologies which resulted in a considerable reduction in sequencing costs and a significant increase in throughput. Mean while in genome diagnostics the total number of genes tested has increased as well. Therefore the implementation of NGS applications in diagnostics is becoming an inevitable step. However the validation of this technique according to diagnostic standards must be established first and is quite a challenge for individual diagnostic laboratories. In the Netherlands all accredited diagnostic laboratories are joint and represented in a coordinating body named the `DutchNational Board for DNA-diagnostics' (LOD). The primary goal of this organization is to ensure the availability of a growing number of high quality diagnostic tests for hereditary and congenital disorders. This study was initiated by the LOD organization through recruiting representatives from all diagnostic laboratories in order to exchange their ongoing validation processes and experiences in NGS applications. Different applications of NGS are addressed including whole exome sequencing and analysis but also targeted analysis of specific gene sets. All stages of the process are evaluated including aspects as DNA preparation, various quality checks at different stages of the process, analysis pipelines, disclaimers and reports. From this study we define a first set of critical criteria and best practice guidelines for validating and implementing NGS applications in genome diagnostics according to accredited diagnostic standards. In addition all Dutch diagnostic laboratories are currently collaborating in a national NGS research study named “project CARDIO” which is initiated by the joint research group, `the Centre for Genome Diagnostics' that is also presenting data at this meeting. Results from this study will provide additional details on critical aspects to consider when using NGS in routine diagnostics.

3651F

Genome phantasmagoria: naming genes in alternative loci. *E. Bruford, R. Seal, M. Wright.* HGNC, EMBL-EBI, Hinxton, United Kingdom.

The HUGO Gene Nomenclature Committee (HGNC) has assigned unique approved gene symbols and names to over 33,000 human loci to date. Over 19,000 of these are protein coding genes, but we also name pseudogenes, non-coding RNAs, phenotypic loci and genomic features. Our website, genenames.org, is a searchable repository of human gene nomenclature and associated resources, including links to genomic, proteomic and phenotypic information, our curated gene family pages and the HGNC Comparison of Orthology Predictions (HCOP) tool. The HGNC assigns nomenclature to genes that are present on the current human genome, which is a combined reference assembly based on only a small number of anonymous donors. Now that the majority of protein coding genes on the reference assembly have been annotated and named, it has become clear that a single tiling path cannot adequately represent regions of complex human variation. The Genome Reference Consortium (GRC) aims to improve the human genome by correcting errors, closing sequence gaps and providing alternative assemblies for some of the hypervariable regions. The GRC generates patches, small regions of genomic sequence that are released independently from the major genome assembly releases, which can either be a fix patch correcting an error or a novel patch containing additional alternate loci. Fix patches are incorporated into the next full assembly release, while novel patches will be represented as alternate reference loci. The HGNC is currently working directly with the GRC to provide gene symbols for genes in novel patches when requested. We will provide some examples of nomenclature for alternate loci and will discuss the issues that we face as complex variation is increasingly represented.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3652W

Rapid Preparation of Targeted Resequencing Libraries from DNA Samples Using the Access Array™ System. *C. Friedlander, X. Wang, R. Ramakrishnan.* R&D, Fluidigm Incorp., South San Francisco, CA.

The dramatic drop in cost and decreased turnaround time associated with next-generation sequencing (NGS) platforms has created a demand for methods that permit rapid library preparation, targeting multiple genomic regions, in large cohorts. The Access Array System is a microfluidic integrated fluidic circuit (IFC) platform capable of simultaneously amplifying 48 samples with 48 primer sets per run. Each reaction contains a target-specific primer pair and a sample-specific barcode primer pair resulting in 2,304 unique amplicons in a single-step reaction and takes four hours to complete. Coverage and throughput can be further increased by a factor of ten by using multiplexed reactions generating 23,040 PCR reactions per run, thereby further reducing the cost. As little as 50ng of DNA sample is enough to generate products which can be used for sequencing on all commercialized next-generation sequencing platforms, including Roche 454, Illumina, and Ion Torrent sequencers. To demonstrate the performance of libraries generated on the Access Array System, we have carried out multiplexed amplifications of about 2,500 genomic regions that map into genes with relevant roles in cancer development. Both normal DNA from cell lines and human liver tissue were used to interrogate these regions. These samples were amplified for the target regions on the IFC, followed by the addition of sample specific barcodes in a subsequent reaction, in a two-step library preparation. On average, 80–98% of the NGS data produced across all platforms map to the human genome and, of these, over 95% map to the target regions. Over 80% of all assays produced reads within five-fold from the average number of total reads. The results obtained from our experiments demonstrate the ability of the Access Array System to generate high-throughput libraries for 454, Illumina, and Ion Torrent NGS platforms in just a few hours, decreasing hands-on time and overall cost per sample.

3653F

MitoExome: A Custom Whole Exome Kit for Mitochondrial Disease Research and Diagnosis that Captures All MitoCarta Genes and the Mitochondrial Genome. *X. Gai[1], E.A. Pierce[2], M. Consugar[2], M. Lvova[3], D.C. Wallace[3], E. LeProust[4], M.J. Falk[5].* 1) Center for Biomedical Informatics, Loyola University Stritch School of Medicine, Maywood, IL; 2) Ocular Genomics Institute and Berman-Gund Laboratory for the Study of Retinal Degenerations, Department of Ophthalmology, Massachusetts Eye and Ear Infirmary, Harvard Medical School, Boston, MA; 3) Center for Mitochondrial and Epigenomic Medicine, Department of Pathology, The Children's Hospital of Philadelphia, and University of Pennsylvania Perelman School of Medicine, Philadelphia, PA; 4) Agilent Research and Development, CA; 5) Divisions of Human Genetics, and Rehabilitation and Metabolic Disease, Department of Pediatrics, The Children's Hospital of Philadelphia, and University of Pennsylvania Perelman School of Medicine, Philadelphia, PA.

Discovering mitochondrial disease variants and genes requires simultaneous interrogation of both the mitochondrial genome and the nuclear genome, which encodes at least 1100 mitochondria-targeted proteins whose exact composition varies between tissues. Currently available whole exome capture kits do not capture the mitochondrial genome and provide insufficient capture for some MitoCarta-catalogued nuclear mitochondrial genes. METHODS: To improve upon this, we formulated a custom SureSelect library by blending RNA “baits” from three separate designs: (A) Agilent Technologies (Santa Clara, CA) SureSelectXT 50 Mb All Exon PLUS Targeted Enrichment Kit, (B) a panel of 16 nuclear-encoded mitochondrial genes that target sequences not included in the All-Exon design, and (C) sequences targeting the entire mitochondrial DNA genome. Designs B and C were created in eArray by standard 1X tiling across the target hg19 mitochondrial loci or 2X tiling across nuclear genes. The final custom formulations consisted of a 1:1 ratio of Design A/All-Exon to Design B baits. Design C/mitochondrial baits were subsequently blended in either at equal molar ratio or reduced concentrations by 10, 50, 100, 200, 500, or 1000 fold less than the nuclear baits. Patient sample capture libraries were 2 × 101 bp paired-end sequenced on an Illumina (SanDiego, CA) HiSeq 2000 Next-Generation Sequencing system using v3.0 SBS chemistry. RESULTS: As expected, the mitochondrial genome coverage varied depending on the mitochondrial:nuclear blend ratio, with 1:100 and 1:200 ratios providing the optimal coverage for both genomes. 1X, 10X, or 20X coverage was achieved for over 99%, 97%, and 95% of all targeted nuclear exons. 1X, 10X, 100X, and 1000X coverage was obtained on average for 100%, 99.99%, 99.9%, and 99% of the mitochondrial genome. Low-level heteroplasmy was sensitively detected. Analysis of nuclear-encoded mitochondrial transcripts is ongoing. CONCLUSIONS: This custom mito-plus whole exome capture kit allows simultaneous detection of both nuclear and mitochondrial DNA genomes, including sensitive detection of low-level heteroplasmy mtDNA variants that commonly cause mitochondrial disease. It represents an ideal experimental platform for both research and clinical-based diagnostic investigations of suspected mitochondrial diseases.

3654W

High Throughput Solution For Illumina Fragment Library Sample Preparations. *A. Jackson, S. Verrow, M. Blair.* Beckman Coulter, Inc., Brea, CA.

Beckman Coulter has released SPRIworks chemistry for low-throughput using the SPRI-TE Nucleic Acid Extractor and has now released chemistry for high-throughput (HT) automation using Biomek® liquid handlers for Illumina compatible Next Generation Sequencing (NGS) library sample preparation. The SPRIworks HT automation solution turns a tedious process that can require fragment size selection through agarose gel separation into a simple process. The HT solution can produce up to 96 libraries in a 96-well format with minimum hands on time. The automation friendly kit uses Solid Phase Reversible Immobilization (SPRI) paramagnetic beads to offer various fragment sizes selection (150–350 bp, 250–450 bp, 350–550 bp, and 150 bp up) with high accuracy as shown by bioanalyzer data. SPRIworks HT has also been utilized to produce libraries for additional applications such as RNA sequencing and Target Capture. Downstream processes, including PCR cleanup, qPCR setup, and normalization and pooling, have been automated to create a suite of methods. This suite of automation methods has a single user interface (UI) to allow for easy selection of options.

3655F

Comparison of two next-generation sequencing technologies on the genomes of a trio family. *E. Jorgenson[1,2], R. Kazma[3].* 1) Department of Neurology, Ernest Gallo Clinic and Research Center, University of California San Francisco, San Francisco, CA, USA; 2) Division of Research, Kaiser Permanente, Oakland, CA, USA; 3) Department of Epidemiology and Biostatistics and Institute for Human Genetics, University of California San Francisco, San Francisco, CA, USA.

Next-generation sequencing provides a new level of scrutiny on the human genome, and its rapidly decreasing cost is making the investigation of the effect of rare genomic variants on complex diseases and traits increasingly affordable. Yet, our knowledge of the relative quality of the various competing technologies is scarce, and only one comparison of a single individual has been published to date (Lam et al., Nature Biotech 2012;30:78–82). Here, for the first time, we compare the two most widely used sequencing technologies, Illumina and Complete Genomics using the genomes of a trio family. We sequenced the genomes of two parents and one male child from Costa Rica using two technology pipelines: Complete Genomics version 1.10.0.28 and Illumina version 1.3.0.0 (MARS). Both pipelines were also used to call single nucleotide variants and small insertion/deletions using the reference genome (NCBI build 37) for alignment. Additionally, the Complete Genomics pipeline called copy number variants and larger structural variants, and provided gene annotations. The average read depth was about 55× for Complete Genomics and 50× for Illumina. Furthermore, the three genomes were genotyped on two Illumina chips: HumanHap650y and HumanOmni 2.5, providing another level of comparison. The use of a trio family for this comparison allowed for crosschecking calls and better estimation than a single genome of the coverage, concordance, and false positives of both pipelines.

3656W

Single-Molecule, Electronic, Solid-State Sequencing of M13 DNA. *J.S. Oliver, B. Bready, D. Dederich, J. Freitas, H. Geiser, Y. He, D. Hevroni, M. Jouzi, H-Y. Lee, P. Mukhatira, M. Nadel, J. Sariadaridis, J. Thompson.* Nabsys Inc, 60 Clifford Street, Providence, RI.

Nabsys' positional sequencing technology is a single-molecule approach that utilizes solid-state nanodetectors. The platform locates, with sub-diffraction-limit resolution, the positions of oligonucleotide probes that have bound to long DNA fragments. The platform does not require amplification of samples. As a result, the method is capable of providing very long reads (in excess of 100 kb) with unbiased, quantitative information. Positional sequencing does not attempt to discriminate individual nucleotide bases passing through an electrical detector. Instead, the Nabsys approach hybridizes short oligonucleotide probes to very long DNA templates, passage of probe-bound templates through solid-state nanodetectors, and electronic detection of the locations of the hybridized probes. By combining information on the positions of probes, it is possible to create detailed genomic maps with sparse probe coverage, or true de novo sequences of large genomes with dense probe coverage. We demonstrate the utility of the method by showing resequencing and map data for M13 using single molecule detection of events in solid-state detectors.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3657F

Efficient Genotyping of Individuals using Overlapping Pool Sequencing and Imputation. *Z. Wang[1], F. Hormozdiari[1], W. Yang[1,2], E. Eskin[1].* 1) Computer Science, Univ California, Los Angeles, Los Angeles, CA; 2) Interdepartmental Program in Bioinformatics, Univ California, Los Angeles, Los Angeles, CA.

Next generation sequencing technologies are rapidly decreasing the cost of obtaining genetic information. The cost for utilizing one of these technologies consists of a sample preparation step and a sequencing step of the prepared sample. The dramatic increase in the efficiency of the sequencing technology makes the costs of the sequencing step negligible for small target regions. Thus the main remaining cost is the sample preparation step. Using overlapping sequencing pools where samples are mixed together into pools which are prepared and sequenced together has been shown to reduce the cost significantly for collecting information on genetic variants which only occur in a few of the samples. These methods utilize ideas from compressed sensing. In this paper, we extend this approach to utilize additional information from reference genetic variation datasets which provide the correlation structure between genetic variants. Utilizing this information, we can significantly increase the efficiency of overlapping pool sequencing.

3658W

Whole-genome sequencing analysis of SNPs and structural variants in DNA from blood vs. EBV-transformed lymphoblasts from the same subject. *X. Zhu[1], C. Laurent[2], M. Haney[1], AE. Urban[1,3], DF. Levinson[1].* 1) Department of Psychiatry and Behavioral Sciences, Stanford university, Palo alto, CA; 2) Department of Child and Adolescent Psychiatry, UPMC, Paris, France; 3) Department of Genetics, Stanford, Palo alto, CA.

Lymphoblastoid cell lines (LCLs) created by in vitro infection of human B-lymphocytes with the Epstein-Barr Virus have been widely used in genetic and genomic studies due to their wide accessibility in repositories. There are concerns about the effect of EBV-induced artifacts on sequencing studies, but the degree and nature of these changes is still unclear. Several direct comparisons of paired blood and LCL specimens based on SNP and CGH arrays suggested that there might be characteristic copy number changes in several chromosomal regions. But in a study using Sanger sequencing of exons of 401 candidate genes, 13 of 28 presumed de novo mutations (observed in LCL from a single proband but not in the parents) were not observed in DNA from the proband's blood, and thus were considered LCL artifacts (PMID: 20797689). We carried out a systematic comparison of SNPs and structural variants in two DNA specimens from the same individual (one extracted from LCL and one from peripheral blood mononuclear cells), using both whole-genome deep sequencing (Illumina HiSeq2000, 30x) and a SNP array (Illumina Omni1-Quad). The discordancy rate for ~700,000 SNP array genotypes (~0.01%) was less than the established error rate for the method. A much greater discordancy rate (~ 0.18%) was observed for the ~4,500,000 SNPs called with adequate confidence from sequencing data, likely due to the higher error rate of this method. Similar proportions of SNPs were called only in the LCL or the blood specimen, and there was no clustering of discordant calls in the genome. Detailed analysis of structural variants is in progress, along with experimental investigation of a subset of the discordant SNP and CNV calls. The very low proportion of discordant SNP array genotypes suggests that very few SNP artifacts are introduced by EBV transformation. However, such artifacts would be problematic for studies focusing on de novo mutations. Validation experiments will permit us to estimate the full extent of SNP and structural variation introduced by EBV transformation in these specimens. This information may be of value to investigators considering or undertaking sequencing-based studies with LCL specimens.

3659F

Genomic capture combined with long read and short read high-throughput sequencing for assembly of the MHC region. *R.J. Bloom[1], A.L. Collins[1], A.E. Byrnes[2], Q. Langdon[1], S. Happe[3], J. Barboza[3], O. Hardy[3], G. Yuan[4], S. Ranade[4], P. Mieczkowski[1], P.F. Sullivan[1,5].* 1) Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC; 2) Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC; 3) Agilent Technologies, Santa Clara, CA; 4) Pacific Biosciences, Menlo Park, CA; 5) Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, NC.

The major histocompatibility complex (MHC) region has emerged as important for many human diseases, including multiple autoimmune diseases and infectious disease. An extended region (chr6:25–34 Mb) has emerged as significantly and replicably associated in schizophrenia, but the complexities of this region have made narrowing of the region intractable. Given the large variability between different individuals, we designed a study to re-sequence and assemble this region in multiple individuals homozygous across the region, with the goal of creating a new reference sequence for each of the most common MHC haplotypes. As proof of principle, we have started with DNA from cell lines with known sequence across the MHC: PGF, SSTO, and DBB. We used a novel variation of Agilent's SureSelect technology to enrich for DNA from this region. We designed baits across the targeted region (25–34Mb on chromosome 6, including the 7 alternative builds in UCSC; including SSTO and DBB). This protocol generates large fragment sizes and allowed us to prepare DNA from the same capture for sequencing with Pacific Bioscience's SMRT technology and Illumina's HiSeq2000. Alignment of the PacBio reads to hg19 confirms that we have enriched for the region (~75% enrichment for each of the cell lines) and that we obtained average insert sizes of up to 1400bp. Mean target coverage for PacBio reads is approximately 10X. Illumina reads have been generated using 100bp paired end sequencing. We will use both sets of reads for hybrid assembly, which combines the advantages of long reads with the accuracy of short reads to allow us to assemble the region from each of the cell lines, independently of the reference sequence. These assemblies will be compared to the existing build from each cell line to verify the method. We therefore have a technique for capturing long fragments from a single Agilent SureSelect enrichment that allows both PacBio and Illumina sequencing from the same capture. These experiments will determine whether the combination of these techniques can allow individual level assembly of complex regions of the human genome, elucidating regions intractable to genetic study in the past.

3660W

Automated pipeline for whole exome/genome sequencing analysis on Mendelian diseases. *Y. Guo, K. Wang.* University of Southern California, Los Angeles, CA.

The development of high-throughput sequencing technologies has dramatically changed the landscape of human genetics research on Mendelian diseases. Currently, there are approximately 1,772 known Mendelian diseases whose genetic basis remain elusive, based on the Online Mendelian Inheritance in Man (OMIM) database. Identifying causal genes and variants for these diseases will greatly improve disease diagnosis and facilitate the development of therapeutic strategies. Compared to complex diseases with complex inheritance patterns and multiple causal genes with incomplete penetrance, Mendelian diseases may be easier to interrogate by sequencing a few cases and family members. Here we present a computational pipeline that performs a series of automated steps to help biologists and clinicians with limited bioinformatics skills identify candidate causal genes for Mendelian diseases. This pipeline perform QC assessments of raw sequencing reads, map single-end or paired-end reads to user-specified reference genome, remove duplicates where necessary, perform local realignments and base quality recalibration, generate SNPs and indels calls by multiple popular algorithms, then compile a set of consensus calls to improve calling accuracy while maintaining high sensitivity. Subsequently, we use a set of filtering and annotation procedures implemented in the ANNOVAR package, including variant function, gene function, allele frequency, conservation scores, functional prediction scores, family information, combined information on multiple unrelated cases, gene-gene networks, phenotype similarity to known Mendelian diseases, to reduce the search space to a small ranked list of candidate genes. We tested the pipeline on whole-exome and whole-genome sequencing data sets on several Mendelian diseases or Mendelian forms of complex diseases. We demonstrated the efficiency of the analysis pipeline to generate candidate genes with minimum human intervention, and discussed the practical challenges in finding causal genes for Mendelian diseases.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3661F

Computational Analysis of Exome-seq for Disease Gene Identification with eDIVA. *S. Ossowski, R. Rahman, O. Drechsel.* Genomic and Epigenomic Variation in Disease Group, Centre for Genomic Regulation (CRG) and UPF, Barcelona, Spain.

Novel or inherited genetic variations can lead to drastic phenotypes including rare and common diseases. Human exome analysis using next generation sequencing (Exome-seq) has recently been established as a key approach to identify genetic variations in protein coding genes. We have developed computational methods for the identification and functional analysis of causal and disease associated mutations in Exome-seq studies of rare and common, mendelian and complex diseases. To facilitate the routine application of Exome-seq for the identification of disease associated genetic variants and somatic mutations, diagnosis and improved personalized treatment we have implemented eDIVA. This pipeline performs read alignment, SNP, indel and CNV prediction, functional annotation of variants and estimation of damaging potential for coding SNPs. To improve variant annotation, data accessibility and comparison between studies we integrated a BioMart based database and disease variant annotation tool. eDIVA-DB incorporates predicted variants together with relevant functional data as well as known variants in the population. It supports rich annotation of variants by adding OMICs information from e.g. dbSNP, 1000genomes, EVS, KEGG, BioGRID and OMIM. Furthermore it supports gene and pathway based mutation enrichment studies for the analysis of complex diseases. We further integrated a rank product based algorithm to prioritize candidate SNPs and indels using all integrated information. The prioritization algorithm has been validated on Exome-seq data from a study on Familial Hyperkalemic Hypertension (FHHt) by successfully identifying the causal variant KLHL3, previously shown to be causal for FHHt in this family. eDIVA allows for automated analysis of large-scale Exome-seq studies and we currently analyze various types of diseases including cardio-vascular (early onset varicose veins), neurological (congenital myopathies and ataxia) and rheumatologic diseases (fibromyalgia) as well as cancer (chronic lymphocytic leukemia).

3662W

The Biologist's Tool for Finding a Needle in a Haystack: CLC bio's Platform to Identify Interesting de novo and Accumulative Variants in the Whole Genome Sequence of Family Trios. *N. Thomson[1], A. Joecker[2], C. Boysen[1], M. Matvienko[1], U. Appelt[2], A. Fejes[2], L. Kahns[2], S. Mønsted[2], J. Grydholt[2], B. Knudsen[2], M. Bungaard[2], J. Buur Sinding[2], H. Handberg[2], A. Hiene[2], M. Nygaard Ravn[2], A. Joecker[2], M. Værum[2], R. Forsberg[2].* 1) CLC bio, Cambridge, MA; 2) CLC bio, Finlandsgade 10–12, Katrinebjerg, 8200 Aarhus N, Denmark.

The debate over the value of Whole Genome Sequencing (WGS) vs Whole Exome Sequencing (WES) revolves around the expense of data generation, the challenge of analysis, and the ability to interpret results. However there are ever increasing reasons to suggest that phenotypic variation may be the result of genotypic variation outside of the exome. Over 90% of variants reported in dbSNP lie outside of the annotated coding regions. Additionally, while less than 2% of the human genome is translated into protein, much larger percentages are conserved, transcribed and assigned functionality. Advances in DNA Sequencing technology eliminate the expense of data generation as one obstacle in pursuing WGS. Mapping sequencing reads to the human genome, a task that previously consumed weeks and huge compute resources, can now be done in a reasonable time on a single compute node with CLC bio's mapping algorithms. The integration of advanced algorithms, access to public resources, and a friendly user interface within the Genomics Workbench create an environment in which biologists can fully leverage WGS data. In this work we will analyze a European family trio from the 1000 genome project, using CLC bio's Genomics Workbench. We will show all steps, from mapping sequencing reads to the human genome, to the identification of potential disease relevant variants. Specifically we will emphasize the method for discovering de novo (only present in the child) and accumulated (heterozygote in the parents, homozygous in the child) variants. This process could reveal candidate variants for disease, when a child is ill, for instance, and the parents are not. Furthermore, we will show how a prioritization of these variants, through annotation of potential functional consequences, will facilitate the identification of variants of interest.

3663F

Ultra-fast clinical sequencing and annotation of human genomes on the Illumina HiSeq 2500 platform. *S. Humphray, J. Weir, Z. Kingsbury, T. James, R. Grocock, P. Saffrey, E. Margulies, K. Hall, D. Bentley, G. Smith, J. Betley, C. Raczy.* Illumina UK, Saffron Walden, Essex, United Kingdom.

We demonstrate an easy to use whole genome sequencing workflow from DNA to annotated variants. In many clinical settings, such as Neonatal, Paediatric and Medical Intensive Care Units, time to answer is critical for patient care. We describe the system advances that have been made in the areas of library preparation, sequencer performance and alignment/annotation software. We illustrate the application of this rapid workflow, and the ability to use low amounts of DNA including FFPE samples, to a real-world clinical environment with a time to answer of less than 48 hours. Whole genome sequencing (WGS) is being increasingly adopted in both research and clinical settings. We are able to use low input amounts of DNA (100ng). Optionally using a PCR-free sample preparation protocol improves uniformity of coverage. The improvements also allow the user to go from genomic DNA to clustering in less than 4 hours. The HiSeq 2500 enables very rapid production of whole genome data with the time from introducing the prepared library to completed paired end 2×100 cycle runs being reduced to less than 27 hours. A new suite of software enables very fast analysis and annotation of genomes for clinical interpretation and reporting. Features include a new fast sequence aligner that, combined with variant calling and generation of genomic VCF (gVCF) files, provides information of quality and depth of coverage of all genomic positions including variant calls. Annotation with Variant Effect Predictor (Ensembl) output and data from 1K Genomes, Encode and other public sources produces a compressed annotated gVCF file (of approximately 1Gbyte) that is highly portable in local and remote (e.g. Cloud) environments for maximum accessibility in clinical and research environments.

3664W

Managing Computing Resources in Large Sequencing Studies: Strategies and Lessons from Sequencing 2,120 Sardinian Genomes. *A. Kwong[1], C. Sidore[1,2,3], S. Sanna[3], H.M. Kang[1], G. Jun[1], M. Trost[1], P. Anderson[1], T. Gliedt[1], R. Cusano[3], M. Pitzalis[3], M. Zoledziewska[3], A. Maschio[1,3], F. Busonero[3], M. Lobina[3], M. Balloi[3], B. Tarrier[4], C. Brennan[4], C. Jones[5], F. Cucca[2,3], G. Abecasis[1], SardiNIA project.* 1) Center for Statistical Genetics, Department of Biostatistics, University of Michigan, Ann Arbor, MI; 2) Dipartimento di Scienze Biomediche, Università di Sassari, Sassari, Italy; 3) Istituto di Ricerca Genetica e Biomedica, CNR, Monserrato, Cagliari, Italy; 4) DNA Sequencing Core, University of Michigan, Ann Arbor, MI, USA; 5) CRS4, Advanced Genomic Computing Technology, Pula, Italy.

Though the cost of sequencing has decreased drastically in the last few years, the computing resources required for processing the data, which includes read mapping, alignment refinement, variant calling, and genotype refinement, remain substantial. Here, we briefly summarize our current results, and describe our strategy for managing, scheduling, and optimizing the computing resources required to conveniently process these data and update results as new samples are sequenced. We have low-pass sequenced the genomes of 2,120 Sardinian individuals (at ~3.5x coverage) to study the genetics of lipid levels and other aging-related traits. From our current dataset of >50 Tb of raw sequence data, an average of 3.4 million SNP variants were called per individual; ~17 million SNP variants were called across all individuals. We discovered 151k coding variants in our samples, and rediscovered 88% of the European variants found in 1000 Genomes with an overall allele frequency > 0.01. Our strategy starts with the design of physical computing resources, with a mixture of large networked file systems that serve as the central repository for project data, and small local file systems for managing intermediate results in each compute node. Jobs are scheduled to maximize the use of local storage for intermediate results and to minimize costly swapping of memory to disk and transfer to network file systems. Samples are aligned in parallel, with one to three samples assigned to each physical computing node. Initial read mapping and alignment refinement requires ~50 CPU-hours per sample (corresponding to >130 CPU-months for all 2,000 samples). After alignment, we performed variant calling and genotype refinement across all 2000 samples together to increase the power to detect rare variants, using parallel jobs that each handle a small chromosomal region (typically 5M bases). Variant calling, a high I/O job, typically requires ~6 CPU-hours per region (corresponding to ~5 CPU-months for all regions), and is run on the cluster with the fastest data access, while genotype refinement, a CPU-intensive job, typically requires ~2 CPU-days per 10,000 SNPs (corresponding to ~100 CPU-months for all regions), and is run on the cluster with the most computational power. The strategies described are instrumental in our large sequencing project, and serve as a useful guideline for sequencing projects of all sizes.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3665F

Clinical implementation of targeted next generation sequencing in a mid-sized diagnostic laboratory. *H. Racher[1], L. Dimnik[1], P. Gordon[2], M.I. Innes[1], F. Bernier[1], J.S. Parboosingh[1].* 1) Medical Genetics, Alberta Children's Hospital/University of Calgary, Calgary, AB, Canada; 2) Genomics & Bioinformatics Facility, University of Calgary, Calgary, AB, Canada.

Sequencing of the entire human exome by next generation sequencing (NGS) is a powerful tool for disease variant detection. This is a useful starting point for translation into the clinical setting as the coding sequence and flanking splice sites harbour ~85% of all human disease-causing mutations (Choi et al., 2009). However, translation of whole exome NGS into the clinic is problematic due to: (1) interpretation and handling of large data sets, (2) incidental findings, (3) inconsistent and/or incomplete coverage, and (4) system accuracy, sensitivity and precision. Using the Life Technologies SOLiD 5500 platform to perform whole exome NGS we determined system performance through orthogonal validation on six exomes with Agilent SNP microarrays. Accuracy was determined to be 98.6%, while sensitivity and specificity were 99.1% and 99.7%, respectively. We explored the variability in sequence coverage between two exon capture kits, Agilent SureSelect v3 kit and Life Technologies TargetSeq kit. The kits, by design, cover 99% ad 98.4% of the 32.9 Mbps of total refGene human coding sequence in addition to varying amounts of flanking intronic/non-coding sequence. ≥20X coverage of coding regions was observed in 85.6% of the SureSelect reads and 91% of the TargetSeq reads. In order to address some of the other challenges, we explored customizing enrichment kits by restricting analysis of the exome data to targets associated with known diseases. We used 31 genes associated with hypertrophic cardiomyopathy (HCM), a common disorder with a prevalence of 1/500, to evaluate the utility of this approach. A coverage level of ≥20X was observed for 99.6% of the regions using SureSelect and 99.8% with TargetSeq. The regions with low coverage varied, however, ~30% of the low coverage reads occurred within exon 1 sequence, a commonly underrepresented region due to GC rich sequence. Our next step in this analysis will be to run a customized kit, specific to the HCM panel of genes, using the SOLiD 5500 platform and compare this data to our whole exome NGS data. Our ultimate goal will be to move to benchtop sequencers, which will improve our turn-around-time by requiring fewer samples per run.

3666W

High-fidelity sequencing for detection of low-frequency single nucleotide variants. *K.M. Squire, Z. Chen, S.F. Nelson.* Human Genetics, UCLA, Los Angeles, CA.

It is often useful to identify very low frequency variants present in a DNA sample, for example, when attempting to identify variants from an unlabeled pool of many unidentifiable individuals, or when attempting to identify mutations that might confer resistance to a subset of cancer cells. Here we propose a method to identify single nucleotide variants (SNVs) present at a very low frequencies within a sample. Previously sequenced ecoli DNA was prepared according to standard Illumina library preparation protocols, with the exception that the DNA was sheared and size selected on a gel so that the insert size was 80±10bp. The DNA was sequenced using 100+100 paired end sequencing on a HiSeq 2000. Because the average insert size was less than 100bp, in almost all cases the same fragment was read twice. We aligned both ends of each read pair against one another, trimmed adapter sequence, and used the overlapping portion of the reads to form a single high-fidelity output read. Bases which were not identical in both reads were covered to Ns. These high-fidelity reads were aligned against a standard ecoli reference, and simple pileup and counting was used to identify variants. The list of variants was compared against previously identified variants from the same DNA pool. Newly identified variants were either low-level variants or sequencing errors, and allowed us to put an upper limit on the sequence error rate of 0.01%, compared with a sequence error rate of 0.1% for the original sequence. This suggests that the proposed method can be used to identify single nucleotide variants present above this rate.

3667F

Species identification by polymorphisms of mitochondrial 12S rRNA and 16S rRNA genes. *R. Li[1], L. Yang[2].* 1) Human Genetics, Cincinnati Children's Hosp, Cincinnati, OH; 2) Dept. of Molecullar & Cellular Physiology, College of Medicine, University of Cincinnati. Cincinnati, OH.

There are two genetic systems in cells of all animals - the nuclear and mitochondrial genomes. The unique properties of mitochondrial DNA (mtDNA), including maternal inheritance, multiple copy numbers, high mutation rate and polymorphism, have made mtDNA a good choice to study biological evolution and species identification. Interspecific and intra-specific variations (polymorphism) of mtDNA have been found by bioinformatic analyses of the mitochondrial genomic sequences of thirteen of species, including Xenopus laevis, Plecoglossus altivelis, Mus musculus, Bos Taurus, Pan Paniscus, Pan troglodytes,Gorilla gorilla, Papio hamadryas, Lemur catta, Cebus albifrons, Tarsius bancanus, Hylobates lar, and Homo sapiens. Some conserved regions have been identified in mitochondrial 12S rRNA and 16S rRNA genes from these thirteen species. Therefore, the universal primers designed in the conserved regions of 12S rRNA and 16S rRNA genes will be capable of amplification mtDNA fragments of 12S rRNA and 16S rRNA genes most animal mtDNA. Using these primers, spanning 1068 to 1076 (forward) and 1341 to 1360 (reverse) for mitochondrial 12S rRNA gene, from 2582 to 2591(forward) and 3044 to 3063(reverse) for the mitochondrial 16S rRNA gene (NCBI: NC_012920.1), to amply DNA samples from culture cells of human and mouse as well as commercial tissues of shrimp, pork, beef and chicken, we successfully obtained ~300 bp (298 bp in human) and ~480 bp (482 bp in human) of PCR products. Sequencing these PCR products, then blast searches at (http://blast.ncbi.nlm.nih.gov/Blast.cgi?-PROGRAM=blastn&BLAST_PROGRAMS=megaBlast&PAGE_TYPE=BlastSearch&SHOW_DEFAULTS=on&LINK_LOC=blasthome#), identified their species origins and similar sequences in NCBI database by their nucleotide variations in these two mitochondrial rRNA genes. This molecular technique combined with bioinformatics provides a reliable method for identification of the maternal origin of animal tissues. It can be used for the study of phylogenetic relationship, species evolution, forensic biology, food industry, protection of extinct animals and classification of novel species.

3668W

Rapid and Efficient Methods for Preparing Globin- and rRNA-Depleted Directional RNA-Seq Libraries. *C. Kinross, J. Hitchen, N. Caruccio, R. Sooknanan.* Epicentre, an Illumina Company, Madison, WI.

Massively parallel sequencing of cDNA libraries (RNA-Seq) is the preferred method for profiling and analysis of novel transcripts, isoforms, alternative splice sites, rare transcripts and cSNPs. However, a persistent challenge is that ribosomal RNA (rRNA) represent >95% of total RNA content and its removal is essential for increased sensitivity and coverage of transcripts when preparing next-generation sequencing (NGS) libraries. Additionally, for whole blood RNA samples, reticulocyte globin transcripts can constitute up to 76% of total mRNA and, unless removed along with the ribosomal RNA transcripts, the detection of low-abundance mRNA transcripts is further compromised. Here, we present an efficient “single-pass” method (GlobinZero™ Gold Kit) for the concomitant removal of rRNA and globin mRNA from intact and fragmented blood total RNA samples (human, mouse and rat), and an improved, more user-friendly version of the ScriptSeq™ RNA-Seq library preparation method. Globin-Zero-treated samples contain <2% of rRNA and globin sequences, while maintaining representation of coding and noncoding transcripts independent of polyadenylation. This reduction in rRNA and globin sequence reads improves sequence depth and coverage of mRNA, and increases the percentage of uniquely mapped reads.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

† Trainee Award Finalist

3669F

Sequencing performance of FFPE DNA in the SureSelect^{XT2} Target Enrichment System. *J. Barboza[1], M. Ramirez[2], A. Giuffre[1], J. Ong[1], H. Ravi[1], M. Guadalupe[1], S. Joshi[1], M. Visitacion[2], C. Pabón-Peña[2], S. Hunt[2], B. Novak[2], D. Roberts[2], S. Happe[1], E. LeProust[2].* 1) Genomics Division, Agilent Technologies, Cedar Creek, TX; 2) Genomics Division, Agilent Technologies, Santa Clara, CA.

Massively parallel sequencing technologies have enabled scientists to discover rare mutations, structural variants, and novel transcripts at an unprecedented rate. Agilent Technologies has developed the SureSelect platform to enable fast, inexpensive and accurate genome analysis. Recently developed, the SureSelect^{XT2} kit has a more streamlined sample preparation protocol as well as the capacity for multiplexed sequencing, greatly reducing sequencing cost per sample. To extend the capabilities of the SureSelect^{XT2} system we probed the limits of gDNA input amounts. Our data shows that reducing the input amount of starting material from 1µg to 0.5µg and 0.25µg did not affect the sequencing quality of the libraries produced. We expanded our studies by analyzing a set of human tumor DNA samples extracted from paraffin-embedded (FFPE) blocks. In agreement with our initial experiments, FFPE samples with varying starting amounts produced sequencing results that were comparable to control samples. In addition, we tested two commercially available FFPE DNA extraction kits. While the two preparations produced DNA that differed in total yield and quality, sequencing results were comparable between FFPE DNA prepared from both kits. Our studies highlight the robustness of the SureSelect^{XT2} Target Enrichment System, and its utility in analyzing DNA from valuable archived material that may be in low abundance and susceptible to damaging modifications.

3670W

Torrent Variant Caller: Enabling Next Levelel of Genomic Analysis. *D. Brinza, Z. Zhang, E. Tsung, A. Joyner, C. Scafe, G. Del-Mistro, F. Hyland, E. Beasley, S. Utiramerur.* Life Technologies, Foster City, CA.

Ion Torrent PGM and Proton sequencing platforms have a very quick turnaround time and produce longer high accuracy reads. These create a computational challenge of making an extremely fast and accurate variant caller that can take advantage of unique error model, and enable detection of longer indels. Here we introduce a Torrent Variant Caller software suite designed to address these challenges. One approach to variant calling is first to align reads to reference genome, then call variants based on pile-up formed by alignments. With longer reads, the challenge is to extend alignments beyond long indels. We address this by first mapping reads to the reference and then de novo assembling non-aligned read portions in regions with evidence for novel sequences. We then quickly traverse all alignments and predict candidate variants based on a very fast heuristic algorithm. The framework models around flow-space alignments of each candidate variant and accurately estimates the length, position, and the score of the variants. The method fast and accurately predicts SNPs, long and short INDELs.

3671F

A Next-Gen Sequencing Software Workflow for Cancer Genomics on a Desktop Computer. *M. Keyser[1], K. Maxfield[1], T. Schwei[1], T. Durfee[1], A. Pollack-Berti[1], D. Nash[1], J. Stieren[1], S. Baldwin[1], R. Nelson[1], K. Dullea[1], J. Schroeder[1], P. Pinnkas[1], G. Plunkett III[1,2], F. Blattner[1,2,3].* 1) DNASTAR, Inc., Madison, WI; 2) University of Wisconsin, Department of Genetics, Madison, WI; 3) Scarab Genomics LLC, Madison, WI.

DNASTAR offers an integrated suite of software for assembling and analyzing sequence data from all major next-generation sequencing platforms supporting key workflows on a desktop computer, including. The cancer genomics workflow that integrates some of the most powerful functionality in the software includes assembling and analyzing multiple samples using one reference template; probabilistic identification of SNPs, small indels and genotype calls with known variants correlated to their dbSNP and COSMIC IDs and GERP reference data; review and filtering of SNPs from multiple samples within a single project; identification of structural variations; and, for large multi-sample projects with hundreds of individual data sets, tools for SNP quantitation, filtering, set comparison, clustering and indication of the gene disruption impact from called SNPs. Interactive views within the software facilitate fast, comprehensive analysis, helping scientists move quickly from raw next-gen sequencing data to genetic and genomic impact, including gene ontology. By using innovative algorithms within the software, scientists can have all of the assembly and analysis capabilities available to them on their desktop computer, supporting large data sets generated by any or all of the next-gen sequencing instruments and platforms.

3672W

Strategy for identification, prediction, and prioritization of non-coding variants of uncertain significance in heritable breast cancer. *P.K. Rogan[1,3], E.J. Mucaki[1], A. Stuart[2], N. Bryans[3], E. Dovigi[1], B. Shirley[3], J.H. Knoll[2], P. Ainsworth[2].* 1) Biochemistry; 2) Pathology; 3) Computer Science, Western University, London ON Canada.

High-throughput sequencing (HTS) of both healthy and disease singletons yields many novel and low frequency variants of uncertain significance (VUS), most of which are likely to be benign. Current approaches to predict pathogeneticity generally do not focus on non-coding variants even though they contribute to Mendelian disorders and contain many VUS. Using cleavable solution microarrays, we are capturing and enriching for variants in a set of genes known to harbor mutations that increase breast cancer risk. Oligo baits covering *ATM, BRCA1, BRCA2, CDH1, CHEK2, PALB2* and *TP53* were synthesized for solution hybridization with a custom cleavable microarray spanning the complete coding and intergenic regions 10 kb upstream and downstream of each gene. Non-exonic sequences are densely populated with repetitive sequences that can affect short read assembly. A novel probe design method was used to capture both repeat-free and divergent repeat sequences that are effectively single copy. Sequencing of multiple samples captured 86.1% of bait sequences with minimal off-target coverage and an error rate of 2.5×10^{-5}/bp. Information theory-based sequence analysis was used to prioritize non-coding variants which occurred within sequence elements recognized by proteins or protein complexes. The novel VUS identified are being investigated for effects on transcription factor-binding site (TFBS) and mRNA splicing mutations. We have developed and apply bipartite information theory based models for exon recognition which predict the relative abundance of natural, cryptic, and mutant splice isoforms resulting from predicted mutations using the combined donor and acceptor site strengths of each mRNA species. We have applied a similar approach to detect mutations in the promoters of *BRCA1* and *BRCA2* that alter strengths of TFBS. Information weight matrices were automatically computed by entropy minimization of c-Jun, FOSL2, GRP20, HSF1, IRF4, NF-κB, PU.1, RAD21, RXRα, TCF12, and YY1 TFBS from the global set of ENCODE ChiP-seq regions embedded within DNAaseI hypersensitive domains. These models are then used to evaluate novel variants discovered by sequence analysis of breast cancer patients for alteration the TFBS binding strengths. This strategy more comprehensively covers non-coding regions in breast cancer genes, and introduces a unified framework for systematic interpretation of VUS that may affect expression.

3673F

Genomic approach for environmental stress assessment in coral, *Scleronephthya gracillimum*. *S. Woo[1], S. J. Hwang[2], S. Yum[1], J. I. Song[2].* 1) South Sea Research Department, Korea Ocean Research and Development Institute, Geoje, Gyungsangnam-do, South Korea; 2) Division of EcoScience, Ewha Womans University, 120–750 Republic of Korea.

Coral ecosystem is the most magnificent and diverse marine ecosystems on Earth. They have had a crucial role in shaping the ecosystems that have dominated tropical oceans over the past 200 million years. Coral ecosystem is the home for fish, millions of invertebrates, algae, and microbes. However, environmental stresses such as climate change, ocean acidification and anthropogenic effects have been recognized as the major destructive impacts on coral ecosystems. Motivated by the precarious state of coral ecosystem, there is currently a keen interest in coral transcriptomics. By investigating changes in coral gene expression that are triggered by particular environmental stressors, we can begin to characterize coral stress responses at the molecular level, which should lead to the development of more powerful diagnostic tools for evaluating the health of corals in the field. Toward this end, we performed NGS (next generation sequencing) of the soft coral, Scleronephthya gracillimum, a geographically spread through the Kuroshio current species and developed oligo microarray using approximately 34,000 clones. Total 33,401 reads were input to analyze KOG categorization and 32.42% of reads were successfully categorized, 2.14%, uncharacterized and 67.58% were not hit in the known information using BLASTx algorithm. For the microarray experiment, We collected specimens from 2 sites in Jeju island, Republic of Korea which are geographically separated and isolated RNA from colonies exposed to heat stress (26°C, 28°C and 30°C) or control condition (24°C) at 6h, 12h, 24h and 48h post-exposure, respectively. Among the genes showing significant changes (2-fold) we selected 18,523 genes which were shared in heat stress groups and investigated their expressions. The clusters of genes showed the similar pattern of transcriptional changes such as upregulation or down regulation between 26 and 28°C groups but over 35% of genes revealed totally opposite pattern of expression in 30°C group. The identified differentially expressed genes have a potential to identify specific stressors in environmental changes and could act as molecular biomarkers for biological responses against external environmental stimuli. And this transcriptomic approaches can be used to develop a diagnostic tool for coral health assessment.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3674W

RNA-seq uncovers the influence of structural variants on transcriptome diversity. *E. Ait Yahya Graison, A. Reymond.* CIG, Univ Lausanne, Lausanne, Switzerland.

Structural variants (SV) of DNA segments have been identified as a major source of genetic diversity, but a comprehensive understanding of the phenotypic effect of these variants is only beginning to emerge. Our group and others established extensive maps of SV in mouse strains, covering ~11 percent of their autosomal genome. In this study, we used the RNA-seq technology to unravel the effects of SV on expression at the nucleotide resolution and simultaneously on alternative splicing. We generated extensive RNA-Seq data from liver and brain (>450 millions single-end reads) from brain and liver of three mouse inbred strains (129S2, DBA/2J and C57BL/6J) using the Illumina Genome Analyzer IIx. We mapped reads on the mouse genome using TopHat, assembled and estimated the abundance of the different isoforms using Cufflinks. To evaluate expression changes for transcripts mapping within and outside SVs, we took advantage of the sequencing of the genome of 17 key mouse strains with high-throughput technologies. We intersected the coordinates of the identified SVs with the transcriptome data we generated to define different categories of expressed features according to their degree of overlap with a rearrangement. Generally, for both brain and liver, gene expression seems impacted for all genes, which are at least partially overlapping with a SV. At transcript level, transcripts lying in large SVs show constant overexpression compared to those outside SV. This is consistent with the fact that larger SVs encompassing whole coding regions of genes in our dataset are mainly copy number gains. We also assessed the impact of SV on alternative splicing by evaluating for each SV category the number of isoforms per gene we could detect in brain and liver for both 129S2 and DBA/2J strains. We observed that genes including exons which contain small SV and genes affected by SV in their upstream or downstream regions display significantly more isoforms per gene than genes not affected by SV. To conclude, our findings suggest that SV affecting the whole gene region have a more direct impact on their expression, whereas SV occurring in regulatory regions (e.g. inside exons or overlapping non-coding regions upstream/downstream of genes) are more likely to influence alternative splicing leading to higher number of isoforms of these genes. This study provides a unique opportunity to extensively gauge the influence of SV on the transcriptome complexity and regulation.

3675F

Detection of Local Signals in Genomics. *D. Siegmund[1], B. Yakir[2], N. Zhang[3].* 1) Dept Statistics, Stanford Univ, Stanford, CA; 2) Dept Statistics, Hebrew University of Jerusalem; 3) Department of Statistics, University of Pennsylvania.

A large class of problems of genomic analysis involve detection of local genomic signals. Examples are (i) linkage and/or association analysis, (ii) detection of intervals of copy number variation based on SNP arrays, (iii) detection of DNA structural variations from paired end reads, (iv) detection of protein-DNA interactions by ChIP-Seq. Theoretical approaches to some of these problems based on change-point methods are described, with particular attention to the problems of multiple comparisons. The theoretical results are illustrated on simulated and on real data.

3676W

Gene Expression Profiling of Prokaryotic samples using LIQA WT kit. *N. Guha.* Genomics, Agilent Technologies, Bangalore, Karnataka, India.

Agilent's Low Input Quick Amp Labeling WT Kit provides a reliable method to amplify and label whole transcripts for the robust generation of cRNA upstream of microarray hybridization. We describe a method to perform microarray analysis of two key prokaryotic organisms, Mycobacterium tuberculosis (MTB), a high GC-rich organism and Escherichia coli (E.coli) using the LIQA WT kit. This kit is also used for analysis of Agilent human, mouse and rat SurePrint G3 Exon Microrrays, catalog and custom arrays. The study demonstrates that the LIQA WT kit can be successfully used to label these prokaryotic transcripts and yield high quality microarray data. The data shows strong correlation (R2>0.95–0.99) between the replicates of the microarrays for both 1- and 2-color. Excellent sensitivity and signal linearity (>5 logs) were demonstrated using spike-in controls. Differential gene expression analysis of E. coli whole transcripts upon treatment with paraquat, a potent inducer of oxidative stress, revealed several key genes to be significantly regulated. The key genes included superoxide dismutase and fumarase C that are known to be strongly induced by paraquat. Overall the data demonstrates that using the LIQA WT Labeling kit along with Agilent's catalog E. coli and custom-designed MTB arrays one can generate high quality biologically relevant expression data.

3677F

In silico identification of microRNA-mRNA target pairs as potential biomarkers in Prostate Cancer. *J. Billaud, C. Tavano, D. Toburen.* Ingenuity Systems, Redwood City, CA.

This study will focus on identifying microRNA-mRNA target pairs as potential biomarkers for Prostate Cancer using in silico techniques. Prostate Cancer (PrCa) progresses in stages; from early stage primary tumors developing in the prostate to lethal metastatic forms in the bones. Two screening methods are used for the detection of tumors, namely their localization, and the presence of metastasis. However, conflicting studies suggest that in particular, PSA testing displays low specificity and inability to distinguish latent from aggressive cancer. In fact, a substantial number of men are over-diagnosed and over-treated for prostate cancer mainly due to the poor performance of the PSA test. Therefore, better biomarkers of diagnostic, predictive, prognostic values are urgently needed. Our in silico research combines the results of two studies that looked for i) a gene signature (158 mRNAs) that map to clinical conditions and ii) the potential use of microRNAs (miRNAs) as diagnostic and/or prognostic markers for prostate carcinoma. Studies have highlighted the role of miRNAs in normal and pathological situations, and as such they add an extra level of complexity to the regulation of gene expression. Indeed changes in their expression have been identified in inflammatory diseases, viral infection as well as human cancer. The miRNAs are a class of non-coding, short RNAs (an average of 22 nucleotides) that function as post-transcriptional regulators by binding to complementary sequences on target mRNAs and resulting either in translation repression or degradation of mRNA and gene silencing. A combination of capabilities within Ingenuity Pathway Analysis (IPA) will help dissect the role of these miRNAs in PrCa by highlighting which specific mRNAs are targeted, their involvement in important canonical pathways and networks as well as in biological processes that would explain tumor progression and/or metastasis. This in silico research will finally suggest hypothetical biomarkers that combine units of microRNA-mRNA pairs as potential biomarkers of disease progression in PrCa. In conclusion, it will be shown that IPA is an important tool for exploration of the molecular signature of cancer, for understanding consequences of miRNA expression changes in cancer by combining different aspects of biology, and exploring the clinical aspects related to biomarker identification.

3678W

Active Learning for Phenotype Mapping. *C. Hsu[1], J.L. Ambite[1], Y. Arens[1], L. Lange[2], S. Sharma[1], S. Voinea[1].* 1) USC/Information Sciences Inst, Marina del Rey, CA; 2) Department of Genetics, University of North Carolina, Chapel Hill, NC.

Many genetic epidemiologic studies of cardiovascular disease have multiple variables related to any given phenotype, resulting from different definitions and multiple measurements or subsets of data. A researcher searching such databases for the availability of phenotype and genotype combinations is confronted with a veritable mountain of variables to sift through. This often requires visiting multiple websites to gain additional information about variables that are listed on databases, and examination of data distributions to assess similarities across cohorts. While the naming strategy for genetic variants is largely standardized across studies (e.g. “rs” numbers for single nucleotide polymorphisms or SNPs), this is often not the case for phenotype variables. For a given study, there are often numerous versions of phenotypic variables. Manually mapping and harmonizing these phenotypes is a time-consuming process that may still miss the most appropriate variables. Previously, we have developed a supervised learning algorithm that learns to determine whether a pair of phenotypes is in the same class. Though this algorithm accomplished satisfying F-scores, the need to manually label training examples becomes a bottleneck to improve its coverage. Herein we present a novel active learning solution to solve this challenging phenotype-mapping problem. Active learning queries users for labels of unlabeled phenotypes that may improve the mapping the most and therefore will reduce the need of labeling efforts. Active learning will make phenotype mapping more efficient and improve its accuracy, along with intuitive phenotype query tools, would provide a major resource for researchers utilizing these databases. Our active learning method explores the connectivity of related phenotype variables to infer a set of sure matches and pairs of phenotypes variables that confuse the program the most, and query users for labels for these variables. We tested our method to a set of 1700 annotated variables from 7 large cohort studies. We show that by applying our active learning method, we can improve the F-scores of mapping from 61% to 66% by adding forty more annotated variables, a huge improvement compared to annotating variables randomly. Remarkably, the set of sure matches are completely correct in our experiments.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3679F

FAVR (Filtering and Annotation of Variants that are Rare): methods to facilitate the analysis of rare germline genetic variants from massively parallel sequencing datasets. *D.J. Park[1], T. Nguyen-Dumont[1], F. Odefrey[1], A. Lonie[2], M.C. Southey[1], B.J. Pope[2].* 1) Genetic Epidemiology Laboratory, Department of Pathology, The University of Melbourne, Melbourne, Victoria, Australia; 2) Victorian Life Sciences Computation Initiative, The University of Melbourne, Melbourne, Victoria, Australia.

Massively parallel sequencing (MPS) is a powerful tool to facilitate research on the genetic basis for observed phenotypes, including predisposition to and progression of human diseases. However, interpretation of MPS data is challenging. A typical single whole -human exome analysis outputs thousands of genetic variant signals. Only a fraction of these are real, the remainder being artefacts from the processes of sequencing and mapping to the reference genome. Many of these are not readily distinguishable from true variants on the basis of `quality metrics'. Of the authentic variants, only a very small proportion might be relevant to a given disease. Here, we present FAVR, a suite of new methods designed to assist the shortlisting of genetic variants under a rare variant-disease model. The methods are designed to work with commonly used MPS analysis pipelines. The FAVR methods use signatures in comparator sequence alignment files to facilitate the filtering of mapping artefacts and common genetic variants, and annotation of genetic variants based on evidence of co-occurrence in individuals. As relevant, FAVR methods can also be used to filter out artefacts derived from imbalanced paired-end sequencing. By applying FAVR methods to whole-exome sequencing datasets, we demonstrate a 3-fold reduced single nucleotide variant shortlist compared to a conventional analysis pipeline with no detected reduction in sensitivity. This analysis included Sanger sequencing of rare variant signals not evident in dbSNP131, assessment of known variant signal preservation, and comparison of observed and expected rare variant numbers across a range of first cousin pairs. The principles described herein were applied in our recent work identifying XRCC2 as a new breast cancer risk gene and have been made publically available as a suite of software tools. FAVR methods are platform-agnostic and significantly enhance the analysis of large volumes of sequencing data for the study of rare genetic variants and their influence on phenotypes.

3680W

Accurate, Efficient Next-Generation DNA Sequencing for Clinical Carrier Screening. *G.J. Porreca, M. Umbarger, C. Kennedy, P. Saunders, B. Breton, N. Chennagiri, J. Emhoff, V. Greger, D. Maganzini, C. Micale, M. Nizzari, C. Towne.* Good Start Genetics, Cambridge, MA.

Next-generation DNA sequencing (NGS) is poised to displace genotyping technology for clinical applications because it promises richer information at low cost. However, to date a number of considerations including sequence accuracy and completeness, as well as workflow scalability, have limited its adoption in the clinical laboratory. Here we describe a NGS-based platform designed for genetic carrier screening in a clinical setting. A set of 14 genes are isolated from genomic DNA by automated multiplex target capture, tagged with molecular barcodes, and pooled and sequenced on the Illumina Hiseq system. Reads from each sample are de-multiplexed, aligned to a reference, and integrated into accurate genotype calls which are then interrogated for pathogenic mutations. A total of 42,660 bp were targeted for capture by a set of molecular inversion probes designed to tile across the target such that each base was captured by at least three different probes. Across a set of 182 DNA samples derived from cell lines or blood, a median of 99.71% of bases were sequenced to sufficient depth and quality for inclusion in genotype calling. Replicate runs exhibited a high level of concordance, with 23 discordant single nucleotide variant (SNV) calls out of 5,684,154 across 148 samples. Concordance of genotype calls with bi-directional Sanger sequence of PCR amplicons derived from a set of 193 samples was 99.97% at SNV positions (1 out of 3,973 Sanger SNV calls discordant), and 99.9998% at non-variant positions (8 out of 5,166,654 Sanger non-variant calls discordant). We identified, in a set of 57 samples, a total of 94 mutations (out of 94) that have been previously reported to be causative of recessive Mendelian disease. We also determined that 10 previously uncharacterized samples contained mutations that were either known or expected to be pathogenic. Clinical carrier screening has traditionally been performed using genotyping technology, and demands high analytical accuracy. Our NGS platform achieves exceedingly high concordance with Sanger sequencing, long considered an accuracy gold standard. Furthermore, we made high-confidence genotype calls across the vast majority of basepairs within our target regions. Taken together, these results indicate that NGS, when paired with the appropriate sample preparation methodology, automation, and data analysis, can deliver the performance required for use in clinical carrier screening.

3681F

High-Throughput *CFTR* Full-Gene Analysis Using Illumina's MiSeq TruSeq Custom Amplicon Technology. *J. Radecki, S. Lee, H. Lu, S. Mexal, A. Elliott.* Ambry Genetics, Aliso Viejo, CA.

Cystic Fibrosis Transmembrane Conductance Regulator (*CFTR*) gene is responsible for the autosomal recessive genetic disorder Cystic Fibrosis (CF). CF is one of the most frequently diagnosed autosomal-recessive diseases in the Caucasian population. Using a genotyping technology that can rapidly identify disease-causing mutations is important for high-throughput general-population carrier screening, confirming clinical diagnosis, determining treatment options, and prenatal diagnosis. Here, we demonstrate Illumina's TruSeq Custom Amplicon (TSCA) technology on the MiSeq platform is capable of sequencing ~92% of known *CFTR* disease-causing mutations with an average coverage of 723X for targeted regions in 48 clinically validated samples. Importantly, using our custom bioinformatics pipeline for *CFTR* variant identification, TSCA reliably detects the IVS9 poly-T TG-repeat polymorphism, thus alleviating the need for low-throughput gel-electrophoresis genotyping. The current version of Illumina DesignStudio produces suboptimal probe design that causes coverage dropout in overlapping regions and is incapable of designing oligos to detect known gross deletions/insertions and complex rearrangements; therefore a secondary assay is needed to detect causative structural variations. Due to the high GC content of the 5'UTR and exon 1, these regions do not meet the minimum quality values 10 mapping and 20 Q-score, creating a need for another secondary assay. In conclusion, TSCA coupled with our custom bioinformatics pipeline provides a robust high-throughput alternative to conventional full gene Sanger sequencing for *CFTR* clinical diagnostics; however, the suboptimal amplicons generated for the 5'UTR and exon 1 regions along with the inability to detect known gross deletions/insertions and complex rearrangements results in additional side assays which increases testing cost and turn-around-time.

3682W

Sensitive detection of minor variants and viral haplotypes using Single-Molecule, Real-Time (SMRT®) sequencing. *A. Sethuraman[1], Y. Guo[1], M. Brown[1], J. Toma[2], A. Newton[2], W. Huang[2], M. Sugiyama[3], C. Petropoulos[2], M. Mizokami[3], E. Paxinos[1].* 1) Pacific Biosciences, Menlo Park, CA. USA; 2) Monogram Biosciences, South San Francisco, CA USA; 3) National Center for Global Health and Medicine, Tokyo, Japan.

Genotypic testing of chronic viral infections is an important part of patient therapy and requires assays capable of detecting the entire spectrum of viral mutations. Most genotypic assays are CE-based and offered as CLIA-validated Lab Developed Tests (LDTs). Single Molecule, Real-Time (SMRT®) sequencing offers several advantages to CE, including superior resolution of mixed populations and long read lengths capable of spanning entire viral protein coding regions. We examined detection sensitivity of SMRT sequencing using amplicon mixtures of various coding regions of both HCV and HBV. Initially, two mutations were introduced separately into a plasmid containing the NS3 coding region of HCV genotype 1b Con 1 by site-directed mutagenesis. Resulting SDMs were mixed at 20%, 10%, 5%, 2% and 1% frequencies of V36M in T54A before amplification of a 563 bp region of NS3. Additionally, larger fragments spanning NS3/4a (2,055 bp) and NS5a (1,341 bp) were amplified from plasmids containing either Con1 or the HCV genotype 1a strain H77. Amplicons were purified, concentrations normalized, and reciprocal mixtures made by serial dilution at the same minor variant frequencies. Finally, we examined detection sensitivity of SMRT sequencing to detect very low level mixtures of known resistance-associated mutations in the Polymerase Gene (P) of HBV. Three mutations were introduced into HBV plasmids by site-directed mutagenesis. Two amplicons (575 bp and 1,389 bp) were generated for both wildtype and mutant plasmids, which were purified, normalized, and mixed by serial dilution at 20%, 10%, 5%, 2.5%, 1.25%, 0.62%, 0.31%, 0.16%, and 0.08% frequencies. SMRTbell™ libraries were constructed and sequenced using standard Pacific Biosciences® C2 chemistry and protocols. We used both a reference-based approach and de novo approaches for variant detection and haplotyping. Average read lengths were 3,500 bp across samples, with 5% of reads longer than 9,000 bp. From a single SMRT ® Cell, minor variants were accurately and reliably detected down to 0.1% with simple analyses. The random error profile and long read lengths make it possible to call variants from as few as 10 molecules, while sequencing individual molecules allows phasing of mutations hundreds of bases apart. SMRT® sequencing can identify species comprising a mixed viral population, with granularity and low cost of consumables allowing for smaller multiplexing of samples and first-in-first-out processing.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3683F

High Throughput Amplicon Sequencing using the Personal Genome Machine. *J. Boland, D. Roberson, M. Cullen, M. Yeager, K. Jacobs, S. Chanock, V. Lonsberry*. Cancer Genomics Research Laboratory, NCI Frederick, Gaithersburg, MD 20855.

Since January 2011, when the first Personal Genome Machine (PGM) from Life Technologies arrived at our facility, we have been in the process of developing an amplicon sequencing pipeline that would be able process 1000s of samples across multiple amplicons, exons or regions of the human genome. Over the past year and a half, we had successfully installed a high throughput amplicon sequencing pipeline that can accommodate project sizes from single samples up to 1500 samples. Through development of barcodes, protocol improvements and the acquisition of several PGMs, we have now accomplished our goal of sequencing upwards of 1500 samples per day. The variant follow-up pipeline currently involves automated primer design (either in-house or outsourced), sample prep processes including but not limited to, Fluidigm Access Array and Life's Ampliseq products, PGM sequencing protocols and variant calling software using the latest Torrent Browser software. This presentation will highlight the developmental milestones that were achieved in the laboratory with regards to increased barcoding ability, improved protocols and reagents and, most importantly, back-end bioinformatic support that has been crucial for the laboratory success. This pipeline has primarily served as our variant validation pipeline from exome data generated from our familial exome sequencing program. In addition, the pipeline has also processed several projects outside of our exome program. These projects included regional studies and studies investigating variants across large population groups. The presentation will highlight two of these >1000 sample studies that were completed by the laboratory using this pipeline. The results will show consistent high quality results across all samples, regardless of plate position, technician or day. This project has been funded in whole or in part with federal funds from the National Cancer Institute, National Institutes of Health, under contract number HHSN261200800001E. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

3684W

A simple method for improving the limit of detection for capillary electrophoresis DNA sequencing: A comparison of methodologies for *KRAS* variant detection. *C. Davidson[1], E. Zeringer[2], K. J. Champion[4], M.-P. Gauthier[1], F. Wang[1], J. Boonyaratanakornkit[3], J. R. Jones[4], E. Schreiber[1].* 1) Life Technologies, Foster City, CA; 2) Ambion, Austin, TX, USA; 3) AcroMetrix, Benicia, CA; 4) Greenwood Genetic Center, Greenwood, SC.

Fluorescent dye terminator Sanger sequencing (FTSS), with detection by automated capillary electrophoresis (CE), has long been regarded as the gold standard for variant detection. However, software analysis and base-calling algorithms used to detect mutations were largely optimized for resequencing applications in which different alleles were expected at heterozygous mixtures of 50%. Increasingly, the requirements for variant detection are an analytic sensitivity for minor alleles of <20%, in particular, when assessing the mutational status of heterogeneous tumor samples. Here we describe a simple modification to the FTSS workflow that improves the limit of detection of cell-line gDNA mixtures from 50–20% to 5% for G>A transitions and from 50–5% to 5% for G>C and G>T transversions. In addition, we use two different sample types to compare the limit of detection of sequence variants in codons 12 and 13 of the *KRAS* gene between Sanger sequencing and other methodologies including shifted termination assay (STA) detection, single-base extension (SBE), pyrosequencing (PS), high resolution melt (HRM), and real-time PCR (qPCR).

3685F

HaloPlex target enrichment from FFPE tissues. *F. Roos, H. Johansson, M. Isaksson, P. Eriksson, L. Forsmark, F. Dahl.* Agilent Technologies, Uppsala, Sweden.

Formalin-fixed paraffin-embedded (FFPE) tissues compose an enormous archive of samples obtained from biopsies of various human cancers, in several institutions around the world. However, as DNA purified from FFPE tissue is highly degraded, damaged, and available only in small quantities, being able to perform molecular analysis of these samples is a common challenge. Extracting molecular information from FFPE samples has the potential to drastically improve our understanding of genetic components that underlie disease states.

The HaloPlex PCR method combines targeted enrichment and library preparation in a simple 6 hour protocol without the requirement of any dedicated instrumentation. The protocol associates all targeted fragments with a common primer motif which enables multiplex PCR using only one primer pair. The resulting complex pool of PCR products is ready for direct sequencing using any Illumina or Ion Torrent platforms.

A simple QC assay was developed with the aim to allow accurate prediction of individual samples' performance in a HaloPlex enrichment assay by assessing the level of fragmentation of the DNA sample. A set of rules was empirically derived to provide guidelines on how to make small alterations to the standard HaloPlex protocol based on the fragmentation state of the sample. By following these guidelines the chance of generating successful enrichment and correct sequence information was significantly increased even in highly fragmented samples.

We demonstrate HaloPlex enrichment data from matched FFPE / fresh-frozen (FF) samples of different quality states showing more than 95% sequencing coverage of the targeted region. Specificities above 90% were observed in all samples and the uniformity of sequenced reads across the target region was comparable between matched samples. The high coverage was shown to be attributed to the fact that several amplicons cover each base in a HaloPlex assay so that smaller amplicons can act as a backup for longer amplicons that are more likely to fail in fragmented FFPE samples.

3686W

Optimized sample and library preparation of FFPE tumor samples for targeted next-generation sequencing. *N. Udar, R. Haigis, E.B. Jaeger.* Illumina Inc, San Diego, CA.

The use of formalin-fixed paraffin-embedded (FFPE) tissues for DNA-based assays is limited by the quality of the DNA that can be obtained. In addition to introducing chemical contaminants that can inhibit molecular reactions, the fixation process leads to crosslinking, DNA fragmentation, and other damage to nucleotides. The extent of this damage can vary widely from sample to sample, but the end result is generally reduced performance of the DNA template. We have optimized performance by selecting an optimal DNA extraction process, ascertaining the amplifiability of the extracted material, maximizing the amount of assay template input, and modifying reagent formulation. This permitted the generation of high-quality data with 90% of FFPE samples screened to date. Four different solvents for the deparaffinization of FFPE samples were tested in combination with eight different FFPE DNA extraction kits from six different manufacturers on a variety of FFPE samples from multiple cancer tissue types. Extracted DNA was used as input for a quantitative PCR reaction, and the delta Cq for each sample was calculated relative to the Cq of the standard non-FFPE human reference gDNA from the Illumina® FFPE QC Kit. To maximize the volume of template added to the library prep assay, the hybridization buffer was modified to remove all dead volume, and reagents were added as powders rather than solutions where possible. Deparaffinization with the Qiagen DPS solution followed by DNA extraction with the Qiagen All-Prep DNA/RNA FFPE Kit yielded the most consistent high-quality sample preps for the TSCA assay. The amplifiability of these sample preps according to our qPCR measurements was completely concordant with sample performance in the library prep assay. Sample input volumes for the library prep assay were increased four-fold, and the sample success rate correspondingly improved from 75% to 90%, or 45 of 50 samples. We have optimized the performance of our library prep assay for use with FFPE tumor samples. We have also developed a qPCR assay that correctly identifies all samples that will provide interpretable sequence data. These samples yielded high quality data that enabled the identification of mutant alleles at frequencies as low as 1%, which were subsequently confirmed by allele-specific PCR. Given the large numbers of tissues preserved as FFPE material, the added robustness of the assay increases its value as a research tool for genetic analysis.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3687F

TargetRich™: Targeted Sub-Exome Sequencing. *I.A. Vasenkova[1], K. Jansen Spayd[1], T. Shvetsova[1], D.A. Kloske[1], R.C. Bachmeyer[1], D.T. Moore[1], K.E. Varley[1,2].* 1) Kailos Genetics Inc., Huntsville, AL; 2) HudsonAlpha Institute for Biotechnology, Huntsville, AL.

Next-generation sequencing (NGS) enables researchers to access large amounts of genetic information from a whole human genome or exome. Sequencing tumor and blood samples obtained from cancer patients reveals that a single tumor may contain as many as 150 mutations in the protein coding region of different genes. Some genes are found to be consistently mutated in various tumors and therefore are believed to be important for the onset or maintenance of oncogenic phenotype. The most frequent mutations in these genes have been cataloged and commercial cancer panels have been developed to amplify and deeply sequence these mutational “hotspots”. As these cancer enrichment panels cover selected portions of the gene's coding sequences, their use in clinical research and the discovery of lower frequency mutations associated with cancer is still limited. In genes such as tumor suppressors, sequence analysis of the entire coding region is especially important since any mutation in these genes could give rise to tumorigenesis.

We have developed TargetRich gene enrichment panels to enable the complete analysis of cancer associated genes. The sixty three gene TargetRich cancer research survey panel is the largest cancer gene panel available and includes oncogenes, tumor suppressor and drug metabolism genes, targeting nearly all exonic bases. TargetRich employs Nested PatchPCR™ technology, a limited-cycle PCR and targeted ligation reaction protocol allowing one to amplify hundreds of fragments together with remarkable specificity in single-tube in less than 9 hours. Platform-specific adapters incorporated during the nested ligation step allow for the high levels of multiplexing and direct sequencing of products on Illumina or Ion Torrent platforms without time-consuming library construction. Here we compare depth of coverage, uniformity and genetic variant detection achieved through whole exome sequencing using hybrid capture, TargetRich and “hot spot” sequencing approaches. We demonstrate utility of the TargetRich approach as a cost effective, scalable solution for development of high-resolution next generation sub-exome sequencing assays for clinical and discovery research.

3688W

A Functional Approach to Sequence Capture Analysis of Disease Pathways. *M. Wijdicks, D. Burgess, M. D'ascenzo, M. Brockman, J. Wendt, C. Skalitzky, D. Green, T. Richmond, L. Brown, R. Slezer.* Roche NimbleGen, 504 S. Rosa Rd., Madison WI 53719.

Enrichment sequencing has quickly become a standard technique for identifying inherited, de novo and somatic alleles responsible for genetic diseases. Whole exome sequencing is often selected for discovering candidate mutations in biochemical pathways that are still poorly understood, while disease specific gene capture panels, for example targeting oncology and neurology candidate loci, are a more efficient embodiment of the technology for high-throughput, focused analysis of better characterized disorders. To increase the efficiency and success of allele identification, we have developed a series of disease specific gene panels to probe the most relevant targets within the human genome for a fraction of the cost, time and data analysis requirements of whole exome and whole genome sequencing. We present two designs. The first is an oncology focused panel that targets 578 known or predicted cancer genes, including content from the MHC where applicable. The second design is a neurology focused panel that targets 256 candidate genes representing 87 different disease classes. We report on the application of these sequence capture panels for the discovery and validation of disease specific genetic variants.

3689F

Comparison of two NGS preparation methods for preimplantation and prenatal screening and diagnostics. *J.P. Langmore, E. Kamberov, M. Mastronardi, T. Tesmer.* Corporate, Rubicon Genomics, Inc., Ann Arbor, MI.

To improve ease of sequencing fetal cells and fetal cell-free DNA for prenatal diagnostics, as well as for sequencing single embryo cells for preimplantation diagnostics, two new methods for producing sequencing templates for NGS were compared. The PicoPLEX™-scD Prep and the ThruPLEX-FD Prep kits are capable of producing sequencing libraries of single cells and small plasma samples. Earlier versions of these technologies have been used extensively for preimplantation testing and prenatal and cancer testing using array and PCR assays. Each technology has been modified to attach adapters with Illumina sequencing and indexing sequences. We compared these two NGS prep kits for specific applications in prenatal diagnostics and preimplantation screening and diagnostics using workflows able to prepare 200 NGS samples per day. Results: ThruPLEX-FD Prep Kit is a 1-tube, 2-hour, 3-step ligation-based process to prepare fragmented DNA for NGS without intermediate purifications. Sequence representation was excellent from 15% to 85% GC in human samples from 20 pg to 100 ng of plasma DNA, yielding mean quality scores >35, <5% unmapped reads, <0.4% mismatch with hg19, and >95% intact PE reads. The high efficiency, low background, and simple workflow made it possible to test single cells and small plasma samples yielding more complex libraries than other prep methods. PicoPLEX-scD Prep Kit is a 1-tube, 3-hour, 4-step process to manually prepare hundreds of single cell or sorted chromosomes for targeted sequencing per day. Greater than 95% of sorted cancer cells or blastomeres were successfully prepared for sequencing. Typically >99% of reads were high quality, 98.5% mapped to human, and <0.6% did not match hg19. Indexed NGS was able to reproducibly detect human single-cell copy number changes less than 20 kb in size. The targeted NGS results produced by PicoPLEX-SC also can be used to prepare cell-free plasma DNA for very reproducible determination of abnormal chromosome numbers. ThruPLEX-FD and PicoPLEX-scD were used to sequence duplicate aliquots of plasma and single cells. Sequence coverage and accuracy was compared for these two types of samples for purposes of establishing the strengths of both for prenatal and preimplantation diagnostics. research and testing.

3690W

Droplet Digital PCR Enables Reliable Discrimination of Copy Number Variation. *J.R. Berman, N. Heredia, J. Regan, L. Montesclaros, S. Hodges, C. Troup, G. Karlin-Neumann.*Digital Biology Center, Bio-Rad Laboratories, Pleasanton, CA., USA.

The ability to quantify and discriminate copy number variation (CNV) at a given locus is increasingly critical to a diverse set of research and diagnostic applications. CNVs encompassing both deletion and duplication of wild-type and mutant alleles have been associated with cancers, neurological, autoimmune, and other diseases, adverse drug response, and other traits. Existing techniques, namely qPCR, are of limited utility for CNV determination due to imprecise discrimination at higher (>3) copy number states. Droplet digital PCR (ddPCR) enables accurate and reproducible copy number determination using a simple, cost-effective workflow amenable to high throughput. Using single-well ddPCR, consecutive copy number states can be distinguished between samples of 5 and 6 copy number at 95% confidence levels. We present use of the QX100 ddPCR system to assess CNV of multiple genes with low to high copy number states, including CYP2D6, MRGPRx1, and CCL3L1. We further demonstrate use of CNV analysis by ddPCR to interrogate loci implicated in cancer and other diseases. Strategies for assay optimization for CNV analysis by ddPCR will be highlighted.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3691F

Integrating high-quality DNA sequence capture with MiSeq for clinical sequencing. *P. Shen[1], W. Wang[2], A. Chi[1], R. Davis[1], C. Scharfe[1].* 1) Genome Technology Ctr, Stanford Univ, Palo Alto, CA; 2) Bioinformatics and Computational Biology Dept, Univ of Texas MD Anderson Cancer Ctr, Houston, TX.

The accurate and complete selection of candidate genomic regions from a DNA sample before sequencing is critical in molecular diagnostics. Several recently developed DNA sequence capture technologies await substantial improvements in performance, cost, and multiplex sample processing. Here we present the utility of long padlock probes (LPPs) for targeted exon capture followed by Illumina MiSeq sequencing. One LPP was sufficient to capture sequences from <100–500 bp in length, while LPPs were pooled together to target 5,471 exons from 524 disease candidate genes in a single-tube reaction. MiSeq sequencing identified a 99% exon coverage at >1x and 95% at >10x (mean coverage 276 reads). More than 98% of the QC-passed reads mapped to our target regions (specificity) and 91% of the targets were within 50 fold coverage uniformity. Our targeted sequencing methodology was highly reproducible and quick (48hrs) and detected DNA variants at a concordance of 98% on the basis of known sample genotypes and Sanger verification. From the same sequencing data we detected copy number variation (CNV) at exon-level resolution in a patient with ornithine transcarbamylase (OTC) deficiency and in another case with Trisomy 13. We also performed multiplex targeted sequencing of up to four samples per MiSeq run for SNP and CNV detection in the 524 candidate genes, making our method particularly interesting for molecular genetic and clinical applications.

3692W

Improved Performance of Solution Based Target Enrichment by Spike-in of Individually Synthesized Capture Probes. *G.R. Mehta, Locus Development, Inc.* Locus Development, San Francisco, CA.

While DNA sequencing costs have decreased and throughput increased at an incredible rate recently, the cost for sequencing a complete genome at high depth is still prohibitively expensive for clinical applications, and thus the vast majority of clinical sequencing-based tests still consist of decoding targeted portions of the genome (exome, cancer panels, etc.). Over the past few years, several commercial methods and technologies of targeted enrichment have emerged. One such example is solution hybrid capture (SHS), where specific biotinylated probes hybridize to genomic DNA targets and then are selectively enriched with the use of streptavidin beads and finally sequenced. There are many advantages to this process, namely the ability to capture precise targets efficiently and in an automatable manner, and recent commercial success of these companies is evident. Unfortunately, there are limitations and challenges to SHS—targets that are high GC, homologs and paralogs, and the inability to accurately assess quality of capture probes on a large scale. For clinical testing, it is imperative that all relevant regions of a gene are assayed in order to provide a high level of diagnostic completeness and accuracy. Here, we present a method to overcome some of these deficiencies by performing spike-ins of individually synthesized and quality controlled oligonucleotide capture probes into our existing SHS probe pools to complement them prior to sequencing. We validated our method by sequencing twelve HapMap samples on the Illumina miSeq platform and directly comparing a fill-in PCR approach of 800 155mer amplicons (our current method of overcoming the gaps left by our capture) to a spike-in strategy (~1200 120mer probes designed to the same regions as the PCR amplicons). Advantages are numerous: we have successfully seen greater on-target capture, reduced GC bias, more uniform coverage, whilst maintaining similar sensitivity and specificity. Furthermore, the spike-in method incorporates the ability to fine-tune our final captured product by the addition or removal of probes easily and is fully automatable / more scalable than PCR.

3693F

Accurate Multiplexing for Clinical Next Generation DNA Sequencing. *M.A. Umbarger, G.P. Porreca.* Good Start Genetics, Cambridge, MA.

Next generation sequencing (NGS) has begun to supplant a number of traditional technologies within the clinic. While the per-base cost of NGS is extremely low, the high capacity of NGS instruments necessitates that clinical labs employing such instruments divide the per-run fixed cost over multiple samples. Such sample multiplexing is frequently accomplished via the use of patient specific DNA barcodes that are covalently attached to patient molecules via ligation or the polymerase chain reaction. Since these barcodes uniquely tag each molecule, it is possible to map a given sequenced molecule to the appropriate patient by simply sequencing its associated barcode.

While the utilization of patient specific barcodes reduces per-sample sequencing costs, this approach is not without risk. For example, if a significant fraction of the molecules from a given patient are accidentally associated with the barcode for another patient, the genotypes of both patients may be incorrectly called. In fact, simulations indicate that at a sequencing depth of 20X, a misassociation rate of 1 in 1,000 yields a clinically non-negligible genotype error rate.

We have developed strategies for identifying and minimizing the contribution of three major sources of incorrect sample and barcode pairing. First, to identify any stock barcode oligonucleotides which are contaminated with other barcode oligos, we have instituted a sequencing-based barcode quality control procedure that allows us to detect contamination down to a level of 1 in 100,000. Second, to minimize the frequency at which sequencing errors lead to a given barcode incorrectly mapping to another barcode, we have developed a quality score-based filtering strategy. Finally, to identify and filter out cases in which recombination between homologous molecules from different patients has yielded chimeric molecules containing the genomic sequence from one patient and the barcode for another, we have developed a dual barcoding strategy in which barocodes are incorportated at each end of a patient DNA molecule. Together, our strategies minimize the rate of incorrect barcode-patient association and thereby render our barcode multiplexing strategy compatible with the high accuracy requirements of clinical settings.

3694W

Enhanced performance of the Illumina MiSeq® next-generation sequencing ecosystem. *K. Hall[1], C. Tregidgo[1], I. Rasolonjatovo[1], A. Breton[1], J. Bwanali[1], A. Jackson[1], E. Vermaas[2], D. Bond[2], M. Siu[2], C. Bruce[3], A. Powell[3], P. Tran[3], D. McBride[1], M. Ross[1].* 1) Illumina Cambridge Ltd., Saffron Walden, Essex, United Kingdom; 2) Illumina Hayward 25861 Industrial Blvd. Hayward, CA 94545 USA; 3) Illumina, Inc. 5200 Research Place San Diego, CA 92122 USA.

Recent improvements to the performance of the MiSeq personal sequencing system has enabled many key metrics to be improved. The yield of the system has increased to around 7 gigabases per run with the read length supported increased to 250 bases with paired end capability. The run time for a paired end 150 cycle run has been reduced to around 20 hours with an average quality score of greater than 75% of Q30 bases. The number of reads passing filters has also been increased to around 12–15 million clusters passing filters. The improvements are derived from a new modified SBS (sequencing by synthesis) chemistry and a hardware upgrade. The system maintains its features of being extremely easy to use with simple workflows, having on board cluster generation, paired end fluidics, plug and play reagents and RFID tracking. As such no further ancillary equipment is needed. Data for most applications is processed to variant calling and alignment is less than 2 hours on instrument with the option of using BaseSpace. These advances have enabled many more applications to be targeted. In this presentation we will demonstrate some of the uses of the advances, these will include the improved de novo assembly performance from the longer reads with a variety of genomes from bacterial and viral sources. Other examples will include resequencing and RNA-Seq applications. In particular we will demonstrate the use of highly multiplexed targeted resequencing of amplicons with the TruSeq Custom Amplicon workflow together with its application in human genomic studies.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3695F

Single-Day, Highly Multiplexed Amplicon Sequencing with MiSeq®. *A. Iyer, A. Tian, K. Chang, E. Guzman, E. Upsall, I. Lewis, M. Won, W. Chang, D. Pokholok, R. Haigis, S. Norberg, M. Ronaghi, K. Gunderson, R. Shen, C. Lin.* Illumina, Inc., San Diego, CA.

Targeted enrichment strategies enable cost-effective analysis of desired genomic regions with next-generation DNA sequencers. We introduce the TruSeq® Custom Amplicon Kits, a new sample preparation method developed and optimized for high-throughput amplicon sequencing on MiSeq, the personal sequencing system from Illumina. Up to 384 user-defined genomic regions can be enriched in a single tube with high specificity (>90% of reads on target) and high uniformity (>85% of targets represented at ≥0.2X mean coverage) in <8 hr (from purified genomic DNA to sequencing-ready library). By comparison, only a dozen or so targets can be amplified in a single tube with multiplex PCR, and hybridization-capture enrichment approaches require 3 days or more from DNA to library. In addition, we have developed a robust and highly scalable sample indexing strategy supporting 96 libraries to be sequenced simultaneously with MiSeq. We also will demonstrate a new library quantification and normalization approach that allows 96 libraries to be adjusted to sequencing-ready concentrations in <1 hr instead of current approaches relying on qPCR and subsequent sample dilution. These innovations streamline amplicon sequencing workflows, and provide unprecedented amplicon and sample multiplexing. Here we present data demonstrating high-accuracy detection of single-nucleotide polymorphisms (SNP) and structural variants, some of which are implicated in drug metabolism and cancer biology.

3696W

WildFire: A simple monoclonal colony generation technology without emulsion PCR. *Z. Ma, K. Lao, R. Lee, S. Goyal.* Life Technologies Inc., Foster City, CA.

We have developed a new massive parallel monoclonal template generation technology, called “WildFire”, where sequencing libraries are in-situ isothermal amplified directly on the surface of a 5500 flowchip. Sequencing libraries are added directly to the 5500-series Genetic Analyzer flowchip, whose surfaces have been coated with a special library-adaptor capture oligonucleotide. A DNA polymerase reaction mix is added, and in a single isothermal step for 30 minutes, single templates are amplified in-situ on the flowchip. The net density of sequencing-colonies created in this manner far exceeds anything currently utilized in next-generation sequencing, reaching 1.5 million colonies per mm2 per flowchip surface. During in-situ amplification, the capture oligonucleotide is “consumed”, and each individual nucleic-acid fragment “spreads” (like a WildFire) inside the flowchip until reaching an adjacent library fragment(s). When the individually-growing fragments “meet”, the amplification step terminates, because all of the surface-bound primer was consumed. These “self-assembled”, spatially resolved, monoclonal colonies, are then sequenced by SOLiDTM chemistry. The resulting colony-sequencing reads maintain the same high accuracy as our bead-based method. Full genomes (from bacterial to human), exomes (human), and transcriptomes (human) have now been sequenced using WildFire technology. WildFire technology greatly improves NGS workflow, increases throughput, and significantly decreases net cost-per-genome.

3697F

Nextera® Enrichment: A New Nextera Library Prep Protocol for Targeted Enrichment Supporting 12-plex Pre-enrichment Sample Pooling. *P. McInerney[1], S. Melnyk[1], M. Chen[1], M. Tsan[1], S. Cooper[1], H. Grunenwald[2], M. Lewis[1], R. Shen[1], J. Whitacre[1].* 1) Illumina, Inc., San Diego, CA; 2) Epicentre, Inc. (an Illumina Company), Madison, WI.

Targeted sequencing is a cost-effective method to sequence specific regions of the genome for large sample collections that require high coverage. Applications include discovery of rare variants in cancer and disease association studies. Currently available targeted sequencing workflows require at least 1ug of DNA as starting material. This requirement often precludes analysis of samples that are precious or limited in quantity, such as tumor biopsies and buccal swabs. We describe a new sample preparation method utilizing Nextera sample preparation technology to create up to 96 uniquely barcoded libraries from only 50ng of input DNA. The new protocol includes reagents for target enrichment, using probes for either exome or custom panels. This protocol generates libraries with optimal insert sizes and improved GC coverage relative to TruSeq libraries. This new low input, streamlined Nextera Enrichment protocol generates 96 enriched libraries in 2.5-days; making it attractive for users who require high-throughput sample processing and a fast time-to-answer. Furthermore, the protocol enables pre-enrichment sample pooling of up to 12 samples, the highest level of multiplexing currently available. We present data showing simultaneous, uniform enrichment of 12 different Nextera indexed samples in a single enrichment reaction using either Illumina's TruSeq® Exome or a panel with ~400 cancer genes. Specifically, we show that the Nextera Enrichment assay exhibits high specificity with >60% of reads on target and excellent coverage uniformity with >85% of targeted bases covered at 0.2X mean coverage.

3698W

Development of Ion Torrent's 400-base Sequencing Technology. *X. Peng, G. Luo, T. Lincecum, E. Tozer, D. Mazur, K. Aguinaldo, G. Lowman, M. Landes, B. Strohecker, T. Nikiforov, P. Vander Horn.* Ion Torrent, LIFE Technologies, Carlsbad, CA.

We have made substantial advancements on Ion Torrent's semiconductor chip-based sequencing system in a number of areas including on-beads library amplification, sequencing polymerase, and sequencing reaction condition. These innovations resulted in high quality and robust 400-base sequencing performance on Ion Torrent PGM platform. The improvement on the sequencing polymerase and reaction conditions is based on a new mutated enzyme that shows high accuracy and high processivity sequencing behaviors. Our efforts on optimizing 400bp insert library emulsion PCR conditions resulted in enhanced on-beads library amplification as demonstrated on PGM by the increased library key signals during sequencing run. We show here the optimal 400-base sequencing performance on an ion 316 chip by sequencing a 404bp insert E. coli dh10b library and compared it with the performance produced by using ion torrent's commercially available 200-base sequencing reagents and protocols. The peak in the AQ20 read length histogram shifts dramatically to the 400bp region using the improved reagents and protocols. The raw error rate @400bp position decreases dramatically to 0.857%, and the overall mean raw read accuracy increases to 99.5%. The %AQ30 base yield holds steadily between 76% and 64% of the total aligned bases during the entire 400bp sequencing run, which translates into 586M AQ20 bases throughput on a Ion 316 chip. The homopolymer accuracy also improved to 99.2% level for the 6-mer homopolymer stretch and the number of Indels per 100 bp also reduced to 0.41 by the new Ion Torrent's 400-base sequencing chemistry.

3699F

From tumor to genome sequence with nanogram quantities of DNA: ultralow target capture and DNA-Seq from FFPE samples. *M. Phelan, L. Pham, G. Miyada, S. Kain, T. Cormier.* NuGEN Technologies, Inc., San Carlos, CA.

With the rapidly increasing number of NGS samples, many laboratories face challenges in reproducibly and cost-effectively generating high-quality libraries for RNA-Seq, ChIP-Seq and DNA-Seq. To that end, NuGEN® Technologies, Inc., has partnered with Advanced Liquid Logic, Inc., (ALL) to develop and commercialize the Mondrian™ SP System - a digital microfluidics system that automates genomic sample preparation for NGS. Several NuGEN Library Systems are available for the Mondrian™ SP Workstation, including the Ovation® SP Ultralow and the Encore® SP Rapid Library Systems. The Ovation SP Ultralow Library System offers scientist the ability to generate NGS libraries from as little as 1ng of human dsDNA (making it well suited for very low input applications like ChIP-Seq, and DNA-Seq from FFPE samples). The Encore SP Rapid Library System is an amplification-free protocol for library preparation that is ideal for use with high-GC genomes and/or in conditions where scientists wish to avoid introducing amplification bias into the final library. In this poster we compare the results of using various human DNAs (including HapMap and FFPE samples) as template for both the Ovation SP Ultralow Library System and the manual Ovation Ultralow Library System. The resulting libraries were used in a target capture and enrichment protocol prior to sequencing on an Illumina® HiSeq™ 2000 instrument using Illumina TruSeq™ SBS Kit v3-HS. Additionally an Ovation SP Ultralow library (whole genome human) was sequenced directly on three lanes of the HiSeq 2000 (without performing target capture and enrichment prior to sequencing) to investigate the levels of coverage and duplication seen in a complex eukaryotic library prepared at nanoliter volumes via microfluidic manipulation.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3700W

GnuBIO Desktop Sequencer - Fully Integrated DNA Preparation Sequencing and Analysis. *T. Raz, A. Aslam, J. Boyce, N. Nerkizian, J. Emhoff, A. Esmall, J. Fanning, H. Ghandour, K. Moulton, P. Stokes, T. Hung, S. Kiani, P. Mary, J. Healy.* GnuBio Inc., Cambridge, MA.

DNA sequencing has become fundamental to biological research, and is poised to be a crucial tool in clinical medicine, once several technical challenges are answered. Sample preparation (enrichment and library creation) is expensive and labor-intensive. While throughput is high, separate samples must be accumulated, bar-coded, and comingled to keep cost from becoming prohibitive. In addition, some current instruments have lower accuracy than would be suited to diagnosis or treatment selection. Certain applications require longer read-lengths than are available from some platforms. Finally, interpretation of the raw output from today's equipment requires specialized personnel, hardware, and some considerable time. GnuBIO is developing an instrument that integrates the sequencing work flow, beginning from genomic DNA, and encompassing enrichment, PCR amplification, sequencing, and analysis, delivering long reads and variant-calls on completion. The user injects a sample into the instrument after which all steps are automated. Microfluidics permits the process to be automated and integrated. Independent patients' samples flow through the instrument serially, obviating the need to accumulate or intermix patients' samples. Hundreds of thousands of emulsion droplets each act as independent reaction vessels, following their way through the steps of enrichment and amplification. Each sample droplet is then sequenced. The data arrive incrementally and once a patient's last probed sample droplet has worked its way through to the last optical detection, analysis is virtually complete. The GnuBIO platform is an innovative synthesis of technologies that are separately well understood including emulsion-based microfluidics, a polymerase extension based assay, diffractive optics and commodity dyes. In prototype instruments, the platform so far shows a per-base error rate of 0.03%, which is independent of position along the read. Reads reach 600nt routinely, and variant calling, averaged against all positions comes in at more than Phred 70. The GnuBIO system will remedy many challenges faced by researchers today, such as obviating the requirement to prepare sequencing libraries and eliminating the need to optimize sample capture reagents. Due to the integration of sample capture, amplification, and sequencing within emulsions (without ever breaking the emulsion), the system will also provide uniform coverage across the target.

3701F

Utilizing next-generation sequencing for exome analysis. *K. Stangier, F. Ernst, Y. Kumar, T. Paprotka.* GATC Biotech AG, Konstanz, Germany.

Next generation sequencing has changed the possibilities for analyzing human exomes. Several commercial kits are available allowing exome enrichment with subsequent library preparation in a cost efficient way. These libraries can be directly analyzed on e.g. an Illumina HiSeq. Enrichment is performed by pull down of coding regions with baits, which differ in overall design depending on the manufacturer. The depth of analysis can be enhanced by increasing the number of reads by using higher sequencing coverage mapped on the genome or the targeted region. We used commercially available exome enrichment kits and designed a pipeline for data quality analysis, mapping and especially SNP detection. The mapping of reads was analyzed in detail to determine the efficiency of enrichment for the targeted exons. Variants were filtered according to several criteria like base quality, read quality, thresholds for coverage, and the overall mapping quality. In order to improve the overall quality of reads for data processing, a quality score which represents the probability of a particular base mismatch in the reference genome was established. This highly increased the probabilities of validating certain mismatches. In addition, FFPE samples were used for exome enrichment and the mapping and SNP detection compared to other starting material. Utilizing exome sequencing in clinical and genetic diagnosis and personalized disease risk profiling is expected in the near future. Optimizing the sensitivity and data analysis pipelines will help to integrate exome analysis into a common clinical setup.

3702W

Novel enrichment reagent for the study of the human microbiome. *F.J. Stewart, G.R. Feehery, E. Yigit, E.T. Dimalanta, B.W. Langhorst, L.M. Apone, P. Liu, D.B. Munafo, C.J. Sumner, J. Bybee, L.M. Mazzola, T.B. Davis, S. Pradhan.* New England Biolabs, Inc., Ipswich, MA.

Nucleic-acid based techniques such as hybridization, PCR, qPCR and Next-Generation Sequencing (NGS) offer a rapid and highly sensitive option for detecting pathogenic bacteria directly from specimens, when compared with culture-based techniques. The overwhelming amount of human DNA in many specimens consumes most of the sequencing capacity of high-throughput sequencers, limiting the detection and identification of pathogens. As a consequence, analyses of a metagenome or microbiome from clinical samples by NGS or PCR are expensive, difficult and time consuming. To address this problem, we have developed a unique method for the separation of large pieces (~15kb) of human DNA from similar sizes (~15kb) of bacterial DNA using Eukaryotic Binding Protein (EBP). As a demonstration of the efficacy of this methodology, DNA extracted from human saliva was added to the EBP beads and incubated for a period of time. A magnetic field was then used to separate the human DNA bound to beads from the supernatant containing bacterial DNA. The enriched bacterial DNA from the supernatant was sequenced on multiple NGS platforms and was found to have a dramatic increase (9 fold) in the number of reads matching bacterial genomes. Likewise, a substantial drop (50 fold) was seen in the number of reads matching the human genome. Bacterial diversity following the procedure remained intact. This simple methodology can be used to analyze entire microbiomes in a cost-effective manner utilizing established NGS platforms, as well as newer single molecule sequencing technologies.

3703F

Structural Variations Identified Using Solid-state Nanodetectors. *J. Thompson, B. Bready, D. Dederich, D. Hevroni, M. Jouzi, H-Y. Lee, D. Lloyd, P. Mukhatira, M. Nadel, J. Oliver, J. Saraidaridis.* Nabsys Inc, 60 Clifford St, Providence, RI.

Many human diseases are caused by insertions, deletions, or duplications of large segments of DNA. Understanding the nature of structural variants and how prevalent they are within a population of cells is sometimes critical for diagnosing and treating patients. However, many structural variants are difficult to detect using classical and next-gen sequencing techniques. Positional sequencing of DNA using nanodetectors can provide the relative locations of various sequences and thus a direct readout on structural rearrangements. Distances between probe-bound sites are determined electronically using solid-state nanodetectors. Long-range information is preserved in these single-molecule reads so structural rearrangements and duplications are easily identified and correctly assembled. Results for the detection of deletions and inversions of various sizes will be presented.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3704W

Automaton of whole exome and targeted sequencing workflows in a high throughput lab. *M. Zilka, B. Marosy, A. Robinson, J. Gearhart, B. Craig, J. Romm, K. Doheny.* Center for Inherited Disease Research (CIDR), IGM, JHU - SOM, Baltimore, MDCenter for Inherited Disease Research (CIDR), Johns Hopkins University, Baltimore, MD.

The Center for Inherited Disease Research (CIDR) provides high quality next gen-sequencing (NGS), genotyping and statistical genetics consultation to investigators working to discover genes that contribute to disease. CIDR began production level automated genotyping in 1996 with Human STRP linkage panels and has continued to automate its production workflows from Illumina GWAS arrays to CIDR's newest offerings, whole exome and targeted sequencing. CIDR requires the use of laboratory automation and informatics tools in order to process samples quickly and to reduce human error in a high-throughput environment. CIDR has created a workflow implementing a Beckman FX®, a Perkin Elmer Multiprobe®, a Covaris E210 and an Agilent© Bravo in its automation of library prep using Agilent SureSelect™ XT Library Prep reagents. Back up programs for the FX and Multiprobe were created on the Beckman Biomek® 2000. For the targeted enrichment phase, CIDR has added the use of a Perkin Elmer Multiprobe, a Beckman Biomek® 2000 and an Agilent Bravo. The Perkin Elmer Janus® is used to normalize and pool samples for clustering on the cBot followed by 76bp paired- end sequencing on the Illumina® HiSeq™ 2000. At several points in the library prep and target enrichment workflow, QC steps are performed to ensure that samples are being processed correctly and should continue in the workflow. CIDR uses a Beckman FX and an Agilent Bravo to make aliquots of the samples that are run on the Agilent BioAnalyzer or the Caliper LabChip®GX. Samples that require pooling before sequencing are pooled by a Perkin Elmer Janus and are then QC'd on the Illumina MiSeq. In addition, CIDR has also integrated the Exemplar LIMS and customized java plugins to help with validation, sample worklist generation and sample tracking. With each addition of automation to the workflow, QC measures were compared to manual processing to verify that the additions of the automated steps did not decrease the quality of the data that was being generated. From our pilot studies, 73 samples were processed with the automated sequencing workflow. The QC measures reviewed were percent target bases 8X, mean target coverage, mean Ti/Tv ratio, mean concordance to GWAS array calls, estimated library size, mean sensitivity and mean percent duplication. Over a period of two weeks, our current workflow can prepare ten 96 well plates for sequencing.

3705F

Automated Genomic DNA QC ensures High Quality Data from Downstream Workflows. *M. Gassmann[1], D. Rabiller[2], A. Padmanaban[3], D. McDade-Walker[2].* 1) Agilent Technologies, Waldbronn, Germany; 2) Agilent Technologies, Loanhead, UK; 3) Life Sciences Application Centre, Agilent Technologies, Bangalore, India.

Assessing the quality of experimental starting material is key to the success of any assay. Ensuring that genomic DNA (gDNA) is of sufficient quality to continue with the Next Generation Sequencing (NGS) or array Comparative Genomic Hybridization (aCGH) workflows saves precious time and money further downstream. The new Agilent Genomic DNA ScreenTape has been developed for the electrophoretic separation, sizing and quantification of genomic DNA. This new ready-to-use ScreenTape device, which runs on the Agilent 2200 TapeStation instrument, provides a reproducible QC method for analyzing the integrity of genomic DNA starting material combined with the convenience of an automated system. With minimal sample preparation, automated loading and a variable throughput system, digital results can be presented in a gel image, data table and electropherogram within approximately 1 minute per sample. In addition, the ability to overlap and compare electropherograms within the 2200 TapeStation software enables the discrimination of sample quality across different degradation states and sample types. We present data that shows the Genomic DNA ScreenTape can easily verify the integrity of gDNA starting material with sensitivity, precision and accuracy.

3706W

Optimization of a miRNA Expression Profiling Workflow for Ion Semiconductor Sequencing. *T. Guettouche[1,2], J. Clarke[3], A. Andersen[1], L. Navarro[2], Y. Cardentey[2], W. Hulme[1], G. Bademci[1], D. van Booven[1], D. Hedges[1], M. Pericak-Vance[1], J. Gilbert[1].* 1) Hussman Institute for Human Genomics, University of Miami, School of Medicine, Miami, Florida; 2) Sylvester Comprehensive Cancer Center, University of Miami, School of Medicine, Miami, Florida; 3) Department of Epidemiology and Public Health Division of Statistical Theory and Methods, University of Miami, School of Medicine, Miami, Florida.

MicroRNAs (miRNAs) are non coding RNA molecules with a size range of 18 to 24 nucleotides that have important functions in regulating numerous cellular processes. Several technologies such as microarrays, real-time PCR and second-generation deep sequencing have been used to profile miRNA expression and characterize variances in miRNA species. This study focuses on miRNA expression profiling using ion semiconductor sequencing with an improved workflow that utilizes multi-channel microfluidics to isolate miRNAs from total RNA. This microfluidics approach replaces urea polyacrylamide gel electrophoresis for size fractionation and enrichment of miRNAs commonly used in sequencing protocols. We compared miRNA expression profiles of commercially available preparations of lung and brain total RNA obtained by ion semiconductor sequencing, second-generation sequencing and real-time PCR array. Using the microfluidics approach and ion semiconductor sequencing we detected 434 and 238 unique known mature miRNAs (mirBase18) in total RNA from brain and lung respectively. Using the urea polyacrylamide gel electrophoresis approach to fractionate total RNA from brain and lung followed by ion semiconductor sequencing we detected 343 and 224 mature respectively. The urea polyacrylamide gel electrophoresis approach produced 60% and 94% of the total number of reads compared to the microfluidics approach for brain and lung respectively. Using a subset of miRNAs detected by sequencing and available on a real-time PCR miRNA array we validated detection of 88% and 79% of brain and lung miRNAs respectively. Our workflow streamlines and significantly improves the utilization of ion semiconductor sequencing for miRNA profiling both in a research and clinical environment.

3707F

Functionalized Nanoparticles For Effective DNA Purification. *A. Lai, A. Fu.* NVIGEN, Inc 265 Sobrante Way, Suite H Sunnyvale, CA 94086.

Multifunctional nanoparticles represent a dexterous platform that can be used to realize innovative solutions for translational research. We have developed a simple and convenient method for the isolation, purification and detection using functionalized nanoparticles. We show that DNA fragments were easily recovered and purified by our high-efficiency magnetic nanoparticles with comparable performance alongside current commercially-available magnetic beads. Purified DNA substrates can be incorporated into experimental workflows, such as for sequencing and detection. Our results serve a basic demonstration in integrating the chemico-physical versatility of biofunctional nanoparticles for biomedical applications.

3708W

POP-ONE™: A New Sieving Matrix for Capillary Electrophoresis That Supports a Wide Range of Applications With a Single Instrument Set-up. *J.A. Romero, B.F. Johnson, F. Mercer, J.A. Fisher, D. Rodriguez, S. Hung, M. Wenz, W. Liao, J. Lee.* Life Technologies, Foster City, CA.

The sieving polymer formulations, POP-4®, POP-6™, and POP-7™, manufactured by Life Technologies for capillary electrophoresis (CE) are designed to provide optimal performance for a variety of separation applications. A new sieving polymer formulation, POP-ONETM, has been designed to capture a broad spectrum of the electrophoretic properties of POP-4®, POP-6™, and POP-7™ so that a single POP formulation can now be used for multiple sequencing applications. For example, POP-ONETM provides better resolution for fragment analysis applications than POP-4®, while providing slightly faster run times. Use of POP-ONETM can provide substantially faster sequencing run times than experiments using POP-6™. Further, significantly better resolution of short sequencing fragments is obtained using POP-ONETM compared to the resolution obtained using POP-7™. For sequencing applications, using either a 36 cm or 50 cm capillary array, a set of three run modules were developed that provide (1) the longest possible basecalling read-length without consideration of run time, or (2) the best compromise between basecalling read-length and run time, or (3) the fastest run time but with a shorter read-length (> 500 bases). For fragment analysis applications, run modules have been developed to support a range of applications from very fast analysis of short fragments (e.g. single-base extension genotyping) to high-resolution analysis of long fragments up to 1kb long. We will compare the resolution performance of POP-ONETM to POP-4®, POP-6™, and POP-7™, and provide examples of sequencing and fragment separations with the different run modules.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3709F

Automation of the Agilent Target Enrichment Portfolio. *M.R. Visitacion[1], J. Karbowski[2], F. Roos[4], B. Arezi[5], B. Novak[1], M. Isaksson[4], A. Giuffre[3], S. Happe[3], D. Roberts[1], E. Leproust[1], F. Dahl[4].* 1) Genomics R&D, Agilent Technologies, Inc, Santa Clara, CA; 2) Agilent Automation Solutions, Agilent Technologies, Inc, Santa Clara, CA; 3) Genomics R&D, Agilent Technologies, Inc, Cedar Creek, TX; 4) Genomics R&D, Agilent Technologies, Inc, Uppsala Sweden; 5) Genomics R&D, Agilent Technologies, Inc, La Jolla, CA.

With the advent of desktop sequencers, turn-around time from sample to sequencing data has been dramatically reduced. As a result, the need to process more samples reproducibly and efficiently has increased. We have developed several options for target enrichment, including SureSelect XT, SureSelect XT2, and HaloPlex. These sample prep kits are designed to be compatible with multiple sequencing platforms, including Illumina HiSeq and MiSeq, SOLiD 5500, and Ion Torrent. We have also developed validated companion protocols for the Agilent Automation solutions NGS workstation. These include support for SureSelect XT for Illumina and SOLiD 5500, SureSelect XT2 (pre-capture pooling) for Illumina, and HaloPlex for Illumina and Ion Torrent. The Automation Solution enables the processing of 96 HaloPlex samples per day and up to 96 SureSelect XT or XT2 samples at one time. The addition of automation allows greatly increased throughput and reproducibility, while maintaining excellent performance. Herein we describe the instrumentation, processing workflow and performance data for the automation system and the respective target enrichment platforms. In Summary, we provide easy to use, automated solutions for target enrichment enabling analysis of discrete genomic regions with unprecedented depth, accuracy and throughput.

3710W

Sequencing of complete HLA haplotypes: resequencing and population studies. *M.S. Won[1], S. Norberg[1], T. Royce[1], T. Dunn[1], T. Mann[1], N. Nemat-Gorgani[2], L.A. Guethlein[2], L. Abi-Rached[2], L. Tian[2], K.L. Gunderson[1], P.J. Norman[2], M. Ronaghi[1], P. Parham[2].* 1) Illumina, Inc., San Diego, CA; 2) Stanford University School of Medicine, Stanford, CA.

The HLA and KIR regions are two of the most polymorphic and structurally complex regions in the human genome. Because of the high functional diversity they impart to the human immune system, HLA and KIR are associated with susceptibility to numerous infectious diseases, autoimmunities and some pregnancy syndromes. HLA matching is required for most forms of allotransplant, and HLA genotyping is becoming increasingly important for predicting adverse drug reactions. To provide in-depth understanding of the HLA and KIR regions, we developed methodology for determining complete haplotype sequences across large population cohorts. This approach includes paired-end library preparation from genomic DNA, a hybridization-capture enrichment assay, and a dedicated alignment/assembly informatics pipeline. The protocol currently achieves >97% coverage of the 4.7 Mb HLA and 200 kb KIR regions. We have completed sequences for eight disease-associated HLA haplotypes previously investigated in the MHC Haplotype Project. We will present data on an additional 96 HLA haplotypes, obtained from homozygous cell lines, which represent a significant proportion of worldwide HLA diversity. This methodology will be beneficial for discovering disease mechanisms and disease diagnostics.

3711F

DNA Sudoku: Hunting Rare Genetic Variations Using Combinatorial Pooling. *D. Esposito[1], D. Golan[2], B. Blumenstiel[3], Y. Erlich[1].* 1) Whitehead Institute for Biomedical Research, Cambridge, MA; 2) Tel Aviv University, Tel Aviv, Israel; 3) Broad Institute, Cambridge, MA.

In the past few years, we have experienced a paradigm shift in human genetics. Accumulating lines of evidence have highlighted the pivotal role of rare genetic variations in a significant number of traits and diseases. Studying rare variations is a needle in a haystack problem, as large cohorts have to be sequenced in order to trap the variations and gain statistical power. Sample preparation techniques have not scaled with the exponential growth in high throughput sequencing technology. Current methods rely on individual barcoding, which requires tedious preparation of large numbers of samples. We are developing a scalable solution, called DNA Sudoku, to reduce preparation time and increase throughput. We use a combinatorial pooling approach, in which we multiplex specimens and then barcode and enrich the pools. Our pipeline substantially increases the feasibility and reduces the costs of large-scale experiments since the pools, rather than the individual specimens, are sequenced. To reduce the possibility of specimen dropout, the DNA samples are first quantified and calibrated so that equal amounts can be pooled and sequenced. We then determine the pooling strategy that best balances the number of specimens and pools while maintaining robust decodability, or the ability to correctly link a sequencing result to its specimen. We have rigorously tested the pipeline by targeted resequencing of samples from the HapMap Project, which we successfully decoded using publicly available 1000 Genomes data. After several successful pilot studies, we plan to employ the DNA Sudoku pipeline to uncover rare human genetic variants with the intent of increasing throughput and significantly reducing sample preparation costs. Our next application of DNA Sudoku will be to take a genetic snapshot of risk alleles for diseases in the Jewish population. Our aim is to utilize our pooling strategy to genotype a cohort of ethnically matched individuals to create a more effective carrier screening diagnostic tool.

3712W

From sample collection to bacterial identification in a single day: 16s rRNA sequencing using novel primers on the Ion Torrent Personal Genome Machine. *G.S. Watts[1], M.M. Oshiro[1], B.W. Futscher[1], D.G. Armstrong[2].* 1) The University of Arizona Cancer Center, University of Arizona, Tucson, AZ; 2) Southern Arizona Limb Salvage Alliance, University of Arizona, Tucson, AZ.

Identification of the bacteria present in clinical infections is important for treatment decision making, including proper antibiotic selection. Culturing samples on agar plates and biochemical assays have historically been used to identify bacteria, but these methods are inherently biased and can take several days. The time required to culture bacteria for identification can be a serious drawback in the setting of a clinical infection. An alternative to culture is offered by 16s rRNA gene sequencing to identify bacteria within samples. Sequencing-based approaches to bacterial identification do not have the inherent bias that culture-based methods possess, and can be more comprehensive. After a sample is collected, polymerase chain reaction is used to amplify a portion of the 16s rRNA gene which is then sequenced. Sequence generated from the amplicon is compared to a reference library to identify the bacteria that were present in the sample. Despite its advantages, 16s sequencing still took several days whether performed using Sanger sequencing or pyrosequencing. The Ion Torrent Personal Genome Machine (PGM) is a new sequencing technology that can rapidly sequence short DNA fragments in a cost effective manner. Current commercially available methods for 16s sequencing utilize ~500 base portions of the 16s rRNA gene for bacterial identification. While next generation sequencing platforms such as the PGM are currently unable to sequence amplicons of this length, recent studies have indicated shorter amplicons can perform well in bacterial identification. To this end, we have developed novel PCR primers to amplify a short amplicon containing the V6 hypervariable region of the 16s rRNA gene. We describe a PCR method that allows amplicon library generation in a bar-coded format directly from bacterial samples without template DNA isolation. The combination of our PCR primers and method with the PGM sequencing platform offers same day bacterial identification from clinical samples. We will present proof of principle of our method on pure bacterial strains as well its application to clinical samples.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3713F

Very high resolution HLA genotyping with the 454 Life Sciences GS FLX system: Simplification of workflow using fusion primers or a four primer system. *B.N. Hoglund[1], C.L. Holcomb[1], T.C. Williams[2], D. Goodridge[3], H.A. Erlich[1].* 1) Roche Molecular Systems Pleasanton, CA; 2) Children's Hospital Oakland Research Institute Oakland, CA; 3) Conexio Genomics Perth, Australia.

Purpose: The HLA genes are the most polymorphic in the human genome, making HLA typing challenging. HLA typing is important for stem cell transplantation and for disease association and population genetics studies. Most current typing methods provide data consistent with many different genotypes ("ambiguity"). We have previously developed very high resolution (VHR) genotyping of HLA using the 454 GS FLX Sequencing System and Conexio Assign ATF 454 genotyping software. The VHR assay was designed to resolve the majority of the thousands of HLA alleles as well as rare expression variants, i.e. the "null" alleles. Amplicon library construction was performed using 22 PCR fusion primers, containing genome target specific sequence contiguous with 454 sequencing adaptors. The VHR assay provided reliable and very high resolution (limited ambiguity) genotyping but required extensive handling of individual amplicons. To simplify the workflow, we prepared the amplicon libraries by three alternative procedures. Methods: PCR amplicons were generated for HLA-A/B exons 1–5; HLA-C exons 1–7; DPA1, DPB1, DQA1 exon 2; DQB1, DRB exons 2 and 3 from 10 cell lines using three different workflows. Amplification required 22 pairs of fusion primers (14 primer pairs in the GS GType HLA primers plates from 454, plus 8 amplicons in a third plate) with 10 Multiplex Identifiers (MIDs), OR 22 primer pairs each containing, at the 5'end, a universal sequence which matched the 3' end of a set of primers that contained 10 MIDs and the 454 sequencing adaptors. Amplicons derived from fusion primers were purified, quantified, diluted to a standard concentration, then pooled OR pooled prior to the other operations. Amplicons generated by the four primer system on the Fluidigm Access Array™ were handled by the latter procedure. Pools of amplicons were sequenced on the GS FLX and genotyped with Conexio software. Results: All three procedures resulted in very high resolution genotyping. Pooling amplicons immediately following the genomic PCR gave a dramatic reduction in the number of pipetting steps necessary to produce an amplicon library; ~80 and 90 percent reduction in pipetting was achieved using fusion primers and the four primer system, respectively. Conclusion: This 22 primer genotyping assay provides very high resolution HLA class I and class II typing. Pooling amplicons early in the 454 library preparation greatly simplifies the workflow of the VHR HLA genotyping assay.

3714W

Highly Multiplexed Amplicon Preparation for Targeted Re-Sequencing of Sample Limited Specimens using the Ion AmpliSeq™ Technology and Semiconductor Sequencing. *C. Li, B. Kong, D. Joun, I. Casuga, M. Shannon, S. Chen, M. Andersen, D. Ruff, R. Bennett.* Molecular Biology, Life Technologies, South San Francisco, CA.

Genomic and exome sequencing has presented researchers a grand view of germ-line and somatic genetic variations that occur in cells. The de-novo identification of cancer mutations or mapping for inherited genetic disorders relies on interrogation of selected regions in the genome across well-annotated specimens. Target enrichment for coding regions of genomic DNA has been proven to be a cost effective way for genetic variation studies without the need of sequencing the entire genome. Typical hybridization-based exon enrichment has extensive sample preparation time and is challenged by applications that require analysis of scarce and highly degraded DNA materials from sources such as single cell, fine needle biopsy or archival FFPE sample. Researchers need a rapid and flexible targeted re-sequencing process that is compatible with small amount of highly degraded DNA derived from a limited tissue source. Ion AmpliSeq™ technology combines the specificity, sensitivity, and efficiency of PCR technology to enrich tens to thousands of genomic targets from less than 10ng of FFPE or gDNA in a reaction for next generation sequencing (NGS). The unique amplicon preparation workflow sharply reduces side-products in multiplex PCR and thereby enables this technology to venture into extremely high levels of multiplex PCR for NGS. The workflow consists of target amplification using primer pools designed by Ion AmpliSeq™ Designer software. The amplicons (150bp ~ 200bp) are subsequently ligated to sequencing adapters and purified. The resulting DNA libraries can typically be completed in 3.5 hours with as little as 30 minutes of hands-on-time using the Ion AmpliSeq™ Library Kit 2.0. Ion AmpliSeq™ Comprehensive Cancer Panel and Inherited Disease Panel have primer designs of up to 4,000 amplicons in a single tube that target defined coding regions in the genome. Sequencing results consistently achieve >98% of all reads on target at >99.5% per base accuracy and uniform coverage of >94% of bases covered ≥20X when normalized to 100X average coverage. By monitoring amplicon PCR efficiency and normalized amplicon uniformity from sequencing, an early demonstration of 18,500-plex single tube reaction suggests the multiplex limit of the AmpliSeq™ technology has yet to be reached. The scalability opens up the possibility of rapid targeted re-sequencing of exons from 25,000 genes in a few reactions with only sub-microgram quantities of DNA. Data will be presented.

3715F

A Fast Solution to NGS Library Preparation with Low Nanogram DNA Input. *P. Liu, G. Lohman, E. Cantor, B.W. Langhorst, E. Yigit, L.M. Apone, D.B. Munafo, C. Sumner, F.J. Stewart, T.C. Evans, E.T. Dimalanta, T.B. Davis.* New England Biolabs, Ipswich, MA.

Next Generation Sequencing (NGS) has significantly impacted human genetics, enabling a comprehensive characterization of human genome as well as better understandings of many genomic abnormalities. By delivering massive DNA sequences at unprecedented speed and cost, NGS promises to make personalized medicine a reality in the foreseeable future. To date, library construction with clinical samples has been a challenge, primarily due to the limited quantities of sample DNA available. To overcome this challenge, we have developed a fast library preparation method using novel NEBNext reagents and adaptors, including a new DNA polymerase that has been optimized to minimize GC bias. This method enables library construction from an amount of DNA as low as 5 ng, and can be used for both intact and fragmented DNA. Moreover, the workflow is compatible with multiple NGS platforms.

3716W

Pre-Capture Pooling for Targeted Enrichment of Libraries for Next Generation Sequencing. *B. Marosy, B. Craig, A. Robinson, M. Zilka, K. Hetrick, S. Griffith, H. Ling, J. Romm, K.F. Doheny.* Center for Inherited Disease Research (CIDR), Johns Hopkins University, Baltimore, MD.

The Center for Inherited Disease Research (CIDR) provides high quality next-generation sequencing (NGS), genotyping and statistical genetics consultation to investigators working to discover genes that contribute to disease. CIDR has developed a NGS lab using the Illumina® HiSeqTM2000 platform in conjunction with the Agilent© SureSelect TM XT library prep and target enrichment kits. Agilent has recently added improvements to their library prep/target enrichment (XT2 reagent kit) including pre-mixed reagents and indexed adapters for ligation to allow multiple samples to be pooled together in a single well for the enrichment step. CIDR has added optimizations (`with Bead' clean-up and `off Bead catch' as published by the Broad Institute) to the protocol to reduce DNA input while maintaining library complexity. Eight unrelated HapMap samples were used to test the updated process using 500ng DNA input into the library prep. Samples were hybridized using a combined total of 1.5ug of constructed library across 8 samples (187.5ng/sample) with the SureSelect Human AllExon v4 + UTRs (71Mb) for 24hrs. These 8 samples were then clustered for sequencing across two lanes of the HiSeq2000 using the Illumina cBOT TM Cluster Generation system. Seventy-five bp paired-end sequencing was performed on the HiSeq2000 which generated on average 5.7 Gb of data (unique passing filter aligned) per sample. Preliminary data analysis was performed on 6 out of 8 samples and yielded an estimated library size of 314 million with a duplication rate of 6.4%. Mean percent of on target and on bait bases covered to at least 8x was 95.4% and 94.6%, respectively. The mean transition/transversion ratio on exon for these samples was 3.11. Sensitivity and concordance to previous genotypes was 96.0% and 99.8% respectively.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3717F

Rapid and accurate semiconductor-based sequencing of human exomes: workflow and performance on a familial trio. *G. Meredith, G. Bee, L. Pickle, M. Dudas, G. Del Mistro, C. Scafe, M. Schorn, J. Miller, M. Minto, B. Reed, G. Fry, J. Gioia, J. Hildebrandt, P. Leong, M. Reddy, K. Atehortua-Khalsa, M. Sedova, S. McLaughlin, V. Sheth, M. Shah, D. Thomas, H. Breu, M. Rhodes, C. Adams, F. Hyland, R. Bennett.* Life Technologies Corp / Ion Torrent, 5791 Van Allen Way, Carlsbad, CA 92008.

Though it is now possible to sequence a human genome in a single day, the need for high sample throughput and reduction in informatics necessitates exome enrichment in many cases. We will discuss three key areas of the human exome workflow, design of the exome enrichment panel, rapid sequencing turnaround, and assessment of the results. We have designed and built a set of efficient hybridization probes that permit the enrichment of protein-coding exons from 6 major gene-model databases (GenCode, RefSeq, Ensembl, UCSC, VEGA, and CCDS), functional RNA genes from miRBase and UCSC databases, predicted micro-RNA binding sites, and COSMIC cancer variants. In total, this set of probes is optimized to enrich ~46 Mb of target regions with probe sequences that densely tile ~52 Mb of the human genome and the workflow yields an on-target read mapping rate of 80%. Following exome-enriched library preparation, templated particles are prepared using an automated device (OneTouch™) or manually and sequenced in one (Proton™) or five (PGM™) instrument runs, with each run taking ~4 hours. Such rapid exome sequencing in a single Ion Torrent Proton™ System run or 5 runs on Ion Torrent PGM™ Systems yielded >5 Gb of aligned base reads and provided an average depth over targets of >50x with >85% of target bases covered at a 20x depth or more. Rapid read mapping, determination of exome coverage metrics, and variant calling is done seamlessly using the Torrent Suite™ Software that is provided with these instrument systems. Annotation of variants was carried out using Ion Reporter™ Software. To validate the performance of this workflow, exome-enriched fragment libraries were sequenced from a familial trio (parents NA12891, NA12892, and child NA12878) from CEPH/Utah Pedigree 1463. Each individual exome dataset, yielded more than 30,000 SNP calls (>18,000 homo and >12,000 het) and more than 98% of all SNPs called exist in dbSNP135. Each exome dataset also permitted the calling of hundreds of indels. Comparison to whole-genome data from the 1000 Genomes project and Complete Genomics along with Mendelian inheritance analysis permitted estimation of the specificity, sensitivity and accuracy of variant calling. The combination of focused exome enrichment and Ion Torrent Systems-based sequencing and analysis provides an efficient, accurate, and rapid means to detect genetic variation in the well-annotated portion of the genome for state-of-the-art genetic disease research.

3718W

TotalScript: A versatile and robust tool for RNAseq applications. *F. Syed, S. Kuersten, A. Radek, R. Vaidyanathan.* Illumina, Madison, WI.

Deep sequencing can provide a digital readout of gene expression and can be used for a variety of distinct applications in human biology and genetics including allele-specific expression profiling, alternative splicing regulation and promoter usage in disease states, non-coding RNA analysis and many others. The diverse and complex world of RNAseq and the rapid development of sequencing technology require the development of simple, versatile tools that competently reflect RNA levels in the cell. We have developed a novel, transposome-based stranded RNAseq method capable of using variety of RNA sample types and a range of input levels to produce high quality RNAseq libraries for illumina-based sequencing. Samples made from this method produce libraries with high complexity, ~97–99 % stranded and low rRNA contamination. The method is scalable (automation) and robust, allowing input amounts down to ~100 pg of polyA+ RNA or ~5ng of total RNA. Furthermore, due to the nature of the technology, input RNA type is quite flexible; polyA+ selected, Ribo-Zero treated or Total RNA samples can be used and both random priming and oligo-dT priming of first strand cDNA synthesis are compatible. The method provides a simple, innovative and integrated solution to RNAseq applications.

3719F

Automating high-throughput creation of sequencing libraries. *J. Bishop, W. Zhang, D. Mandelman, M. Allen, A. Harris, R. Bennett.* Life Technologies, Carlsbad, CA.

Obtaining the best results from high throughput sequencing studies often depends on the ability to produce high quality sequencing libraries, yet creating these libraries is a repetitive, labor-intensive, error-prone process with often variable results. To achieve greater throughput and consistency, we have developed automated protocols for creating DNA and transcriptome libraries for use in semi-conductor sequencing. To create DNA fragment libraries, our protocols automate shearing, end-repair, adaptor ligation, and library size-selection from as little as 50 ng of input DNA. User-control options also allow customization of the insert-sizes produced during automated shearing and size selection. The resulting DNA libraries are of sufficient yield and quality to be used directly in downstream steps without requiring amplification or additional manual manipulation. To create strand-specific transcriptome libraries, our protocols automate adaptor ligation, reverse transcription, and library size selection from as little as 50 ng of fragmented RNA. Semiconductor sequencing demonstrates that both fragment and transcriptome libraries produced by these automated protocols are complete and unbiased relative to manually-constructed libraries. The technologies described here provide automation solutions for increasing the speed, throughput, consistency, and efficiency of semiconductor sequencing applications.

3720W

Highest sample quality for molecular analysis through ambient stabilization technologies. Improved genome, transcriptome and proteome analysis from saliva, blood and tissue samples. *R. Muller, V. Liberal, S. Wilkinson, A. Stassinopoulos, J. Muller-Cohn.* Biomatrica, San Diego, CA.

Biospecimen stabilization during collection and transport is of critical importance in the fields of biomedical research and molecular diagnostics. The emergence of personalized medicine has placed an even greater emphasis on the significance of protecting the quality and reliability of clinical samples. The results obtained during biomarker discovery is directly linked to potential disease diagnostics and therapeutic intervention. Current methods for blood and saliva collection depend heavily on cold-chain logistics. The reliance on cold and ultra-cold based sample preservation is fraught with potential complications to specimen integrity that can result in inconsistent result generation impacting the discovery of suitable and reliable biomarker and their transition to clinical applications. Chemical stabilization of biomolecules at ambient temperatures is far superior in preserving the sample quality and integrity over conventional low temperature based stabilization methods. In this report we describe the development of chemical stabilization technology that addresses the problem of sampling induced inconsistencies in human blood and saliva. The formulations described preserve nucleic acids in these biospecimens even under extreme environmental shipping conditions and remove the need for cold-chain logistics during sample collection, transport or storage. Here we show that the genomic DNA in blood collected in DNAgard Blood Tubes is stabilized during extreme temperature fluctuations ranging from -20°C to +45°C. Samples remain stable for up to 14 months at room temperature. The analogous DNA stabilization technology for saliva specimens, DNAgard Saliva, preserves sample integrity for over 12 months at ambient temperature. We also compare freezer storage with the room temperature stabilization formulation RNAgard Blood and other technologies for ambient temperature blood RNA and protein stabilization. The presented results are demonstrating that transcriptome and protein based analysis is vastly improved through ambient biostabilization. This data is relevant for large scale cohort studies as well as moving biomarkers to the clinic.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3721F

Increased sensitivity in whole-genome bisulfite sequencing (WGBS): A novel "post-bisulfite conversion" library construction method for sub-nanogram inputs. *R. Sooknanan[1], A. Adey[2], J. Hitchen[1], J. Shendure[2], N. Caruccio[1].* 1) Epicentre® (an Illumina® company), Madison, WI; 2) Department of Genome Sciences, University of Washington, Seattle, WA, USA.

Genome-wide analysis of 5-methylcytosines is possible with whole-genome bisulfite sequencing (WGBS), where unmethylated cytosine residues are converted to uracil. However, a major challenge in WGBS is the degradation of DNA that occurs during bisulfite conversion under conditions required for complete conversion. Typically, ~90% of input DNA is degraded and thus, is especially problematic with limited starting amounts of DNA. Additionally, regions that are rich in unmethylated cytosines are more sensitive to strand breaks. As a consequence, a majority of DNA fragments contained in di-tagged NGS DNA libraries treated with bisulfite "post-library construction" can be rendered inactive due to strand breaks in the DNA sequence flanked by the adapter sequences. These mono-tagged templates are then excluded during library enrichment resulting in incomplete coverage and bias when performing whole genome bisulfite sequencing. Here, we describe a novel "post-bisulfite conversion" library construction method for preparing NGS libraries from genomic DNA prior to the addition of one or both adapters. This "post-bisulfite conversion" library construction method uses the resulting untagged or mono-tagged single-stranded DNA as template for the subsequent addition of adapter sequences required for NGS. Thus, single-stranded DNA fragments independent of size and position of strand breaks remain as viable templates for library construction, eliminating the loss of fragments and the selection bias associated with a "post-library construction" bisulfite conversion strategy. This novel "post-bisulfite conversion" library construction method exhibits increased sensitivity and efficiency (500 picogram input), and improved coverage required for WGBS.

3722W

Highly efficient miRNA isolation method using Solid Phase Reverse Immobilization (SPRI) Technology and Biomek automation. *BN. Lee.* Beckman Coulter 250S Kramer Blvd., Brea, CA 92822.

miRNA is a short ribonucleic acid with the average size of 22 nucleotides and is found abundantly in many human cell types as well as in other organisms. Over 1000 unique miRNA have been identified in human genome and it is reported that miRNA may target and regulate over 60% of human genes in cellular responses. In human cells, miRNA complementary binds to the 3'UTR of mRNA and binding of the miRNA plays a role in gene silencing, transcriptional/translational regulation in cellular responses, developmental processes, signal transduction and cancer or disease control. In the past two years biomedical research has placed much importance on characterizing miRNA to further understand its biological function and how it can be used as a biomarker for cancer development, and be utilized as therapeutic targets for cancer treatment. Sample preparation for this process can be very challenging. It often requires isolation of high quality miRNA from FFPE samples, cell culture and tissues. Most commercial nucleic acid extraction kits and protocols are inefficient in recovering small molecules such as miRNA. We describe efficient and robust miRNA extraction methods using Agencourt SPRI* (Solid Phase Reverse Immobilization) technology and Biomek automation*. The two methodologies described are (1) miRNA extracted and purified from FFPE samples using the Agencourt FormaPure Kit* and (2) miRNA extracted and purified from cell culture using the Agencourt RNAdvance Cell v2 kit*. The TaqMan** qPCR expression of let-7c demonstrates successfully purified miRNA from FFPE and cell culture samples. * Biomek/AgencourtFormaPure/Agencourt RNAdvance Cell v2 kits are registered trademarks of Beckman Coulter, Inc. ** TaqMan qPCR is registered Trademark of Roche Molecular System, Inc.

3723F

MiRNA extraction from human plasma using high capacity miRNA Anti-Probe Magnetic beads. *A. Zampetaki[1], T. Xu[2], A.V. Vlassov[3], N. Bernard[4], M. Mayr[1].* 1) Cardiovascular Division, King's College London, LONDON, United Kingdom; 2) Life Technologies, 850 Lincoln Centre Drive Foster City, California, 94404, USA; 3) Life Technologies, 2130 Woodward st, Austin TX 78739, USA; 4) Life Technologies Europe B.V. - Belgisch bijkantoor Regus Gent Zuiderpoort Gaston Crommenlaan 4 bus 0501 9050 GENT Belgium.

Background: Circulating microRNAs (miRNAs) have emerged as promising biomarkers. Objective: In this study, we assessed the efficiency of the miRNA extraction using Anti-probe Magnetic Beads or the conventional extraction technology based on phenol/ guanidine sample lysis and glass-fiber binding purification of RNA. Methods and Results: MiRNA Anti-probe Magnetic Beads were used to capture the miRNA pool from 50ul of human plasma (n=10). Beads Panel A that targets a total of 377 human miRNAs that are based on the Megaplex Pool A and 6 non-human positive and negative controls. Following the initial sample lysis, anti-probe beads were hybridized to the specific miRNA targets. The extracted miRNA was reverse transcribed using the Taqman MicroRNA RT kit and the Megaplex Pool A with RT primers. In parallel, RNA was extracted using the conventional method of extraction using 100ul from the same human plasma samples. In both protocols, the exogenous cel-miR-39 was spiked-in during extraction and served as a normalization control. MiRNA expression was validated by QPCR using the Taqman miRNA assays. In all samples an extremely high association was detected between the two extraction protocols (r= 0.856), with several miRNA targets reaching almost perfect association (e.g. miR-92a, r=0.992). Conclusions: Our study demonstrates that the miRNA Anti-probe Magnetic beads can offer high sensitivity and reproducibility in miRNA extract using a short and simple workflow with minimal biological sample. These findings suggest that miRNA Anti-probe Beads could be used to establish an automated protocol to facilitate the screening of miRNAs in large cohorts.

3724W

Single Molecule Real-Time (SMRT®)Sequencing of genes implicated in autosomal recessive diseases. *Y. Guo[1], A. Bashir[2], E. Gould[2], R. Kornreich[2], L. Edelmann[2], T. Brandt[2], Y. Kasai[1], J. Chin[2], E. Paxinos[1], A. Kasarskis[2], E. Schadt[2].* 1) Pacific Biosciences, Menlo Park, CA., USA; 2) Mount Sinai School of Medicine, Dept. of Genetics and Genomic Sciences, New York, NY USA.

In today's clinical diagnostic laboratories, the detection of the disease causing mutations is either done through genotyping or Sanger sequencing. Whether done singly or in a multiplex assay, genotyping works only if the exact molecular change is known. Sanger sequencing is the gold standard method that captures both known and novel molecular changes in the disease gene of interest. Most clinical Sanger sequencing assays involve PCR-amplifying the coding sequences of the disease target gene followed by bi-directional sequencing of the amplified products. Therefore for every patient sample, one generates multiple amplicons singly and each amplicon leads to two separate sequencing reactions. Single Molecule, Real-Time (SMRT®) sequencing offers several advantages to Sanger sequencing including long read lengths, first-in-first-out processing, fast time to result, high-levels of multiplexing and substantially reduced costs. For our first proof-of-concept experiment, we queried 3 known disease-associated mutations in de-identified clinical samples. We started off with 3 autosomal recessive diseases found at an increased frequency in the Ashkenazi Jewish population: Tay Sachs disease, Niemann-Pick disease and Canavan disease. The mutated gene in Tays Sachs is HEXA, Niemann-Pick is SMPD1 and Canavan is ASPA. Coding exons were amplified in multiple (6–13) amplicons for each gene from both non-carrier and carriers. Amplicons were purified, concentrations normalized, and combined prior to SMRTbell™ Library prep. A single SMRTbell library was sequenced for each gene from each patient using standard Pacific Biosciences® C2 chemistry and protocols. Average read lengths of 4,000 bp across samples allowed for high-quality Circular Consensus Sequences (CCS) across all amplicons (all less than 1 kb). This high quality CCS data permitted the clean partitioning of reads from a patient in the presence of heterozygous events. Using non-carrier sequencing as a control, we were able to correctly identify the known events in carrier genes. This suggests the potential utility of SMRT sequencing in a clinical setting, enabling a cost-effective method of replacing targeted mutation detection with sequencing of the entire gene.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3725F

Detection of Rare Somatic Mutations Using a Simplified, Specific digital PCR Workflow with Zero Dead Volume. *M.C. Pallas, D. Keys, J. Wilde, J.C. Nurse.* qPCR R&D, Life Technologies, Foster City, CA.

The detection and enumeration of low frequency somatic mutations is important to continued advances in the understanding of cancer. Conventional PCR is limited in its sensitivity to the rare mutants due to large background of the wild type. Digital PCR utilizes PCR technology to precisely count the number of nucleic acid templates. Its basis is the dilution of the nucleic acid template to a point where each PCR reaction vessel contains a single copy or no copies. The single copies amplify while no amplification is detected in the vessel lacking DNA template. Using a large number of reaction vessels with the assumption of Poisson distribution of copies, the average number of copies per vessel can then be calculated. This results in a highly precise and specific method of detecting and quantifying DNA. We have developed a simple, three step, digital PCR workflow that allows for highly specific detection of rare somatic mutations with zero dead volume. It is difficult to achieve high sensitivity mutation detection with conventional qPCR because the high concentration of non-mutant template outcompetes the target for amplification reagents, masking the mutant signal. In dPCR, individual reaction vessels containing mutant DNA strands have no, or very low, numbers of non-mutant template. This allows the signal from the mutant to emerge, and the overall frequency to be precisely determined. Our digital PCR workflow consists of three steps, loading the sample, thermal cycling, and detection of the digital results. Operator interaction with the sample exists in only the first step, loading the sample. We demonstrate the utility of this method using well characterized cell lines containing six KRAS mutations. DNA from these lines was titrated with wild-type DNA to give mutant concentrations as low as 0.025%. dPCR runs on these samples, as well as mutant-free control samples, demonstrate high sensitivity, specificity, and precision of this simplified workflow.

3726W

Rapidly profiling thousands of large non-coding RNAs and mRNAs from nanogram amounts of total RNA using a single microarray design. *A. Bergstrom Lucas, K. Swaminathan, V. Kulkarni, E. LeProust, S. Fulmer-Smentek.* Genomics R&D, Agilent Technologies, Santa Clara, CA.

Large intergenic non-coding RNAs (lincRNAs) are emerging as key regulators of diverse cellular processes. As researchers face the challenge of investigating the function of lincRNAs, there is a need for tools that can accurately and rapidly measure the expression of both lincRNAs and mRNAs in the same assay. In 2011, a reference catalog of more than 8,000 human lincRNAs was annotated from 4 billion RNA-Seq reads across 24 tissues and cell types by scientists at the Broad Institute of MIT and Harvard. Using this new reference catalog of lincRNAs we have updated the content of the SurePrint G3 Human microarrays to enable systematic profiling and simultaneous detection of coding and non-coding gene expression from a single RNA sample. We used low nanogram amounts of total RNA from matched tumor and adjacent normal tissues to detect both large and subtle differences in gene expression profiles that were consistent with the current literature. GeneSpring GX software rapidly identified differentially expressed lincRNAs and protein-coding RNAs resulting in expression measurements from total RNA in less than two days. The SurePrint G3 Human Gene Expression v2 microarrays provided data that demonstrated good reproducibility, wide dynamic ranges and high sensitivity. The tumor versus normal ratios generated from the v2 microarrays showed high correlation with ratios generated from whole transcriptome sequencing from the same matched RNA samples. In this study we demonstrate that the new v2 microarrays can provide accurate differential expression measurements of both protein coding and non coding RNA from very low amounts of total RNA, rapidly providing biologically relevant gene expression measurements that are equivalent to whole transcriptome sequencing data.

3727F

Development of a system based on SMART technology for robust transcriptome library preparation from small quantities of degraded sample. *M. Bostick, C. Chang, A. Farmer.* Clontech Laboratories, Inc, Mountain View, CA., USA.

By providing sequence data on millions of short DNA fragments in parallel, Next Generation Sequencing has revolutionized biomedical research. In particular, the technique has enabled RNA expression analysis over the entire transcriptome with superlative sensitivity and dynamic range. With this has come a drive to utilize smaller and smaller sample inputs - with the ultimate goal of analyzing the transcriptome of a single cell. Achieving this requires highly sensitive and robust sample preparation methods. One powerful method for cDNA preparation is Clontech's proprietary SMART technology (Switch Mechanism At the 5' end of the RNA Template), which utilizes the template switching activity of reverse transcriptase to enable the direct addition of a PCR adaptor to the 3' end of the first-strand cDNA; thus avoiding inefficient ligation steps. The result is a single-tube protocol that greatly enhances library amplification efficiency as compared to traditional technologies, while minimizing the chance for contamination or degradation of the sample, making it ideal for library preparation from small amounts of starting material. Indeed, the currently available SMART kit allows researchers to readily obtain high quality data from as little as 100 pg of total RNA. Although the SMART system provides many benefits, one drawback of the current system is the inability to work with samples comprising degraded RNA, owing to the method's dependence on an oligo dT primer for first-strand synthesis. In many samples where input amount is low - e.g., FFPE specimens - the starting RNA is often degraded; limiting the applicability of such poly A based strategies. Accordingly, we have adapted the SMART protocol to work with degraded starting material - even with starting amounts as low as 1ng total RNA. Data will be presented out-lining various optimization steps, including the use of random primers of varying lengths, as well as methods for both purifying RNA from FFPE sections and depleting rRNA. In addition, the use of novel polymerase formulations to further enhance yield in the amplification step has also been investigated. These newer formulations are compared to other currently available enzymes with respect to total DNA yield, mutation rate and sample bias. Based on these modifications, we have developed a system that is capable of generating cDNA libraries for transcriptome profiling from FFPE samples starting with as little as 1 ng of total RNA.

3728W

Selective sequencing of mature transcripts in the human brain by cytoplasmic RNA-Seq. *A. Zaghlool, A. Ameur, J. Halvardson, L. Cavelier, L. Feuk.* Dept Immunology Genatic & Patholohy, Science for Life Laboratory Uppsala, Rudbeck Laboratory, Uppsala University, Uppsala, Sweden.

The recent data from total RNA and polyA RNA sequencing experiments have significantly increased our understanding of the human transcriptome and led to identification of previously unknown exons and splice isoforms. However, previous studies have also indicated that a large fraction of sequencing reads align to introns and intergenic regions, which may lead to a reduced efficiency of these methods to detect novel exons and new splice junctions. To provide deeper insight into the dynamics of the transcriptome, and to avoid the biases associated with polyA selection, we developed a technique to extract pure RNA populations from cytosolic and nuclear fractions and sequenced these separately. We show that our technique generates RNA populations with substantially lower cross contamination as compared to commercially available kits. We then sequenced cytosolic, nuclear, polyA selected and total RNA from human frontal cortex using SOLiD5500 to investigate the sequencing profiles of these fractions. The results show that cytoplasmic RNA is highly enriched for mature transcripts with low levels of intronic reads, as compared to polyA and total RNA-seq. The exonic/intronic ratio of the reads in the CytoRNA-Seq was 0.74, while the corresponding ratios were 0.20 and 0.27 for total RNA-Seq and polyARNA-Seq respectively. The NucRNA-seq data showed the lowest ratio of 0.13. Our results also show that CytoRNA-Seq yields substantially larger amounts of splice junction reads compared to the NucRNA-Seq. We propose that our CytoRNA-Seq offers a new and more accurate method to selectively sequence and measure the levels of mature transcripts compared to conventional RNA-seq approaches.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3729F

Targeted library preparation for Ion Torrent sequencing using HaloPlex PCR. *H. Johansson, E. Agne, K. Zettermann, M. Isaksson, P. Eriksson, B. Skarpås, F. Roos, F. Dahl.* LSG Genomics, Agilent, Uppsala, Sweden.

We have developed a method for multiplex target capture and library preparation of genomic regions. The protocol can target up to 250 Kbp regions, requires 200 ng of gDNA and produces ready-to-load PGM sequencing libraries in less than six hours. Benchtop sequencing instruments such as the Ion Torrent PGM sequencer is capable of generating several gigabases of sequence data in less than 24h. This capacity allows for comprehensive studies of genetic material to be carried out within a few days. However, since the entire human genome is too large to be sequenced on this platform, without spending an unreasonable amount of time and sequencing reagents, it is often advantageous to enrich for regions of interest prior to sequencing. Methods used for genetic research studies as well as for clinical applications need to meet a few important criteria. The coverage of targeted regions should be close 100 percent and capture efficiency of different regions should be uniform to preserve initial allele representation. High specificity is important to direct sequencing capacity to regions of interest without discriminating against novel variants. Furthermore it is also useful if the protocol is fast to make use of the short turnaround time of the PGM. The HaloPlex protocol consists of four main steps. In the first step the genomic DNA is fragmented using restriction enzymes followed by the second step where HaloPlex probes are hybridized to targeted fragments. During hybridization, both strands of the restriction fragment are targeted to achieve bidirectional sequencing and thereby improve coverage. The third step involves DNA ligation of a common primer motif to all targeted fragments which is then used in the final step for multiplex amplification all targeted fragments using one single primer pair. Samples are barcoded during ligation and can thereby be pooled to optimize use of the sequencing capacity on the sequencing chip. The protocol can be applied for enrichment of exonic or intronic regions that are either continuous or interspersed. To demonstrate the performance of the new protocol we have enriched three different target regions (25Kb, 60 Kb and 250 Kb) on genomic sample material and sequenced them on the PGM. More than 95% of targeted bases were covered for all designs and all samples were enriched with >80% specificity. Furthermore, over 90% of the targeted bases were covered at >10% of average depth.

3730W

SNP Genotyping using the Affymetrix® Axiom® 2.0 platform. *M. Shapero, M. Purdy, H. Dong, S. Hsiung, R. Kurapati, J. Law, H. Lee, H. Loi, D. Nguyen, P.H. Wang, A. Yan, C.S. Yu, M. Shirazi.* Affymetrix, Inc. Santa Clara, CA.

The use of high-density DNA microarrays for accurate, cost-effective genotyping of SNPs and in/dels continues to play an important role in the identification of the genetic basis of common and complex human diseases through genome-wide association studies (GWAS). As next-generation sequencing expands the catalog of both common and rare variants in multiple worldwide populations, microarrays are well positioned to capitalize on this evolving information content. Here an overview of recent advances related to the biochemistry and lab automation of the Axiom® 2.0 platform is presented.

The Axiom 2.0 platform had previously validated the use of multiple sources of human genomic DNA (gDNA) including blood, saliva, whole-genome amplified DNA, and DNA from cell lines. At an input mass of 100 ng to 200 ng per sample per array, performance consistently achieves an average sample call rate ≥99.0%, average sample concordance to independent DNA genotype information (HapMap) ≥99.5%; and intra- and inter-run reproducibility ≥99.8%. The Axiom® Assay performance profile has continued to be evaluated relative to several new human sample types including FFPE gDNA, mouthwash-derived gDNA, and ultra-low gDNA input mass. Additionally, a protocol is described in which amplified DNA target can be hybridized across multiple arrays.

The Axiom 2.0 platform enables complete automation of DNA target preparation, DNA amplification, and enzymatic fragmentation of post-amplification products on the Beckman Coulter Biomek® FX^P Target Prep Express platform. Here we describe an alternate liquid handling workstation as well as development of a new array format which offers the capability to dramatically increase the weekly throughput of samples while retaining full compatibility with the existing Axiom® Instrumentation platform and all downstream data analysis.

In summary, new Axiom 2.0 additions to sample compatibility, lab automation, and array formats further extend the platform's capabilities to genotype hundreds of thousands of SNPs in a single assay, offering a sample throughput coupled with minimal manual intervention that is consistent with the needs of both large-scale GWAS and targeted studies being conducted as part of the search for the underlying genetic basis of complex human disease.

3731F

Information Tracking Through the Use and Customization of the Exemplar LIMS in a Medium Scale Sequencing Laboratory. *B. Craig, D. Newcomer, S. Griffith, B. Marosy, A. Robinson, M. Zilka, KF. Doheny.* Center for Inherited Disease Research (CIDR), Johns Hopkins University, Baltimore, MD.

The Center for Inherited Disease Research (CIDR) provides high quality next-gen sequencing (NGS), genotyping and statistical genetics consultation to investigators working to discover genes that contribute to disease. The rapidly changing sequencing landscape brings the complicated task of how to quickly adapt existing methods of workflow information tracking. CIDR's in-house developed genotyping Laboratory Information Management System (LIMS) handles the tracking of detailed information for large numbers of samples going through the Illumina® genotyping assays. When changes to genotyping protocols are required, software developers implement them by modifying validation criteria, adding or removing elements for tracking or creation of a new LIMS module to accommodate new genotyping services. A sequencing LIMS was required that allowed for even faster development of modules for different sequencing protocols and technologies. For this challenge, CIDR evaluated the Exemplar LIMS© from Sapio Sciences. The Exemplar LIMS allows for the creation of user specific data types and incorporation of these elements into tasks and workflows. These workflows create sequential tasks that record information about reagent lot numbers, robotics used and timestamp task completion. Each of these tasks can be made to enforce predefined order to ensure that tasks are not skipped in the lab. Exemplar LIMS allows for some native validations and reporting elements but also has a robust Java application programming interface (API) that gives the ability to create custom plugins to extend CIDR specific validations and tasks. With the Exemplar LIMS, a workflow can be created to integrate with a new sequencing wet bench protocol in as little as one day. CIDR has created workflows to track NGS library and target capture preparation for Agilent© SureSelect™ XT and XT2 methods and clustering and sequencing preparation using the Illumina cBot™ and HiSeq™ 2000 systems. A framework using the API has been created to quickly add plugins to validate user input, update values in the Exemplar database, upload QC data in bulk for experiments and integrate with existing sample information databases. Current work focuses on creation of workflows for sequencing on the Illumina MiSeq™ and Ion Torrent© PGM™, preparation of libraries for whole genome sequencing and target capture preparation for Roche Nimblegen© SeqCap EZ.

3732W

Locus Reference Genomic sequences: reference sequences for the reporting of clinically relevant sequence variants. *J.A.L. MacArthur[1], A. Astashyn[2], E. Birney[1], R. Dalgleish[3], P. Flicek[1], L. Gil[1], P. Larsson[1], D.R. Maglott[2], W.M. McLaren[1], R.E. Tully[2], F. Cunningham[1].* 1) European Bioinformatics Institute, Cambridge, UK; 2) National Center for Biotechnology Information, Bethesda, MD, USA; 3) Department of Genetics, University of Leicester, Leicester, UK.

Locus reference genomic (LRG) sequences are internationally recognized reference-sequences designed specifically for the reporting of clinically relevant sequence variants. An LRG provides a stable genomic DNA sequence for a region of the human genome which will never be changed and so is not versioned. They are created by the NCBI (http://www.ncbi.nlm.nih.gov) and EBI (http://www.ebi.ac.uk/), as part of the GEN2PHEN (http://www.gen-2phen.org) and RefSeqGene (http://www.ncbi.nlm.nih.gov/refseq/rsg) projects, in collaboration with the community of diagnostic and research labs, locus specific database curators and mutation consortia. The LRG project aims to create an LRG reference sequence for every genomic locus with clinical implications. Currently over 350 LRGs have been created, with this number expected to increase significantly by the end of this year. LRGs have been created for a panel of genes involved in inherited cardiac disease, in collaboration with the British Heart Foundation. LRGs are also in preparation for genes involved in cancer predisposition, including BRCA1 and BRCA2. All information relating to an LRG is contained within a single XML file, which comprises separate sections for the fixed and updatable annotation. The fixed (stable) section contains the genomic sequence and annotation for exons and coding regions of transcripts used as reference standards. The updatable section contains chromosome and coordinate mapping information for the LRG sequence, as well as annotation of additional cDNAs, co-located genes, variation, database cross-references and alternative (legacy) exon and amino-acid numbering systems. The stable nature of these reference sequences enables unambiguous reporting of variants in LRG coordinates, as recommended by the Human Genome Variation Society (HGVS) and European Molecular Genetics Quality Network (EMQN). Variant reporting in accordance with HGVS conventions using the LRG as the reference standard is possible in genomic DNA (e.g. LRG_1:g.8463G>C), coding DNA (e.g. LRG_1t1:c.572G>C) or protein coordinates (e.g. LRG_1p1:p.Gly191Ala). LRGs can be viewed in the Ensembl and NCBI genome browsers, while tools such as Mutalyzer (https://mutalyzer.nl) also support LRGs. All LRG records are available on the LRG website (http://www.lrg-sequence.org), which also provides instructions on requesting LRGs and the complete LRG specification.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.

3733F

QC Measures for Whole Exome and Targeted Sequencing Library Prep and Enrichment in a High Throughput Lab. *A. Robinson, M. Zilka, B. Marosy, B. Craig, J. Romm, K. Doheny.* Center for Inherited Disease Research (CIDR), Johns Hopkins University, Baltimore, MD.

The Center for Inherited Disease Research (CIDR) provides high quality next-gen sequencing (NGS), genotyping and statistical genetics consultation to investigators working to discover genes that contribute to disease. CIDR uses Agilent© SureSelect™ XT library prep kits for whole exome sequencing and custom target enrichment. Quality control measures (QC) are evaluated after shearing (QC1), prior to hybridization (QC2), after post hyb amplification (QC3), and after pooling, prior to clustering (QC4) using the Agilent© Bioanalyzer™, the PerkinElmer LabChip®GX and the KAPA SYBR® FAST qPCR kits. CIDR utilizes separate lab spaces for the workflow to maintain pre/post PCR separation between the two amplifications. Since CIDR physically separates its lab space and enforces a strict flow of samples, separate sample aliquots and/or dilution plates must be made to utilize the BioAnalyzer and LabChip GX. The first QC is performed after sample shearing to ensure samples have been sheared to the necessary fragment size. QC2 evaluates adapter ligation and concentration and is used to normalize samples for hybridization. QC3 is done after capture to check size and molarity. Samples are then normalized using the PerkinElmer Janus® to 1–2nM and retested to confirm the concentration before pooling. Once the pools are created (between 3–96 samples), each pool is tested (QC4) to ensure the samples do not become over-clustered. Pool balance is important, such that one sample is not over-sequenced in comparison to the other pooled samples. A final QC is done on the Illumina® MiSeq™ to ensure the evenness of the pools and check optimal cluster density before finally being clustered on the Illumina® cBOT Cluster Generation System™ and set up on Illumina® HiSeq 2000™. While QC tasks can add time to the lab processing workflow, ensuring only high quality libraries are sequenced will save time and money in the long run. A LabChip GX chip can run up to 96 samples at a time which takes ~3 hours and costs about $125 per chip, or $1.30/sample. A Bioanalyzer chip can run up to 12 samples at a time which takes 45 minutes and costs $19–$32, or $1.58–$3/sample (depending on chip type). The Bioanalyzer is more cost and time effective for smaller sample sets. The MiSeq run takes another 12–24 hours and costs about $695 for each run. Overall, this can add 1–2 days to the lab process but can save up to 14 days and >$6K in sequencing time and reagents.

3734W

Additional annotation enhances biologically relevant sub classification of the Illumina HumanMethylation450 BeadChip array. *E. Price[1,2,3], A. Cotton[3], L. Lam[2,3,4], E. Emberly[5], C. Brown[3], W. Robinson[2,3], M. Kobor[2,3,4].* 1) Obstetrics & Gynaecology, University of British Columbia, Vancouver, British Columbia, Canada; 2) Child and Family Research Institute, Vancouver, British Columbia, Canada; 3) Medical Genetics, Vancouver, British Columbia, Canada; 4) Centre for Molecular Medicine and Therapeutics, Vancouver, British Columbia, Canada; 5) Department of Physics, Simon Fraser University, Burnaby, British Columbia, Canada.

Over the past six years, Illumina has been a leading source for DNA methylation (DNAm) microarrays. To facilitate biologically significant data interpretation, we set out to enhance probe annotation of the newest HumanMethylation450 array (with >450,000 probes covering 99% of RefGenes). Annotation that was expanded on includes a) CpG island and b) gene feature classification; while annotation that was added includes c) SNPs in target CpGs and d) cross-hybridizing probes. a) Probes were annotated into four CpG island groups based on Weber's 2007 definition: high density CpG islands (HC; n=142,474); intermediate density CpG islands (IC; n=110,225); IC shores (i.e. ICs adjacent to HCs; n=31,724) and non-islands (n=201,059). The majority of HC probes were in Illumina-annotated CpG islands, while only 18% of Illumina-annotated shores were in IC shores. Shores may be important regions for changes in DNAm and a comparison to Weber's islands may elucidate shores with differing biological roles. We assessed tissue differences in DNAm ($\Delta\beta$ >0.20) using human buccal (n=4), blood (n=4) and placental (n=4) samples. Within Illumina-annotated shores, tissue differences in DNAm were noted in 20% of IC probes and 18% of IC shore probes but only 8% of HC probes. b) To aid in understanding how DNAm at different gene regions may be associated with a given condition, the following annotation was added for each probe: gene features (5'UTR, 1st exon, exon, intron and 3'UTR) and distance to closest transcription start site. c) We specified probes with SNPs in the C or G of the target CpG (~4%), since these are hypothesized to affect assessment of DNAm. Illustrating this, a trichotomous pattern of DNAm was observed in samples at a probe (cg09101062) with a predicted C/T SNP (rs4866831) at the target CpG; individuals assessed as unmethylated had a TT genotype (n=5), ~50% methylated had a TC genotype (n=5) and fully methylated had a CC genotype (n=2). d) In silico analysis showed 9–10% of probes hybridized to multiple genomic locations, thus we suspect these sites do not measure DNAm at one CpG. Specifically probes that cross-hybridize to sex chromosomes are of concern when male and female samples are analyzed together. As the assessment of DNAm in association with aspects of human health becomes more widespread, it is important for array data to be comprehensively analyzed in order to classify biologically relevant changes in DNAm.

3735F

Computer-aided detection of defects in Affymetrix CEL data files. *D.C. Edwards[1], A. Konkashbaev[1], A. Pluzhnikov[1], N.J. Cox[1], J.B. Tomblin[2], K.L. Mueller[2].* 1) Dept. of Medicine, The University of Chicago, Chicago, IL; 2) University of Iowa, IA.

Background: Genotyping array images such as Affymetrix CEL files can contain imaging defects that impair allele calling for all individuals in the calling batch. Manual identification of defective images can be time-consuming. We are developing an automated method to identify defective CEL files. Materials and Method: For algorithm development, we used a set of 144 CEL files from a recent GWAS conducted using the Affymetrix 6.0 platform, 36 of which contained visible defects such as dark spots, bright smudges, and bright streaks. Many of these defects were found to be more uniform than a typical region of a non-defective CEL file. To detect such defects, the variance in a square window is calculated at each pixel in the CEL file; regions in which this resulting variance is below a threshold are reported as detections, where the threshold is a fixed number k of standard deviations below the mean of the windowed variance values throughout the CEL file. The parameters of this detection algorithm, namely the window size (20 pixels) and number k (2), were chosen to minimize the number of false-positive (FP) image detections while still detecting the majority of such uniform defects. The detection algorithm was then applied to a set of 1337 CEL files from HapMap3, which was known to contain a number of such defective images, but would have been prohibitvely time-consuming to manually identify all of them. Results: For the given choices of algorithm parameters, 23 out of the 36 defective CEL images from the training GWAS were identfied, for a true-positive (TP) rate of 64%. Of the remaining 108 normal images, 36 contained false-positive detections, for a FP rate of 33%. In the HapMap3 dataset, we sampled 181 of the 832 images in which the algorithm produced detections; 54 of these were TP, and 127 FP. We also sampled 35 of the 505 images without detections; 10 of these were FN (containing at least one undetected defect), and 25 TN. The TP rate was thus 84%, and the FP rate 83%. Conclusion: It should be noted that the algorithm parameters were chosen manually, and some types of common defect (bright linear streaks) were not detectable with the variance threshold method. However, it is already capable of identifying important classes of defect, and we expect to be able to extend and improve the algorithm to address these issues.

Copyright © 2012 The American Society of Human Genetics. All rights reserved.